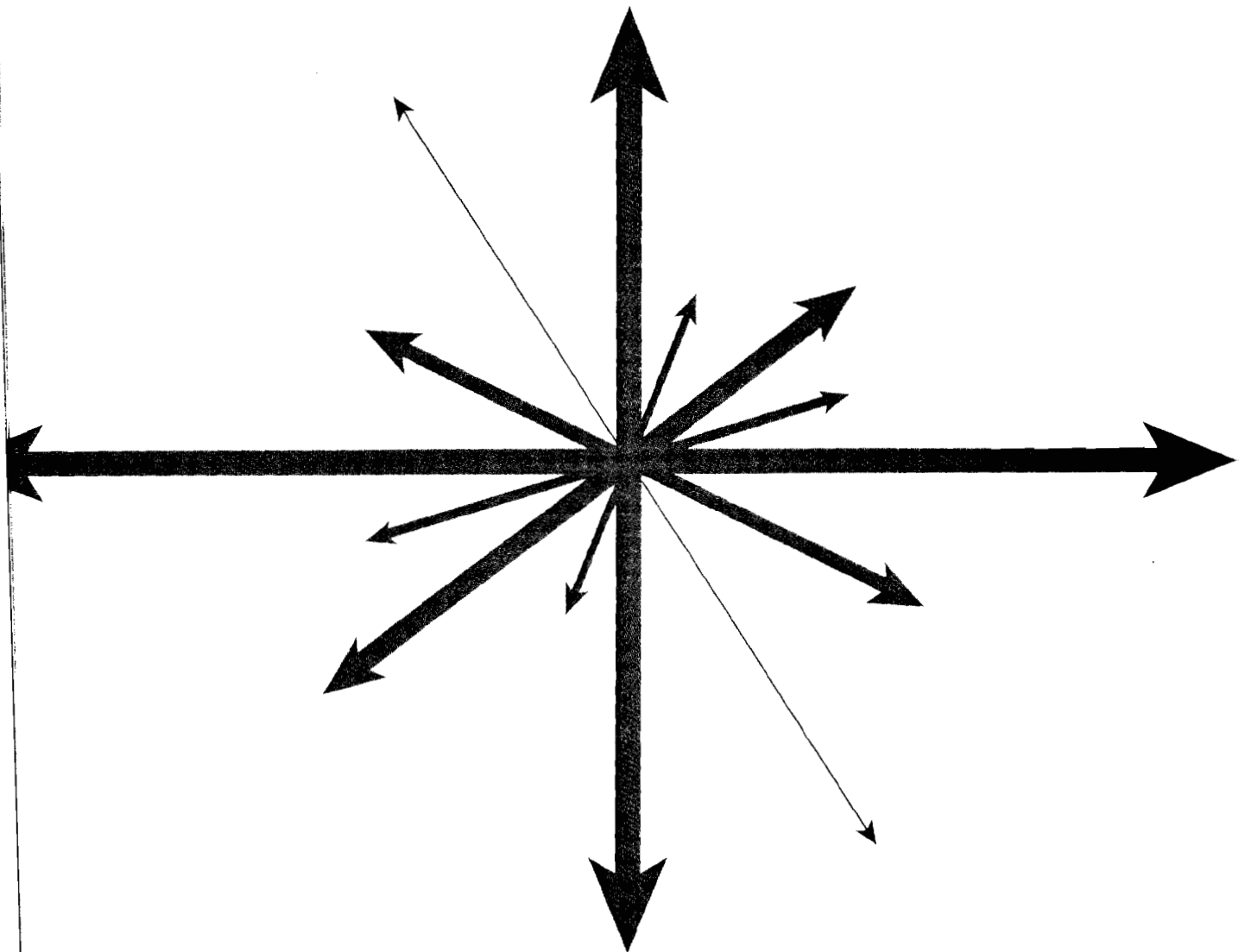


Émile QUINET

Principes d'Économie des Transports



Préface de Marcel BOITEUX

Émile QUINET

**Principes d'Économie
des Transports**

Préface de Marcel BOITEUX

Préface

Si Émile Quinet m'a demandé de préfacer ce livre, c'est sans doute en raison de quelques interventions en économie des transports que j'ai eu l'occasion de faire récemment sous les auspices du Commissariat Général du Plan, ou pour le compte de la Région Ile-de-France. Mais la lecture de son travail m'a conduit à faire retour sur mes expériences antérieures, et notamment sur la lointaine époque où j'avais eu l'occasion d'aborder le secteur des transports à propos de problèmes de tarification et de choix d'investissements.

Deux points me frappent lorsque je fais la comparaison avec ces temps.

Le premier concerne le champ privilégié que constitue la gestion des infrastructures. Les grands principes économiques n'ont pas changé. Mais les exemples de leur application se sont considérablement multipliés et ont permis de prendre une vue beaucoup plus précise des conditions de mise en œuvre desdits principes, de leur portée et de leurs limites. Je pense ici, par exemple, aux développements contenus dans le chapitre VII et relatifs à l'introduction des externalités, aux considérations sur la prise en compte de l'incertitude, à l'influence des modalités de financement sur le choix des méthodes. Disposant d'une telle accumulation d'expériences, l'économiste n'a plus d'excuse aujourd'hui à ne pas faire profiter les décideurs de ses lumières.

L'autre fait remarquable est l'extension des domaines auxquels l'approche économique s'est attachée. À côté de la gestion des infrastructures, qui constituait l'essentiel des préoccupations de l'économiste il y a vingt ou trente ans, des développements très importants ont porté sur les interactions entre les transports et l'espace, et sur les problèmes de concurrence et de réglementation. Non que cela ait été totalement ignoré dans le passé. La théorie de la rente foncière de Von Thunen date de plus d'un

réflexion et à observation. Et finalement, ces thèmes tiennent maintenant une plus grande place que la seule gestion des infrastructures. Émile Quinet leur consacre à juste titre des développements à peu près deux fois plus longs.

* *

Autre comparaison instructive, celle de l'économie des transports avec les problèmes du secteur énergétique. La confrontation des deux secteurs est intéressante dans la mesure où ceux-ci comportent de nombreuses similitudes, et une différence majeure.

Les similitudes sont visibles à tous. Il s'agit, dans les deux cas, de services de base, essentiels à l'exercice de toute activité économique ou sociale, et qui méritent le qualificatif de services publics ; les pouvoirs publics y sont largement impliqués ; la production y présente souvent des rendements croissants. Dans les deux domaines, des moyens alternatifs existent pour fournir le service – le train ou la route, l'électricité ou le gaz – et ces moyens parfois s'opposent, parfois se complètent, d'où suit qu'on navigue entre concurrence et monopole, entre compétition et coopération.

Quant aux différences, certaines sont évidentes ; ainsi les lois de Kirchoff n'ont rien de commun avec les lois de circulation des véhicules. Une autre différence est moins apparente, mais a des conséquences fondamentales pour la conduite du secteur. Pour l'énergie, les services rendus sont facturés, de sorte que la diversité des situations se trouve automatiquement agrégée en termes de recettes, recettes que l'on peut analyser ensuite par catégorie d'usagers en fonction des études poursuivies. Dans le secteur des transports au contraire, les prix marchands n'interviennent que partiellement (le billet, le péage) ou pas du tout (le réseau routier). Il faut alors affronter directement l'atomicité des acteurs et la diversité des situations dans lesquelles ils se trouvent. Et, comme le rappelle Émile Quinet il n'y a pas un usager moyen, mais autant d'usagers que d'individus et de situations dans lesquelles ils se déplacent.

Ce point est particulièrement net dans l'analyse de la demande. où l'on voit bien la difficulté qu'il y a à enfermer dans des formulations communes les situations spécifiques et leur diversité. Il apparaît aussi dans l'analyse des fonctions de coût, auxquelles Émile Quinet consacre une attention toute particulière. On y voit que la caractéristique de rendement croissant, si fondamentale pour l'analyse économique, n'est pas liée à un mode de transport particulier : selon les circonstances, on peut ou non la rencontrer dans chaque mode, ce qui pose de vrais problèmes pour l'organisation du secteur. On pourrait ajouter que cette diversité se retrouve aussi au niveau des décideurs ; la politique des transports est bien moins dans la main du ministre des Transports que celle de l'énergie ne l'est dans celle du ministre de l'Industrie. Les autorités locales interviennent de façon beau-

coup plus fondamentale et peuvent ainsi contrarier les effets d'une politique nationale ; la gestion de l'activité s'en trouve affectée, d'autant que ces différents centres de décision ont en général des intérêts divergents.

*
* *

Peut-être est-ce là une des causes de l'insatisfaction que ressent l'observateur à voir l'écart qui subsiste entre les connaissances économiques accumulées et les politiques pratiquées, insatisfaction qu'Émile Quinet met bien en évidence dans sa conclusion.

Mais il y a, me semble-t-il, d'autres raisons pour expliquer cette insatisfaction.

La première porte sans doute sur les insuffisances de l'économie appliquée : les modèles sont encore incertains, difficiles à calibrer, les estimations statistiques restent floues. Cela tient certes à des difficultés intrinsèques, liées à la diversité des situations et à la masse des données statistiques nécessaires pour en rendre compte. Mais ces difficultés, au lieu de susciter plus de rigueur et d'attention, poussent trop souvent à baisser les bras et à se contenter d'évaluations sommaires.

La seconde est plus grave et trouve moins sa source du côté des économistes que des autres acteurs dans les processus de décision. Elle réside dans une assimilation incertaine des raisonnements et des principes économiques de base. Cette situation n'est pas propre aux transports, mais le phénomène y est sans doute plus frappant qu'ailleurs en raison des nombreux pièges qu'y a déposé la nature, sources de paradoxes et d'erreurs possibles que seule une analyse économique précise peut surmonter. Qu'on songe aux problèmes de l'actualisation dans le long terme, aux conséquences des externalités, aux difficultés qu'il y a à traiter correctement de la notion de coût marginal ou de celle de rendement croissant.

*
* *

D'où l'intérêt du livre que vous allez lire.

D'abord, parce que son origine et sa vocation première sont précisément de faire face à ces insuffisances : il est pédagogique, et montre comment les principes et concepts de l'analyse économique doivent être utilisés, quelles sont les portées et limites des résultats obtenus, comment il convient de les mettre en œuvre.

Mais, avec un souci qui est bien dans la tradition des ingénieurs-économistes, cet ouvrage peut aussi contribuer à apporter un éclairage économique plus précis dans les débats de politique des transports. Celle-ci est souvent marquée par des conflits d'intérêts où les protagonistes utilisent sans doute des arguments scientifiques, mais en les biaisant parfois à leur profit. D'où l'importance de mettre davantage de rationalité dans les

... la science économique et c'est là une raison suffisante pour
... la science économique et que je salue la synthèse qu'il représente.
... les sciences de l'économie des transports connaissent
... le succès que mérite tant l'importance du sujet que
... le traitement.

Marcel BOITEUX

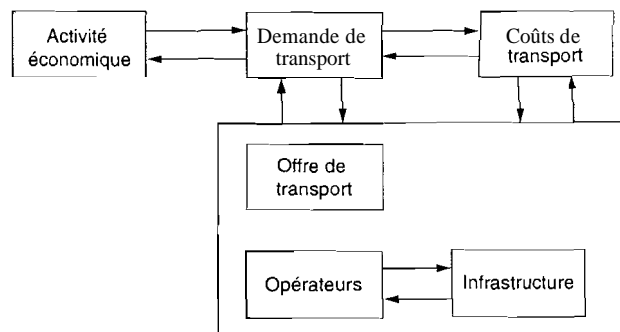
Introduction

La science économique est bâtie, comme les sciences de la nature, autour d'hypothèses dont l'outil mathématique tire les conséquences logiques. Mais elle s'applique à un domaine complexe, celui des sociétés humaines. De nombreuses causes s'y entremêlent, aucune n'est prédominante. Comme l'expérimentation est rare, il n'est en général possible ni d'isoler les causes dont on veut évaluer l'effet, ni de contrôler les autres. Les schémas et modèles explicatifs reposent donc surtout sur des réflexions théoriques, chacune correspondant à des hypothèses spécifiques. Quand on veut en appliquer les enseignements, il est difficile de savoir quelles hypothèses répondent le mieux à la situation concrète à laquelle on est confronté. On ne peut donc mettre en œuvre sans discernement les outils de l'analyse économique : il faut connaître les conditions et limites de leur validité ; il faut savoir lesquels choisir parmi tous ceux qui sont disponibles.

C'est ce que cherche à illustrer ce livre en montrant comment la théorie économique peut servir à la compréhension et à la résolution de problèmes concrets, en l'occurrence ceux qui se présentent dans un secteur à la fois varié et fondamental dans la vie économique et sociale. Varié car on y rencontre l'initiative privée et le management public, des activités marchandes et non-marchandes, des consommations intermédiaires et finales, la vie des particuliers et la gestion des entreprises ; c'est en quelque sorte un kaléidoscope de tout ce sur quoi l'analyse économique a porté son attention. Fondamental pour la vie économique et sociale dans la mesure où toutes nos activités passent par les transports, et où leur qualité et leurs caractéristiques façonnent nos vies ; qu'on songe par exemple au développement des loisirs lointains, ou à la mondialisation des échanges, ou encore à la marque imprimée par les transports à la vie urbaine.

Ces considérations générales justifient l'organisation de chacun des chapitres qui suivent : un rappel des enseignements de l'économie sur le sujet traité est suivi des compléments et adaptations que nécessite leur application aux transports. Puis sont présentées les modalités de mise en œuvre concrète, avec exemples illustrant la portée et les limites des résultats. Enfin, sont discutés à la lumière de ces résultats quelques problèmes concrets ou questions de politique des transports.

L'ordre des chapitres s'inspire de la représentation du système des transports proposée par Florian et Gaudry¹ et composé d'une série d'équilibres interactifs : l'activité économique contribue à déterminer la demande de transport, conjointement avec les coûts de transport, qui traduisent les performances du système. Ces coûts résultent de l'interaction de la demande et de l'offre ; et ils déterminent l'offre d'infrastructure en même temps que le comportement des opérateurs du secteur. Ce comportement est lui-même la résultante à la fois de la situation en matière d'infrastructures et de la demande.



Dans une première partie, « **Les transports dans l'économie** », on examine les liens entre les transports et l'économie dans son ensemble.

Le chapitre I « *Transport et activité économique* » analyse à la fois la manière dont l'activité économique détermine les flux de transport, et aussi la forme et l'intensité de la contribution des transports à la productivité et par là à la croissance économique.

Le chapitre II « *Transport et localisation* » se penche sur les interactions entre les transports et la répartition spatiale des activités. Ces interactions sont de deux sortes. La première est d'ordre mécanique : les agents économiques se localisent, toutes choses égales par ailleurs, en fonction des commodités d'accès fournies par les transports. La seconde est d'ordre straté-

gique : les commodités d'accès que rencontre un agent économique dépendent des décisions des autres agents, et, de ce fait, les localisations d'activité sont interdépendantes, à travers la structure des marchés et le degré de compétition qu'on y rencontre.

Les effets de croissance et de localisation prennent un tour particulier dans le cas des agglomérations, objets du chapitre III « *Les transports et la ville* », où est examinée la dualité entre transport et valeurs foncières, avec les conséquences qui en découlent pour la politique urbaine.

La deuxième partie, « **La demande et les coûts** », analyse d'abord dans le chapitre IV « *La demande de transport* » la modélisation des trafics, en mettant en lumière les principes et les conditions de mise en œuvre des modèles, dont les raffinements méthodologiques sont fondamentalement tributaires de la qualité des informations statistiques.

Le chapitre V « *Les coûts de transport* » est l'occasion de présenter la multiplicité des coûts engendrés par les transports, (coûts monétaires et non monétaires, coûts sociaux et privés...), les problèmes de leur mesure et les enseignements tirés de leur comparaison.

La troisième partie, « **L'organisation de l'offre** », débute par une analyse générale, dans le chapitre VI, de « *La nature des marchés et l'intervention des pouvoirs publics* », qui constitue une toile de fonds pour les chapitres ultérieurs en présentant les formes que revêtent la gestion privée et l'intervention publique pour la fourniture des transports et les raisons économiques de leur mélange.

Le chapitre VII « *Décisions publiques optimales dans les transports* » s'applique surtout à la gestion des infrastructures, tarification et investissement ; il traite des objectifs et moyens de la puissance publique visant à atteindre un optimum collectif.

Les « *Modalités et effets de la concurrence entre opérateurs* » sont l'objet du chapitre VIII. Les marchés de transports sont nombreux mais présentent entre eux de très fortes analogies ; elles sont apparues en pleine lumière à travers les conséquences des déréglementations diverses qui ont marqué les dernières décennies, et ont montré les difficultés à établir une concurrence efficace dans les transports.

Le chapitre IX est consacré au thème « *Monopole et service public dans les transports* » ; fréquents dans les transports, les monopoles sont souvent concomitants de services publics. Ils prennent des formes variées allant de la gestion directe par l'État à la dévolution privée ; et les résultats des nombreuses réformes menées dans la plupart des pays ont permis de jeter une lumière nouvelle sur les conditions de leur contrôle.

Enfin un dernier chapitre, intitulé « *De l'économie à la politique des transports* », rassemble les résultats précédents et analyse les modalités de leur utilisation pour la politique des transports.

Comme tout travail de synthèse, ce texte doit à beaucoup, d'abord à l'enseignement de Marcel Boiteux, aux échanges avec mes étudiants, aux contacts avec mes collègues du Ministère des Transports notamment Christian Brossier, et enfin aux avis et réactions de ceux qui m'ont fournis

1. Florian et Gaudry (1980), « A conceptual Analysis for the supply side in Transportation Systems », *Transportation Research*, part B, 14, 1-2.

Florian et Gaudry (1983), « Transportation system analysis : Illustration and Extensions of a conceptual Framework », *Transportation Research*, part B 17, 2.

des conseils et des pistes d'amélioration : Bernard Belloc, Bruno Bieder, Claude Charmeil, André De Palma, Jean-Raymond Fradin, Marc Gaudry, Joël Maurice, Dominique Schwartz, et mon fils Alain. Les erreurs et oublis restent à ma charge, et les opinions exprimées n'engagent que moi.

La lecture de ce texte ne nécessite que des connaissances économiques et mathématiques de base. Les rappels théoriques et les approfondissements techniques, qui peuvent être sautés sans nuire à la compréhension, sont placés dans des annexes et dans des passages en petits caractères. On s'est par ailleurs efforcé de faire en sorte que chaque chapitre puisse, dans la mesure du possible, être lu indépendamment des autres. Certains d'entre eux traitent de sujets voisins et mettent en jeu des mécanismes et concepts similaires. On ne s'étonnera donc pas des répétitions, ou des renvois lorsque celles-ci auraient été trop longues.

PREMIÈRE PARTE

Les transports dans l'économie

Chapitre I

Les transports et l'activité économique

Les transports entretiennent des liens multiples avec l'activité économique. Ils constituent d'abord un des secteurs de cette activité, et représentent de ce fait une part de la production. Ils sont aussi un moyen indispensable pour la réalisation de cette production ;leur développement suit la croissance économique, et il est utile d'apprécier la nature et l'intensité de ce lien. Enfin, selon une causalité inverse, l'amélioration de leurs performances est un facteur de croissance dont l'importance est sujette à débat. Examinons successivement ces trois aspects.

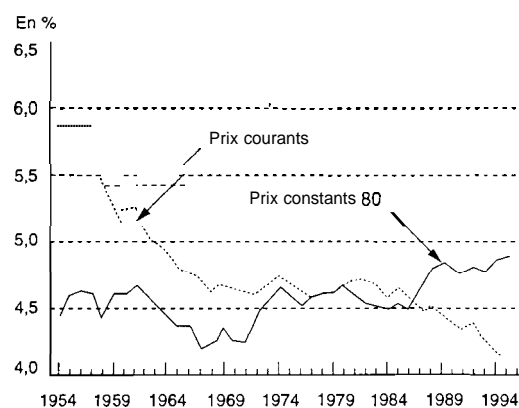
1. La place des transports dans l'activité économique

1.1. Transports et comptabilité nationale

Les comptes nationaux montrent qu'en 1995 en France, la valeur ajoutée de la branche Transports s'élevait à 262 GF, alors que le Produit Intérieur Brut (PIB) se montait pour la même année à 7 675 GF. Cette valeur ajoutée représentait donc 3,8 % du PIB, proportion assez stable puisqu'elle était de 3,7 % en 1985. La proportion de la valeur ajoutée de la branche Transport dans le PIB marchand est un peu plus forte, et aussi un peu plus fluctuante : 5,0 % en 1995 contre 4,5 % en 1985.

Mais cette constance relative ne se constate que dans l'expression des agrégats en prix constants. Le même ratio exprimé en prix courants a considérablement diminué, comme le montre le graphique joint, tiré de Vacher (1997) ; le mouvement en ciseaux que montre ce graphique résulte de la baisse relative du prix de la valeur ajoutée du transport par rapport au PIB, manifestation des gains de productivité dans la branche.

FIGURE 1-1
Ratio de la valeur ajoutée de la branche transport au PIB marchand



La branche « transports » de la comptabilité nationale ne recense que les transports publics et laisse de côté les transports privés : déplacements automobiles et deux-roues des particuliers, et transports de marchandises pour compte propre. Aussi les chiffres précédents ne représentent pas bien la place des transports dans l'activité nationale. Un autre concept fournit une image plus réelle, c'est celui de dépense courante de transport, qui représente la somme des consommations intermédiaires et finales et des transferts nets, ces derniers provenant essentiellement des administrations. La dépense courante est dite nationale lorsqu'elle se réfère aux dépenses effectuées par les nationaux, et intérieure lorsqu'elle se réfère aux dépenses effectuées sur le territoire national. La dépense nationale s'élève à 14,6 % du PIB en 1992, se décomposant selon le tableau suivant :

TABLEAU 1-1
Décomposition de la dépense courante nationale de transport en 1992 en GF (source SES 1996)

	Consommation finale		Consommation intermédiaire	Transferts courants	Total
	Ménages	APU			
Route	392,0	74,7	368,7	1,1	836,5
Fer	28,8		18,0	16,5	63,3
Transports collectifs urbains	17,5		1,0	10,2	28,6
Aérien	22,4		34,7	0,6	57,7
Maritime	1,8		25,3	0,8	27,9
Fluvial	1,3		2,2	0,0	3,6
Conduites	-		2,1		2,1
Total	463,9	74,7	451,9	29,3	1 019,8

On peut de même analyser la dépense de transport en capital, fournie par la même source :

TABLEAU 1-2
Dépense en capital des transports en 1992 (en GF)

	FBCF	Transferts	Total
Route	267,6		267,6
Fer	30,2	3,4	33,6
TCU	10,1		10,1
Aérien	13,1		13,1
Maritime	7,0	0,1	7,1
Fluvial	0,6	0,1	0,7
Conduites	0,3		0,3
Total	328,9	3,6	332,5

Cette dépense en capital ne peut pas être directement comparée à la FBCF nationale, car elle comporte les achats d'automobiles par les ménages, comptés comme consommation finale dans la comptabilité nationale.

On peut néanmoins, revenant à la définition de la branche « transports » des Comptes de la Nation, déterminer sa FBCF (96,9 GF en 1992, 82,2 GF en 1995) et la rapporter à la FBCF nationale ; on obtient alors 7 % en 1992, 6 % en 1995. Ces chiffres sont supérieurs à la part de la branche dans le PIB, manifestant le caractère capitaliste de l'activité.

1.2. L'emploi et le commerce extérieur

On peut aussi mesurer la place du secteur dans l'activité économique à travers les effectifs qu'il emploie. Ceux-ci s'élèvent à 849 000 en 1995 en ce qui concerne la branche « stricto sensu », soit 3,8 % des effectifs totaux, pourcentage en légère augmentation puisqu'il se montait à 3,7 % en 1985. Mais si l'on y ajoute à la fois les effectifs employés dans le transport privé (conducteurs de camions et bus...) ainsi que l'ensemble des activités touchées directement par les transports (construction et maintenance de matériel et d'infrastructure, activité pétrolière liée aux transports), on arrive à un chiffre total de 1,5 million environ.

Les transports sont également un élément important des échanges internationaux ; ils contribuent au solde de la balance des biens et services, dans des proportions variables d'une année à l'autre comme le montre le tableau 1-3.

TABLEAU 1-3
Solde des échanges de services de transport (FAB-FAB) en GF

	1990	1995
Ferroviaire	1,4	1,9
Routier	- 1,8	- 1,6
Fluvial	- 0,3	- 0,3
Maritime	- 1,6	- 1,9
Aérien	- 0,5	4,2
Total	- 2,8	2,2

1.3. les entreprises de transport

Ces activités productives sont réalisées par des entreprises de nombre et de taille très diverses selon les modes, comme le montre le tableau suivant :

TABLEAU 1-4
Les entreprises en 1994

Sous-secteur	Nombre d'entreprises	Effectifs	Chiffre d'affaire (GF)
Aérien	127	58 700	64,6
Maritime	161	10 700	27,0
fluvial	51	1 300	1,2
Routier marchandises	35 398	262 300	138,5
Routier voyageurs	26 656	163 000	50,9
Ferroviaire	14	186 800	54,3
Manutention entrepôt	969	27 700	18,1
Organisation tran. fret	2 049	92 700	96,8

Source : « Les transports en chiffres, 1995 » DAEI-SES-MELTT).

1.4. les trafics

Les services rendus par les transports peuvent d'abord être évalués en quantités physiques, c'est-à-dire en voyageurs x km et en tonnes x km. Les tableaux ci-dessous sont à cet égard instructifs.

Ils mettent en évidence plusieurs faits : la part élevée et croissante de la route, notamment pour les marchandises ; le développement du trafic routier de marchandises pour compte d'autrui par rapport au trafic pour compte propre, manifestation dans les transports de l'« *outsourcing* », tendance générale de l'industrie ; la croissance rapide du transport aérien ; le déclin du trafic ferroviaire malgré le développement du réseau et du trafic TGV ; enfin, la récente chute du trafic des transports publics en Ile-de-France.

TABLEAU 1-5
Les trafics et leur évolution

Voyageurs	en milliards de voyageurs-kilomètres			
	1985	1990	1995	1996
Voitures particulières	489,6	585,6	664,3	674,3
Autobus autocars	37,0	41,3	41,0	41,3
Urbain (hors Ile-de-France)	4,5	4,5	5,3	5,1
Interurbain (hors Ile-de-France)	5,6	6,0	5,2	4,9
Ile-de-France (urb. + interurb. hors RATP)	0,9	1,6	1,4	1,4
Scolaire	5,3	5,6	5,9	5,5
Personnel	4,5	3,1	2,3	2,2
Occasionnel	14,1	18,3	18,9	20,0
Autobus RATP	2,1	2,2	2,1	2,3
Transports en commun ferroviaires	71,1	73,9	64,5	69,3
SNCF	61,9	63,7	55,6	59,8
dont. TGV	8,7	14,9	21,4	24,8
réseau principal hors TGV	44,1	38,9	25,7	26,1
TER ¹	5,6	6,1	6,7	7,1
banlieue parisienne	9,1	10,0	8,5	8,9
RATP (métro + RER)	8,9	9,7	8,3	8,8
Métro de province	0,3	0,5	0,7	0,7
Transports aériens	7,4	11,4	12,7	13,8
Ensemble	605,2	712,2	782,6	798,7
	en milliards de tonnes-kilomètres			
Marchandises	1985	1990	1995	1996
Transport ferroviaire ²	54,2	49,7	46,6	48,3
<i>Transport ferroviaire hors transit</i>	48,6	44,0	39,0	39,6
National	32,4	29,0	24,6	24,9
Importations	6,3	6,2	5,9	6,2
Exportations	10,0	8,8	8,5	8,6
Transit	5,6	5,7	7,5	8,7
Transport routier	126,4	191,5	221,9	222,2
<i>Transport routier hors transit</i>	126,4	169,8	193,5	193,3
1. Pavillon français ³	114,3	153,9	175,4	174,2
1.1. National (CU > 3 t)	92,0	119,8	137,3	136,0
1.2. International (CU > 3 t)	10,0	18,3	21,3	21,5
1.3. Tr. CU < 3 t	12,3	14,5	15,5	15,6
1.4. Transit ⁴	n.d.	1,3	1,4	1,1
2. pavillon étranger	12,0	37,6	46,5	48,0
2.1. Transit pavillon étranger	n.d.	20,4	27,0	27,8
2.2. Autres pavillon étranger	12,0	17,2	19,5	20,2
Navigation intérieure ⁵	7,6	7,2	5,9	5,7
National	4,5	4,3	3,2	3,2
Importations	1,3	1,3	1,2	1,1
Exportations	1,7	1,6	1,6	1,4
Oléoducs ⁶	24,1	20,5	22,3	21,9
Ens. modes terrestres y compris transit	212,2	268,9	296,6	298,2
<i>Hors transit</i>	206,7	241,5	260,7	260,6

Marchandises répartition modale	en % des tonnes-kilomètres			
	1985	1990	1995	1996
Transport ferroviaire	23,5	18,2	15,0	15,2
Transport routier	61,1	70,3	74,2	74,2
Navigation intérieure	3,7	3,0	2,2	2,2
Oléoducs	11,7	8,5	8,6	8,4
Tous modes en Md t-km	206,7	241,5	260,7	260,6
Hors Oléoducs, VU de moins de 3 tonnes ou pavillon étranger :				
Transport ferroviaire	30,7	23,2	19,2	19,5
Transport routier	64,5	73,0	77,9	77,6
Navigation intérieure	4,8	3,8	2,9	2,8
Tous modes en Md t-km	71,1	189,2	203,4	202,9

Sources : DAEI/SES, RATP, SNCF, DGAC, SES 1996.

(1) Y compris les « Express d'Intérêt Régional » à partir de 1991.

(2) Source : SNCF. Transport total : t-km taxés sur le réseau ou en provenance d'autres réseaux ; transport par wagons non compris les wagons de particuliers et les transports routiers ni les expéditions.

(3) Source : DAEI/SES (enquête TRM). Transport intérieur des véhicules immatriculés en France, d'au moins 3 t de PTCU et de moins de 15 ans d'âge. Les chiffres publiés ici ont été rétro-polés jusqu'en 1989, pour tenir compte d'une modification de l'enquête à partir de 1990. Par ailleurs, l'enquête ayant été rénovée au 1^{er} janvier 1996, des coefficients de raccord ont été appliqués sur les résultats de l'enquête jusqu'en 1995.

(4) Il s'agit du transit des véhicules français de plus de 3 t de PTCU.

(5) Source : Voies Navigables de France. Le transit rhénan est exclu pour toutes les années.

(6) Source : Ministère de l'Industrie, Direction des Hydrocarbures. Les deux dernières années sont provisoires.

Un autre aspect doit être souligné : c'est le développement des échanges internationaux. Il se manifeste de multiples manières. Les données disponibles montrent ainsi que pour les marchandises, le trafic routier international croît à peu près deux fois plus vite que le trafic national routier. En matière ferroviaire, le trafic international s'est à peu près maintenu en tonnes x km alors que le trafic intérieur chutait.

Les chiffres de trafic exprimés en passagers ou en tonnes x km ne traduisent qu'un des aspects de l'importance économique des différents modes ; un même voyageur x kilomètre ne correspond pas à la même prestation selon qu'il représente un trajet en TGV ou en omnibus. Pour tenir compte de cette différence, on peut pondérer les voyageurs x kilomètres et les tonnes x kilomètres par leurs coûts, ou mieux par la valeur ajoutée du service de transport correspondant. L'exercice correspondant, effectué par la commission des comptes de transport de la Nation en 1996, ne conduit pas à de sensibles différences par rapport aux résultats traditionnels, comme le montrent les chiffres ci-dessous :

Indice 1995 base 1990 = 100	fer	route
en tonne* kilomètres	87	118
en valeur ajoutée	91	119

Une autre vision de l'importance respective des différents modes peut être obtenue en pondérant les trafics par la valeur des produits transportés. Le calcul correspondant a été effectué par Gerondeau (1996) et fait apparaître de fortes différences par rapport à la pondération usuelle en tonne x km :

TABLEAU 1-6
Poids économique relatif des différents modes,
selon la valeur ajoutée des marchandises transportées

	En tonnes x km	En valeur ajoutée
Chemin de fer	22,4	4,5
Routes	74,5	95,3
Voies navigables	3,1	0,2

Ceci résulte de ce que les modes présentent des spécificités quant à la nature des produits transportés : la route transporte davantage de produits à forte valeur ajoutée que le rail et la voie d'eau. Ce point est confirmé par le tableau suivant exprimé en tonnes x km :

TABLEAU 1-7
Structure des trafics intérieurs par groupe de produit
en milliard de tonnes x km en 1994

Modes de transports	Matières premières	Produits énergétiques ¹	Produits agroalimentaires	Produits manufacturés
Voie navigable	3,2	1,1	1,2	0,1
Fer	19,4	4,7	9,5	13,5
Route	40,7	6,1	41,8	33,5

Source : SES 1996.

La spécificité des modes se traduit aussi dans leur contribution à l'acheminement du commerce international, fournie par le tableau suivant :

TABLEAU 1-8
Commerce extérieur par mode (1994)
en millions de tonnes

	Union Européenne		Pays Tiers	
	Introductions	Expéditions	Importations	Exportations
Fer	7,3	14,2	2,5	3,1
Route	60,7	59,4	14,0	11,2
Voie navigable	4,0	11,4	4,9	1,7
Mer	32,7	26,7	135,6	29,3
Air	0,0	0,0	0,2	0,3

Source : SES 1996.

La spécificité des modes se marque enfin à travers les distances moyennes de transport, fort différentes d'un mode à l'autre :

TABLEAU 1-9
Distances moyennes de transport intérieur de marchandises en km

	1975	1990
Route	56	78
Fer	296	359
Voie navigable	126	115

Source : SES 1996.

1.5. Les prix

Les prix des transports sont mal connus, et les statistiques sont loin de couvrir tous les modes. On a déjà vu, en 1.1, à travers la comparaison de différentes statistiques, que les prix de la valeur ajoutée de la branche transport augmentaient moins vite que le prix du PIB. L'écart moyen de 1 % par an – particulièrement marqué de 1954 à 1974 et depuis 1987 – conduit à une différence finale de 60 points d'indice sur une base 100 en 1954.

Le tableau 1-10 joint fournit les indices de prix, non pas de la valeur ajoutée, mais de la production effective. Il permet de différencier les évolutions selon les modes. On voit que les transports maritimes et aériens ont connu depuis 1980 une baisse des prix relatifs particulièrement forte, les transports routiers aussi, mais à un moindre degré, et ceci à l'inverse des transports urbains et interurbains routiers de voyageurs.

TABLEAU 1-10
Indices des prix des branches

	1980	1985	1990	1995	1996
Transport ferroviaire	100	155	169	179	183
Transport routier de marchandises	100	166	165	169	171
Autres transports terrestres (essentiellement routiers de voyageurs)	100	165	195	223	229
Navigation intérieure	100	163	174	188	187
Transport maritime	100	172	158	139	134
Transport aérien	100	162	146	136	128
Auxiliaires, ports, aéroports	100	168	183	194	194
Ensemble de la branche	100	165	170	175	175

Source : Exploitation de bases de données du SES.

Le tableau 1-11 fournit l'évolution des prix des transports utilisés par les ménages depuis 1980 : l'automobile coûte de moins en moins cher à l'achat, mais de plus en plus cher à l'utilisation ; les transports collectifs voient leur prix augmenter nettement plus vite que la moyenne, et les transports aériens nettement moins vite.

TABLEAU 1-11
Évolution des prix de certains postes de dépenses transport des ménages

	1970	1975	1980	1985	1990	1995	1996
Achats d'automobiles	36	60	100	154	179	192	192
Caravanes cycles et motocycles	43	63	100	150	190	216	219
Achats de véhicules	37	60	100	153	179	194	194
Réparations et pièces détachées	36	56	100	165	225	272	282
Pneumatiques chambre à air	37	59	100	155	163	190	191
Pneus accessoires hais de réparation	36	56	100	165	221	267	276
Carburants et lubrifiants	34	53	100	168	162	180	193
Péages, parkings, locations, auto-écoles	40	59	100	163	203	239	248
Assurances transports	39	64	100	251	355	277	271
Autres dépenses d'utilisation	40	62	100	203	254	252	256
Dépenses d'utilisation véhicules	35	55	100	170	195	225	235
Transports ferroviaires de voyageurs	46	58	100	160	188	221	225
Transports urbains de voyageurs	40	56	100	168	196	241	251
Transports routiers, taxis, déménagement	33	54	100	165	196	236	242
Autres transports terrestres	36	55	100	167	196	238	245
Transports maritimes et aériens	41	62	100	170	188	192	185
Achats services de transport collectif	40	57	100	165	191	221	221
Total transports	36	57	100	165	190	216	222
Consommation finale des ménages	39	61	100	157	182	204	208

Source : Exploitation de bases de données du SES.

Des mouvements financiers importants se produisent entre le secteur des transports et les budgets publics. Les transports versent aux budgets publics de nombreuses contributions spécifiques, en particulier la taxe sur les hydrocarbures, qui s'élevait en 1994 à 131 GF sur un total de contributions de 168 GF, en majorité payées par les trafics routiers. Dans l'autre sens, les transports reçoivent des budgets publics une somme qui s'élevait en 1994 à 198,5 GF, composée essentiellement des fonds consacrés par l'État et les collectivités locales aux infrastructures routières, aux transports publics urbains et à la SNCF Orfénil (1997) # # # # #

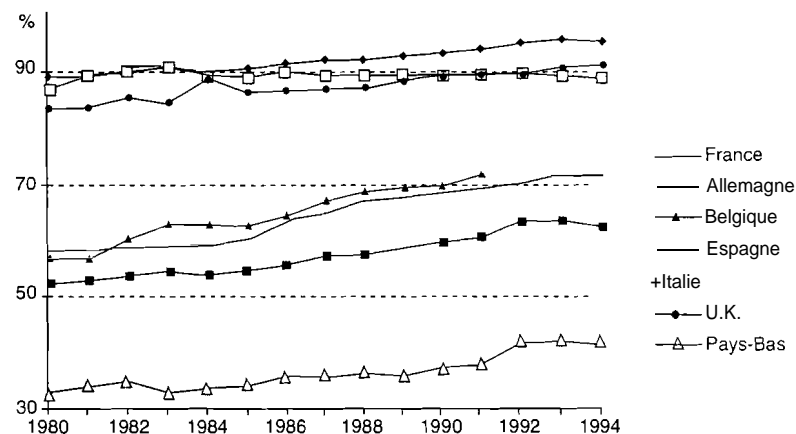
recettes spécifiques routières et dépenses publiques pour la route pour 1991 (en milliards de francs) :

Recettes	Milliards de francs	Dépenses	Milliards de francs
Liées à l'acquisition	11,4	Investissement	44,8
Liées à la possession	26,7	Maintenance	33,1
Liées à l'usage	118,9	Exploitation	27,5
Total	157	Total	105,4

1.6. Comparaisons internationales

La situation française qui vient d'être décrite n'est pas très différente de celle que l'on rencontre dans les autres pays européens, ne serait-ce qu'en raison de la convergence des économies. Les données structurelles sont assez similaires, par exemple la part de la dépense de transport dans le PIB. Seule diffère assez sensiblement d'un pays à l'autre la proportion d'utilisation des différents modes. Ainsi, en matière de marchandises, les répartitions sont les suivantes :

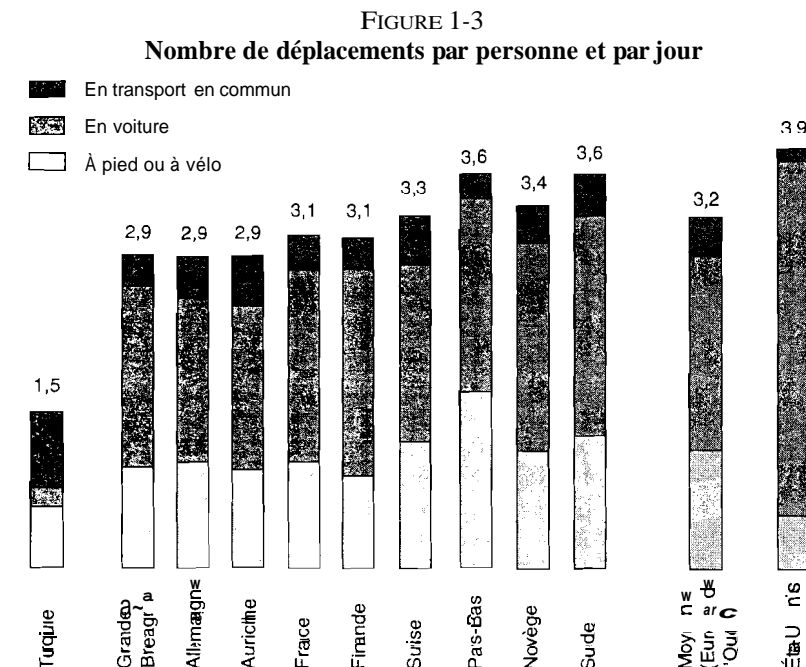
FIGURE 1-2
Part de la route dans le transport de marchandises



Source : SES 1996.

Ces différences s'expliquent par des différences dans les qualités de l'offre des différents modes et par des données géographiques telles que le relief ou l'orographie. De même, en ce qui concerne les transports de personnes, comme le nombre de déplacements par jour ainsi que

leur répartition selon les modes utilisés présentent en Europe de l'Ouest une grande homogénéité ; en revanche, les différences avec d'autres pays sont marquées, comme le note Gerondeau (1996) d'où est tiré le tableau suivant :



Source : Gerondeau, 1996.

2. L'évolution des transports expliquée par la croissance économique

Dans la section précédente on a mesuré la place des transports dans l'activité économique, et on a vu à travers quelques comparaisons internationales que des pays présentant de grandes similarités comme ceux de l'Europe de l'Ouest présentent aussi des similarités fortes pour ce qui est de leurs transports. Cela donne à penser qu'il y a de fortes relations entre économie et transports, et invite à en démonter les mécanismes.

Ceux-ci jouent dans les deux sens. Tout d'abord, l'activité économique nécessite un certain volume de transport. Selon cette causalité, le volume et la nature des transports sont expliqués par l'activité économique et par sa structure ; c'est ce à quoi on va s'attacher dans la présente section. La section suivante sera consacrée à l'analyse de la causalité inverse, selon laquelle une amélioration des transports suscite le développement écono-

Dans une première sous-section quelques résultats d'une analyse sur longue période seront présentés ; les deux sous-sections suivantes seront consacrées aux modélisations actuelles mises en œuvre respectivement pour les transports de marchandises et de voyageurs.

2.1. L'analyse sur longue période

L'analyse historique sur longue période fournit sur ces sujets des enseignements intéressants, dans la mesure où une perspective longue permet de mieux faire apparaître certaines causalités, et en particulier de distinguer les situations où les transports ont accompagné le développement économique et celles où leur progrès l'a permis.

Il serait intéressant et instructif d'étudier les transports et leur évolution depuis l'Antiquité jusqu'à l'Ancien Régime. Mais les données les plus fiables sont plus récentes. Elles ont été rassemblées par J.C. Toutain (1967) en un travail inestimable, qui accompagne heureusement l'histoire des transports depuis le début du 19^e siècle et qu'on résumera pour ce qui concerne les marchandises.

La révolution industrielle amena des bouleversements profonds et rapides dans la structure antérieure. D'abord elle a accru les besoins de transports, notamment pour les produits lourds en provenance des mines et de la sidérurgie ; mais elle fournit aussi le moyen technique d'assurer ce transport : le rail.

TABLEAU 1-12

Évolution du transport de marchandises au 19^e et au début du 20^e siècle

Années	Marchandises transportées (milliards de t x km)				
	Route	Canaux	Fer	Cabotage	Total
1830	2,0	0,7	–	1,1	3,8
1831-1844	2,3	1,1	0,06	1,2	4,7
1845-1854	2,6	1,7	0,46	1,3	6,1
1855-1864	2,7	2,0	3,0	0,9	8,6
1865-1874	2,8	1,9	6,3	0,8	11,8
1875-1884	2,6	2,1	9,4	0,8	14,9
1885-1894	2,7	3,3	10,9	1,4	18,3
1895-1904	2,8	4,6	14,9	1,9	24,2
1905-1913	2,9	5,5	21,0	1,9	31,3
1921-1924	4,3	4,4	31,3	0,8	40,8
1925-1934	7,2	6,9	36,6	1,4	52,1
1935-1930	11,7	8,1	28,7	3,0	51,5
1956-1960	17,8	9,7	53,4	–	80,9
1961-1965	28,2	11,8	62,0	–	102,0
1966	40,2	12,6	64,1	–	116,9

Source : Toutain, 1967.

En France la première ligne de chemins de fer vit le jour en 1828 entre St-Étienne et Andrezieux. Elle était privée et devait assurer le transport du minerai de charbon du lieu d'extraction au lieu de transformation. Mais très vite le territoire se couvrit de lignes et le transport ferroviaire supplanta le transport routier. La domination du rail dura jusque vers 1930, date où le transport par camion redonna un essor à la route, comme l'illustre le tableau ci-dessus.

L'analyse économétrique montre la corrélation entre le trafic total et l'indice de la production nationale. Cependant la liaison entre les deux n'est pas stable et connaît des déformations. Sa qualité se détériore vers la fin de la période recensée. L'interprétation généralement donnée de ce phénomène est que, notamment en période de crise ou d'accroissement de la concurrence, les chargeurs révisent leur politique logistique, ajustent les stocks, modifient les sources et les conditions d'approvisionnement, recherchent de nouvelles clientèles, toutes actions qui découplent la liaison entre le volume de transport et le volume de la production.

2.2. Activité économique et transport de marchandises

La modélisation des transports de marchandises s'exprime en général par des relations multiplicatives dont les coefficients représentent des élasticités¹ :

$$y = kx_1^\alpha x_2^\beta x_3^\gamma \dots$$

Les relations les plus usuelles expliquent soit le trafic total tous modes confondus, soit le trafic de chaque mode, exprimés en tonnes x kilomètres. Les variables explicatives sont la plupart du temps des variables représentatives de l'activité, le PNB ou la production industrielle par exemple ; plus rarement sont introduits les prix des différents modes.

Ainsi le tableau ci-dessous, tiré du rapport « Transport 2010 » (1992), fournit les élasticités utilisées par le Ministère des Transports. On constate que les élasticités aux prix n'y figurent guère Le SES (Ministère des Transports), dans le cadre des travaux d'élaboration des Schémas Directeurs d'infrastructures, a utilisé les élasticités-prix suivantes :

- élasticité du trafic route au prix routier : – 0,4,
- élasticité du trafic fer au prix routier : + 0,6.

Ajoutons que l'OEST a distingué des élasticités de court terme et de long terme du trafic marchandises à la croissance économique (Girault, Blain, Meyer, 1995, Meyer, 1997). En utilisant la méthode de la cointégration, les auteurs trouvent que l'élasticité de court terme (réactions de court terme à une modification de l'activité économique) est assez élevée, de l'ordre de 2. L'élasticité de long terme est assez nettement inférieure, de l'ordre de la moitié de la précédente. En outre, ces élasticités ont tendance à augmenter au cours des années récentes.

1. α est l'élasticité de y par rapport à la variable x_1 .

TABLEAU 1-13
Déterminants de la demande de transport

Mode/entreprise	Unité	Source	Déterminants	Élasticité	Tendance 1970-1989 (% annuel)
Route	tonnes-km	MINIDMS	Production industrielle	1,47	+ 1,6 %
		D.P.	PIB marchand	1,87	
		(1966-1990)	tendance- 3,1 %		
OEST	tonnes-km	(1985-1990)	Prod. industrielle	1,06	
			Prix TRM	- 0,64	
Fer	tonnes-km	SNCF	Demande de biens industriels	0,67	- 1,6 %
		(1962-1980)	tendance- 3,3 %		
		D.P.	PIB marchand	1,18	
OEST	tonnes-km	(1966-1990)	tendance- 4,6 %	1,40	
			Prod. industrielle	tendance- 3,6 %	
Voies navigables	tonnes-km	D.P.	PIB marchand	1,49	- 3,9 %
		(1966-1990)	tendance- 7,5 %		
Ports maritimes métropole	tonnes (débarquées) + embarquées)	D.P.			+ 0,6 %
Groupe Air France	tonnes-km				+ 10,6 %
Air France	tonnes-km	MINIDMS	Trafic passagers Air France	0,88	
		(1970-1981)	tendance 7,0 %		
ADP	tonnes	D.P.	PIB marchand	+ 2,76	+ 6,8 %

Source : « Transports 2010 », 1992.

Van de Voorde et Mersmann (1997), utilisant également un modèle de cointégration, trouvent dans le cas de la Belgique les élasticités suivantes, concernant l'effet de la production industrielle sur les trafics :

	Court terme	Long terme
Route	0,89	2,38
Rail	1,45	0,45
Voie navigable	1,44	- 0,34

À côté de ces analyses sur séries chronologiques il est intéressant de présenter une étude statistique en coupe instantanée effectuée par la Banque Mondiale (Bennathan, Fraser, Thompson, 1992) sur les pays développés. Les auteurs ne fournissent pas d'élasticité-prix, mais uniquement des élasticités PNB :

0,67 pour le trafic total (route + fer)

1,25 pour la route

1,14 pour le fer

Quinet (1994) fournit des fourchettes pour les élasticités aux prix, directes et croisées, des trafics fer et route à longue distance, sur les segments où la concurrence est la plus forte :

Trafic	Prix	
	Fer	Route
Fer	- 1	1,3
Route	0,5 à 0,7	- 0,7 à - 0,9

Les relations économétriques sont ajustées sur les constatations passées et ne prennent pas en compte les facteurs susceptibles de créer des inflexions ou ruptures par rapport aux tendances passées. Ces inflexions ou ruptures peuvent être évaluées à travers des méthodes de prospective, par exemple par des méthodes *Delphi*¹ de consultation d'experts. C'est à une enquête de ce type qu'à procédé l'OEST en 1996, dans le cadre de l'élaboration des Schémas Directeurs d'infrastructures. Les facteurs cités concernent les mesures de politique des transports, une possible indifférence croissante aux coûts de transport, un attrait accru pour la flexibilité, le développement de l'internationalisation, les incertitudes sur la construction européenne, une sensibilité accrue à l'environnement, une crise possible du financement public.

2.3. Activité économique et transport de personnes

Le rapport « Transports 2010 » (1992) avait, comme pour les marchandises, établi un recensement des études disponibles en France sur le sujet, et en avait résumé les résultats dans le tableau 1-14.

Plus récemment une analyse effectuée par Blain et N'Guyen (1994), complétant et actualisant des travaux de Taroux et Buchmuller (1980) et de Madre et Lambert (1989), aboutit aux résultats du tableau 1-15.

1. On demande à un panel d'experts leur opinion sur un sujet donné, en les interrogeant indépendamment. On les interroge ensuite de nouveau en leur communiquant les résultats du premier tour d'enquête.

TABLEAU 1-14
Déterminants de la demande de transports voyageurs

Modelentreprise	Unité	Source	Déterminants	Élasticité	Tendance 1970-1989 (% annuel)
Circulation routière réseau national (parcours)	véhicules-km	OEST (1971-1990)	Consom. des ménages Prix carburants	+ 1,21 - 0,10	+ 4,8 %
Circulation routière des ménages (véhicules particuliers)	véhicules-km	OEST (1975-1989)	PIB Prix carburants	+ 1,21 - 0,29	
Autocars	voyageurs	MINIDMS	Tendance	+ 1,4 %	+ 0,7 %
SNCF réseau principal (ensemble du trafic à l'exception du trafic banlieue)	passagers-km	MINIDMS (1967-1980)	Revenu des ménages Prix	+ 0,57 - 0,54	+ 2,4 %
		OEST (1970-1990)	Consom. des ménages Prix SNCF Prix carburants	+ 0,73 - 0,69 + 0,13	
SNCF réseau express	passagers-km	SNCF	Consom. des ménages 1 ^{re} classe (RE) 2 ^e classe (RE) Pondérée (RE) Prix 1 ^{re} classe 2 ^e classe Pondérée	+ 1,3 + 0,6 + 0,7 - 0,8 - 0,4 - 0,4	
RATP	voyageurs	MINIDMS	Prix	- 0,10	+ 2,2 %
RATP + SNCF banlieue	voyageurs				+ 2,1 %
Ports maritimes métropole	passagers (embarqués et débarqués)	D.P.	Consom. des ménages	+ 2,46	+ 6,6 %
Aérien France-Étranger	passagers	D.P.	Consom. des ménages	+ 1,94	+ 5,3 %
Aérien Europe-Amérique Nord (trafic passagers US) (trafic passagers non-US)	passagers	Air France	PIB US Prix relatif (Europe/US) PIB EUR Prix relatif (US/Europe)	+ 1,3 - 0,8 + 1,2 - 1,2	+ 4,9 %
Aérien France-Europe	passagers	Air France	PIB Recette unitaire	+ 1,92 - 0,35	
Groupe Air France	passagers-km	D.P.	Consom. des ménages	+ 1,89	+ 6,3 %
Air France	passagers-km	MINIDMS (1966-1981)	Revenu des ménages	+ 1,42	
Air Inter	passagers-km	MINIDMS (1966-1981)	Revenus des ménages Prix	+ 1,37 - 0,69	+ 9,7 %
		D.P. (1969-1988)	Consom. des ménages Prix	+ 2,9 - 0,7	
	passagers	OEST	Consom. des ménages Prix Prix SNCF	+ 2,22 - 1,03 + 0,57	
Partage SNCF grande ligne sur Air Inter	passagers-km	D.P.	Consom. des ménages Prix relatif Offre TGC (% rames)	- 0,10 - 0,06 + 1,37	
ADP	passagers mouvements	D.P.	Consom. des ménages	+ 2,20	+ 8,1 % + 2,0 %

TABLEAU 1-15
Relations explicatives du trafic voyageur

Variable à expliquer	Variables explicatives et élasticité
Trafic sur le réseau routier national	<ul style="list-style-type: none"> • PIB 1,20 • Rapport : prix route/prix train - 0,22 • Parc automobile 0,25
Trafic sur autoroutes concédées	<ul style="list-style-type: none"> • Trafic sur autoroutes concédées de l'année précédente 0,25 • Produit intérieur brut marchand 0,92 • Prix du carburant - 0,47 • Coût du péage - 0,96 • Longueur du réseau autoroutier 0,65
Trafic aérien intérieur	<ul style="list-style-type: none"> • Produit intérieur brut marchand 0,75 • Produit moyen tarifaire aérien - 0,53 • Trafic TGV en voyageurs x km - 0,017 • Nombre de siège-kilomètre offert 0,56
Trafic ferroviaire sur le réseau principal	<ul style="list-style-type: none"> • Consommation finale des ménages 0,43 • Produit moyen tarifaire de la SNCF - 0,70 • Prix des carburants + 0,20 • Vitesse moyenne des trains + 1,08

Comme on le voit, l'explication du trafic des autoroutes concédées fait apparaître un effet de retard, concrétisé par la variable endogène retardée comme variable explicative. Sa présence manifeste des adaptations progressives au changement des variables exogènes et permet de calculer la progressivité à la cause. Ainsi l'élasticité à court terme au prix du péage est de - 0,96, et l'élasticité à long terme est de :

$$- 0,96 / (1 - 0,25) = - 1,28$$

Madre (1997) effectue une analyse fine des résultats des études économétriques existants en France, et conclut que les croissances constatées dans le passé se déformeront probablement dans le futur, dans le sens d'un tassement, mais qu'elles dépendront étroitement de facteurs tels que la politique d'urbanisme ou la politique tarifaire.

Donnons également quelques indications sur les résultats d'études étrangères. Ainsi, Goodwin (1992) a analysé les élasticités au prix du pétrole, qui prennent les valeurs suivantes :

Élasticité	À court terme	À long terme
Du trafic	- 0,2	- 0,3

Johansson et Schipper (1997) ont déterminé des élasticités dans 12 pays de l'OCDE à partir de données portant sur les 20 dernières années. Ils distinguent les élasticités du parc, de la consommation unitaire en carburants, du kilométrage annuel, de la demande de carburant et du volume de trafic automobile. Les élasticités sont calculées par rapport au revenu et au prix du carburant.

Élasticité	Au prix du carburant	Au revenu
Du parc automobile	-0,1	1,0
De la consommation unitaire	-0,4	0,0
Du kilométrage annuel	-0,2	0,2
De la consommation totale de carburant	-0,7	1,2
Du trafic automobile total	-0,3	1,2

Selvanathan et Selvanathan (1994) procèdent à une intéressante étude sur les élasticités directes et croisées du transport individuel, du transport public et des communications (poste, téléphone...) à partir de l'estimation d'une ensemble complet de fonctions de demande, utilisant les particularités de ces fonctions déduites d'une fonction d'utilité, et dans lequel les 4 biens sont le transport public, le transport privé, les communications et l'ensemble des autres biens. L'étude faite sur le Royaume-Uni et sur l'Australie fait apparaître les élasticités globales suivantes :

	Élasticité au revenu		Élasticité-prix directe	
	UK	Australie	UK	Australie
Transport privé	2,0	2,3	-0,5	-0,5
Transport public	1,8	0,8	-0,4	-0,7
Communications	1,2	0,5	-0,1	-0,6

Les élasticités croisées sont positives et de l'ordre de 0,1 sauf pour les élasticités du transport public et des communications au transport privés qui sont de l'ordre de 0,3.

Les élasticités présentées jusqu'ici étaient issues de corrélations portant sur des données agrégées constituées soit par des séries chronologiques, soit par des coupes instantanées comparant des pays ou des régions. On peut aussi calculer des élasticités à partir des modèles de trafic qui déterminent le trafic sur chacun des arcs d'un réseau géographique (ces modèles sont l'objet du chapitre IV) en faisant varier un des paramètres, par exemple le prix d'un mode (modifié par exemple proportionnellement sur tous les arcs du réseau) ou le prix du carburant. Les élasticités qu'on en déduit sont évidemment très dépendantes des conditions

locales de concurrence entre les modes ; les valeurs qu'on en tire sur un réseau donné sont différentes d'un réseau à l'autre et constituent elles-mêmes des moyennes de situations diversifiées à l'intérieur du réseau.

Les élasticités-prix directes des transports publics urbains ont été recensées par Kechi (1996) pour la France ; Goodwin (1992) ; Oum, Waters II et Yong (1992) ont effectué un tour d'horizon plus général. Il en ressort que ces élasticités sont en moyenne de -0,3, et en général dans une plage de -0,1 à -0,6 à court terme, et qu'elles atteignent le double à long terme.

Oum *et al.* (1992) recensent aussi les élasticités-prix d'autres modes. Pour le transport aérien, ils trouvent des valeurs dispersées allant en gros de -0,5 à -2 ; la plage relative aux élasticités du transport ferroviaire interurbain est à la fois un peu plus resserrée et plus basse, de -0,5 à -1,5. Ces valeurs sont issues de modèles agrégés ; les valeurs issues de modèles désagrégés sont plus faibles.

Une autre étude fournit, pour le trafic interurbain, des valeurs des élasticités croisées, calculées par Morellet à partir du modèle Matisse, et cités dans Quinet (1994).

TABLEAU 1-16
Élasticités directes et croisées
dans le trafic interrégional des voyageurs tirées du modèle Matisse

	Prix du			
	Carburant	Rail (< 150 km/h)	Rail (> 150 km/h)	Air
Trajet 100-300 km				
Route ordinaire	-1,2	0,1	0,0	0,0
Autoroute	-1,3	0,4	0,0	0,0
Rail (< 150 km/h)	1,4	-1,3	0,0	0,0
Rail (> 150 km/h)	0,0	0,0	0,0	0,0
Air	0,3	0,4	0,0	-
Trajet supérieur à 700 km				
Route ordinaire	-0,4	0,05	-	0,05
Autoroute	-0,6	0,1	0,05	0,1
Rail (< 150 km/h)	0,2	-1,0	-1,4	0,4
Rail (> 150 km/h)	0,2	0,05	0,3	0,4
Air	0,1	0,05	0,3	-1,4

Ces relations traduisent les évolutions passées. La possibilité d'inflexions et de ruptures a été étudiée qualitativement par interviews d'experts, comme pour les transports de marchandises. Les facteurs d'incertitude jugés les plus importants sont apparus être, outre les décisions de politique des transports, les modifications démographiques, la concentration des localisations, la réorganisation des temps de travail, l'internatio-

nalisation, le développement du tourisme, une possible crise de la construction européenne accompagnée d'une tendance au repli national, une sensibilité accrue aux problèmes d'environnement, et la crise des finances publiques.

2.4. les résultats des prévisions

On rapportera ici les résultats de l'utilisation de ces modèles pour la prévision des trafics en France à l'horizon 2015, tirés de Girault et Kail (1997) et de Girault et Blain (1997) : ces prévisions ont été faites dans le cadre de l'élaboration des Schémas Directeurs d'infrastructures. À partir d'un scénario médian sont établies des variantes portant sur le rythme de la croissance économique, sur le développement des réseaux d'infrastructures, et sur la politique des transports, volontariste ou libérale. Les résultats sont présentés dans les tableaux suivants :

TABLEAU 1-17
Transports de voyageurs

a) Scénario médian

	Milliards de voy.-km		Croissance annuelle moyenne	
	1992	1970-1992	1992-2015	2015
Routes	217,4	4,6 %	2,5 %	387,2
dont autoroutes concédées	68,2	11,5 %	3,8 %	160,3
Ferroviaire	52,9	2,0 %	1,8 %	80,4
Aérien	8,8	9,1 %	3,5 %	19,5
Total	279,2	4,0 %	2,4 %	487,0

b) Hypothèses alternatives

	Milliards de voy.-km	Écart en milliards de voy.-km par rapport au scénario médian				
		Hypothèses alternatives				
		Scénario médian	Croissance Faible	Croissance Forte	Politique de régulation Libérale	Politique de régulation Volontariste
Routier	387,2	- 27	21	9,0	- 9,5	5,1
Ferroviaire	80,4	- 3,2	2,4	3,3	- 3,2	6,4
Fluvial	19,5	- 2,8	2,3	3,2	- 0,7	- 2,1
Total	487,0	- 33	26	15,5	- 13,5	9,4

TABLEAU 1-18
Transports de marchandises

a) Scénario médian

	Milliards de t-km	Croissance annuelle moyenne		Milliards de t-km
	1992	1971-1992	1992-2015	2015
I. - Transport intérieur tous modes	216	2,0 %	2,1 %	346
II. - Évolution des transports				
Transport routier	161	3,5 %	2,6 %	288
- dont transport national	102	1,9 %	1,2 %	135
- dont échanges extérieurs français	34	6,6 %	3,8 %	80
- dont transport de transit	25	16,6 %	4,8 %	74
Transport ferroviaire	48,2	- 1,3 %	0,5 %	53,5
Transport fluvial	6,9	- 2,8 %	- 2,0 %	4,4

b) Hypothèses alternatives

	Milliards de t-km	Écart en milliards de t-km par rapport au scénario médian				
		Hypothèses alternatives				
		Scénario médian	Croissance faible	Croissance forte	Politique de régulation libérale	Politique de régulation volontariste
Routier	288,0	- 41,0	63,0	21,0	- 25,0	6,5
Ferroviaire	53,5	- 7,9	9,1	- 5,5	6,1	- 4,2
Fluvial	4,4	- 1,7	1,8	- 1,0	1,1	0,0
Total	346,0	- 50,0	74,0	14,5	- 18,0	2,3

3. Les transports, facteur de croissance économique

Dans cette section, on prend un point de vue inverse de celui adopté dans la section précédente, en s'attachant à l'effet d'une amélioration des transports sur l'activité économique. Cet effet peut prendre plusieurs formes. La plus évidente est celle qui résulte de la réalisation des infrastructures et de l'activité que suscite cette réalisation. La seconde réside dans l'abaissement des coûts de production, conséquence mécanique de l'abaissement du coût des transports, qui entraîne un développement des marchés et de l'activité économique. La troisième, qui constitue depuis quelques années un champ de recherche en développement rapide, passe

par les effets d'externalités que peut entraîner un meilleur réseau de transports. Nous examinerons successivement ces différents points.

3.1. L'effet d'entraînement dû à la réalisation des infrastructures

La plus immédiate conséquence macro-économique de la réalisation des infrastructures résulte de cette réalisation même. Une attention particulière est portée à cet aspect durant les périodes de basse conjoncture. On s'intéresse alors aux effets des travaux de construction ou d'entretien envisagés sur l'emploi ou PNB. Un travail pionnier dans ce domaine est celui de Charmeil (1968), qui a posé les bases méthodologiques de ce type d'analyse et calculé, pour la période 1965-1970, l'effet sur le PNB de la réalisation des investissements routiers, qu'il trouve égal à 0,25 %.

Les méthodes actuellement en vigueur passent par l'utilisation de modèles macro-économiques sommaires qui se centrent sur les effets de multiplicateur et d'accélérateur et sur les échanges interindustriels, la partie monétaire étant traitée assez sommairement. Taroux (1995) a effectué des simulations de ce type en utilisant le modèle IMPACT, qui ne comporte pas de volet monétaire et ignore les effets multiplicateurs du revenu. Les calculs ont porté sur l'emploi. Les résultats en sont que 1 milliard de F dépensés en infrastructures en 1990 suscitaient environ 2 700 emplois-ans dont le modèle fournit une répartition en branches d'activité, selon une décomposition en 40 branches. Le modèle « DÉPENSE PUBLIQUE » (1997) ne considère que 4 secteurs définis par leur exposition à la concurrence et au chômage. De type keynésien, il intègre les effets de distribution de revenu. Une injection de 1GF en 1997 créé environ 2 300 emplois-ans (le double avec financement par déficit budgétaire). Ces résultats ne prennent pas en compte les décalages entre décisions et apparition de l'effet, liés à la durée des travaux qui, pour des grands projets, s'élèvent souvent à plus de cinq ans. Ils ignorent aussi les modalités de financement de la mesure testée, qui peuvent être diverses : augmentation du déficit budgétaire, impôt, emprunt. Il faudrait des modèles plus complexes pour en intégrer les effets, qui diffèrent selon la modalité du financement et qui vont toujours dans le sens d'une diminution des chiffres précédemment cités.,

3.2. L'effet d'abaissement des coûts de transport

Cet effet et ses répercussions sur l'ensemble de l'économie sont faciles à comprendre : en ce qui concerne les biens, leurs coûts de production et leur prix sont réduits, les quantités échangées augmentent ; pour les déplacements de personnes, la diminution de leur coût change les structures de consommation et de déplacement ainsi que les revenus distribués, et un nouvel équilibre s'établit progressivement. Deux voies ont été utilisées pour explorer ces transformations : les modèles macro-économiques dyna-

Burzau et Cipriani (1986) ont étudié l'impact macro-économique du TGV Paris Sud-Est, à travers le modèle mini-D.M.S.-Transport¹, dans le cadre d'un financement par l'emprunt de cette infrastructure, et en tenant compte d'une contrainte extérieure, l'équilibre de la balance des paiements courants étant. à l'époque de l'étude, un problème majeur en France. Les auteurs vérifient que la variation de PIB ainsi calculée est voisine de la variation de surplus du calcul de rentabilité de l'investissement : 325 MF contre 352 MF.

Morisugi et Hayashiyama (1997), calculent une évaluation ex-post de l'impact du réseau ferroviaire sur la croissance japonaise durant la période 1875-1940. Ils utilisent pour cela un modèle d'équilibre général calculable représentant une économie composée d'individus tous identiques, d'entreprises privées, d'un secteur transport, et d'un gouvernement qui en assure l'équilibre financier. La fonction d'utilité des individus comporte un bien courant et le bien transport de voyageurs dont la qualité dépend du niveau d'équipement ; la fonction de production du bien courant dépend aussi du bien transport de marchandises, du travail et du capital. La fonction de production est à rendements constants (fonction à élasticité de substitution constante). L'ensemble des paramètres est calé sur des données historiques ; le modèle fait apparaître que la contribution du réseau ferroviaire à la croissance du PNB, liée à la baisse du prix des transports qu'il a permis (et défalquée des dépenses de financement que la construction de ce réseau a nécessité) a été de 0,5 % en 1875, date du début de la construction du réseau, et a cru pour atteindre 12,3 % en 1940.

Ce travail constitue en quelque sorte le pendant japonais du travail fait par Fogel (1964) sur la contribution du réseau ferroviaire à la croissance américaine ; mais Fogel n'utilisait pas de modèle formalisé ; il reconstruisait par inférence vraisemblable quels auraient été, en l'absence du chemin de fer, les moyens de transport alternatifs, leurs coûts, et leurs conséquences sur l'activité économique. Il aboutit au résultat global qu'en l'absence des chemins de fer, le PIB américain aurait été réduit de 10 % à 20 % environ.

3.3. Externalités positives et rendements croissants

Dans le paragraphe précédent, le transport est considéré comme un facteur de production usuel, entrant dans des fonctions de production tout

1. Le modèle mini-DMS - Transport est un mini-modèle macro-économique dynamique. simplification du modèle DMS, auquel on a ajouté une description du secteur des transports détaillant les modes et distinguant voyageurs et marchandises. La structure du modèle est de type keynésien : les mécanismes principaux sont ceux du multiplicateur et de l'accélérateur, l'indexation prix-salaire, la compétitivité extérieure, et, sur le marché du travail, la relation de Phillips. Les fonctions de production sont du type Léontieff. En ce qui concerne les transports. chaque sous-secteur est doté d'une fonction de production à coefficients fixes et d'une demande liée aux prix. la politique publique (investissements taxation) est exo-

apparente à un bien public ou à un générateur d'effets externes positifs, et qui est lié aux théories de la croissance endogène (voir encadré).

ENCADRE 1-1

Les théories de la croissance et la croissance endogène

Le modèle classique de la croissance de Solow-Swan a fondé pendant des décennies notre analyse de la croissance. Sous sa forme la plus simple, il met en jeu les variables suivantes :

K le capital du pays,

L la force de travail, supposée suivre une croissance exogène à un taux géométrique n constant,

Y le produit,

δ le taux de dépréciation du capital,

s le taux d'épargne des ménages,

C la consommation des ménages,

I l'investissement brut.

On suppose que le produit Y dépend des entrants K et L à travers une fonction de production :

$$Y = F(K, L)$$

qui possède les propriétés suivantes :

$$\frac{\partial F}{\partial K} \geq 0 \quad \frac{\partial F}{\partial L} \geq 0 \quad \frac{\partial^2 F}{\partial K^2} \leq 0 \quad \frac{\partial^2 F}{\partial L^2} \leq 0 \quad F(\lambda, \lambda K) = \lambda F(L, K)$$

La dernière relation exprime la constance des rendements de la fonction de production agrégée nationale. Elle est nécessaire pour faire le raccord avec les analyses classiques de l'équilibre des marchés, fondées elles aussi sur la constance des rendements dans le secteur productif.

Des relations précédentes on tire :

$$\frac{dK}{dt} = sY - \delta K = sf(K, L) - \delta K \quad (1-1)$$

Il est alors usuel de raisonner en termes de capital par tête et de produit par tête, en posant :

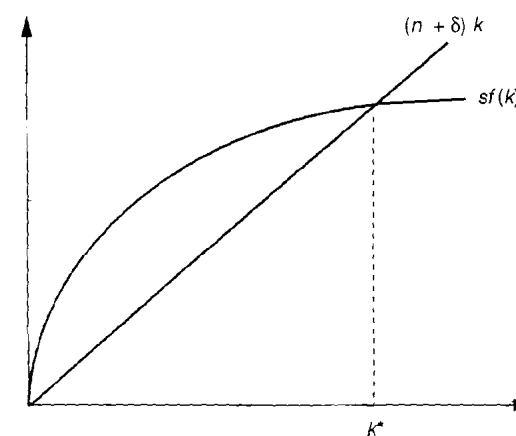
$$y = \frac{Y}{L} \quad k = \frac{K}{L}$$

La relation (1-1) devient alors

$$\frac{dk}{dt} = sf(k) - (n + \delta)k \quad (1-2)$$

En connaissant la fonction f , on pourrait résoudre complètement cette équation différentielle. Mais on peut graphiquement prendre une vue qualitative de l'évolution de k dans le temps, à travers les courbes figurées sur le graphique. On y constate intuitivement que k tend vers la valeur k^* . En effet si :

$$k \leq k^*$$



alors :

$$\frac{dk}{dt} \leq 0$$

et réciproquement.

Ce modèle aboutit au résultat qu'on tend vers un état stationnaire où le capital par tête et donc le produit par tête se stabilisent à une valeur unique. Ce résultat est contredit par les faits, qui montrent une croissance des revenus par tête, et l'absence de rattrapage des pays à faible revenu. Comment se sortir de cette contradiction ?

Une première réponse est de faire intervenir un facteur exogène, un *deus ex machina*, en l'occurrence le progrès technique, mais ceci ne rend pas compte de la divergence des revenus par tête.

Une autre réponse est de modifier le modèle pour y introduire un facteur de croissance endogène. Pour cela, un moyen est de poser :

$$y = A k$$

On rompt ainsi l'hypothèse de décroissance des rendements, mais ceci se soutient plus aisément si l'on considère que K et k représentent non pas seulement le capital physique, mais aussi le capital humain, le stock de connaissance, ou encore le stock de capital public, et qu'ils se rapportent non pas à chaque entreprise mais à l'ensemble de l'économie. Alors, en tenant compte de (1-2), on voit que le taux de croissance de l'économie est constant et positif, et vaut :

$$sA - (n + \delta)$$

Une forme analogue est celle dans laquelle la fonction de production s'écrit :

$$Y = AL^{1-\alpha} K^\alpha G^\beta$$

expression dans laquelle G représente le capital public, A , α , et β sont des paramètres.

Cet effet s'exprime à travers la fonction de production des entreprises, qui dépend alors, non seulement du capital privé et des consommations intermédiaires – ces dernières disparaissent quand on agrège les fonctions de production au niveau de l'ensemble du pays – mais aussi du capital public, et ici plus particulièrement du capital en infrastructures de transport. Pour chacune d'entre elles, dans cette hypothèse, la fonction de production s'écrira :

$$y_i = f(k_i, l_i, t_i, x_i) * A_i(G) \quad (1-3)$$

expression dans laquelle

k_i est le capital de l'entreprise i

l_i est sa force de travail

t_i est un vecteur représentant le coût et la qualité du transport

x_i est un vecteur représentant l'ensemble des autres entrants

G est un paramètre relié aux commodités de transport, représenté par le capital d'infrastructures de transport ou par un indice d'accessibilité (par exemple le nombre d'agents susceptibles d'être atteints en un temps donné ou la taille de l'agglomération, ou un autre indice d'accessibilité).

Dans cette expression, les transports interviennent comme usuellement en tant que consommation intermédiaire, par t_i . Cet effet est celui analysé dans la section 3-2 précédente. Mais ils interviennent aussi par l'intermédiaire du facteur $A_i(G)$. Ce facteur peut représenter différents mécanismes dont la liste n'est pas exhaustive :

L'amélioration des communications permet une meilleure information de l'entreprise sur les procédés de production, une accélération de la diffusion des progrès techniques, à travers des effets externes de réseau : ainsi le représentant qui rend visite à un client potentiel et qui supporte les frais de transport apporte au client, même si ce dernier n'achète pas, des informations sur le marché et les nouveaux produits; lui-même retire du contact infructueux des informations sur le marché.

D'autres externalités passent par le marché. L'amélioration des communications donne accès à des marchés plus vastes permettant à l'entreprise de mieux satisfaire à ses besoins, par exemple par un meilleur appariement entre les postes de travail et leurs titulaires grâce à un bassin d'emploi accru.

D'autres mécanismes enfin résultent de l'existence de rendements croissants (on verra au chapitre III que l'existence de rendements croissants est une des raisons fondamentales qui expliquent l'existence des agglomérations) : l'entreprise peut développer ses économies d'échelle en raison d'une extension des aires de marché de l'entreprise.

Les relations de ce type ont fait l'objet de nombreux tests statistiques, avec des agrégations aux niveaux géographiques ou sectoriels.

3.3.1. Tests au niveau national ou régional

Les tests au niveau national ont pour la plupart utilisé comme variable explicative l'ensemble du capital public; peu d'entre eux ont spécifié

l'effet du capital d'infrastructures de transport. Le tableau suivant, tiré de J. Hakfoot (1996), fournit les résultats de quelques études récentes, et indique l'élasticité au capital public; la plupart des fonctions testées étaient du type Cobb-Douglas, et étaient appliquées à des séries chronologiques.

TABLEAU 1-19

Auteur de l'étude	Année de parution	Aire géographique	Élasticité du PIB au capital public
Ratner	1983	USA	0,06
Costa	1987	USA	0,20
Aschauer	1989	USA	0,39
Duffy-Deno et Eberts	1991	USA	0,08
Munnell	1990	USA	0,15
Berndt et Hansson	1991	Suède	0,69-1,60
Toen-Gout et Jongeling	1993	Hollande	0,48

Les résultats de ces études sont, on le voit, assez dispersés.

Plusieurs critiques leur ont été adressées (cf. Quinet (1992) pour une recension de ces critiques) :

- Les données de base sont incertaines. Les séries de capital sont toujours difficiles à établir, notamment celles du capital public. Par ailleurs la frontière entre capital public et capital privé n'est pas claire, et est variable d'un pays à l'autre.

- L'hypothèse d'une fonction de production de Cobb-Douglas, qui fonde ces analyses, n'est certainement pas la plus adaptée aux économies actuelles des pays développés.

- Enfin les résultats numériques fournissent une élasticité de la production au capital public de l'ordre de 0,3 à 0,4, qui paraît élevée. Passe encore pour les portions de capital public directement productives comme les transports ou l'énergie, mais ces valeurs sont plus difficiles à admettre pour les équipements scolaires, hospitaliers...

Tout ceci doit provenir pour une large part de ce que l'estimation statistique sur séries chronologiques mêle aux effets d'offre des effets de demande, et notamment l'effet de multiplicateur de court terme, ainsi que le fait que l'investissement public est souvent la conséquence, non la cause, de la croissance.

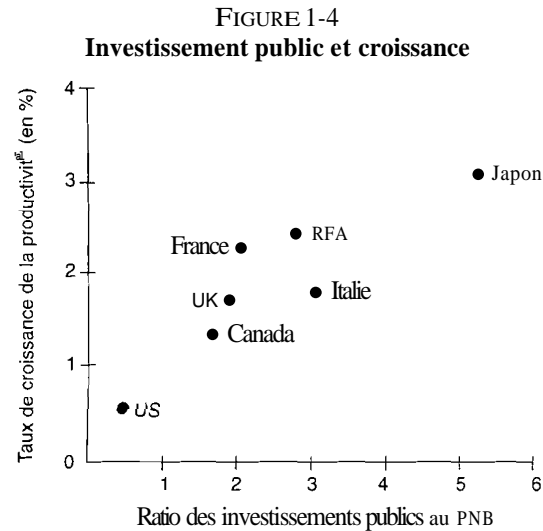
On est alors conduit à se pencher sur les résultats d'analyses en coupe instantanée, où l'on prend comme unité statistique des zones géographiques différentes et où les taux de croissance sont calculés, non pas sur un an, mais en moyenne sur une plus longue période, annihilant ainsi les effets de court terme.

Des études de ce type ont été faites par Barro (1989), Looney et Frederiksen (1981), Ford et Poret (1991), Eberts (1986), Costa, da Silva, Ellson et Martin (1987).

Au total les études en coupe instantanée qui éliminent les effets de causalité inverse (plus le PIB augmente, plus on a à la fois le besoin et les moyens financiers nécessaires pour investir) fournissent des valeurs des élasticités dispersées (car les

autres sources d'erreurs qui ne sont pas biaisées n'ont pas disparu : les séries de capital restent incertaines), mais dans l'ensemble beaucoup plus faibles.

Aschauer (1989) a également effectué des présentations très parlantes quoique moins rigoureuses entre les pays du groupe des 7 sur une même période de temps. Ainsi le graphique ci-dessous tiré de l'article cité met en relation l'investissement public et la productivité au cours des années 80 entre les pays du groupe des 7 ; il est particulièrement parlant même si rien ne prouve que la rentabilité « naturelle » des investissements publics, telle qu'elle résulte des calculs de rentabilité usuels ne suffise pas à expliquer l'excellente corrélation traduite par ce graphique (en fait la preuve est apportée par les raisonnements économétriques développés plus bas).



Peu d'études ont été menées qui spécifient l'effet des infrastructures de transport. Looney et Frederiksen spécifient l'investissement routier, comme on l'a déjà signalé, et lui trouvent un rôle significatif, de même d'ailleurs que Méra (1973).

Prud'homme (1996) et Fritsch et Prud'homme (1997) ont étudié spécifiquement l'effet des infrastructures routières sur la productivité pour la France, en effectuant une analyse en coupe instantanée sur les 21 régions. Ils trouvent une élasticité du PIB régional au capital d'infrastructures routières de l'ordre de 0,08 à 0,10.

3.3.2. Tests sectoriels

Des tests sectoriels ont été faits, notamment par Nadiri et Mammuneas (1995). Ceux-ci estiment des fonctions de coût par secteur et examinent leur évolution dans le temps, reliant l'évolution de la productivité par secteur à celle du capital routier. Ils trouvent une contribution des investissements routiers à la productivité variable selon les secteurs, l'élasticité

étant en moyenne de 0,05 %, un chiffre nettement plus faible que ceux fournis dans les autres études.

3.3.3. Existe-t-il des externalités positives ?

Qu'enseignent ces résultats quant à l'existence d'effets externes positifs ? On ne peut pas en tirer de conclusions directes. En effet, les corrélations calculées fournissent l'effet global de l'amélioration des infrastructures – publiques ou de transport – sur l'augmentation de la production ; mais elles ne font pas le partage, dans cet effet global, entre ce qui est dû à la réduction des coûts de transport et à leur effet à l'intérieur de la fonction de production, et ce qui est dû à l'effet externe.

On peut toutefois faire ce partage, par un détour de raisonnement : les élasticités au capital public se traduisent par une rentabilité ex-post des investissements leur concernant. En retranchant de cette rentabilité calculée ex-post, au niveau macro-économique, les rentabilités calculées ex-ante de ces mêmes investissements, on obtient une évaluation de grandeur de l'effet externe recherché, puisque, dans tous les calculs de rentabilité l'hypothèse d'absence d'effets externes de ce type est posée¹. Les élasticités au capital public mises en évidence dans les études passées en revue se situent entre 0,05 et 0,30. Quinet (1992) et Prud'homme (1997) ont calculé les rentabilités ex-post qui en résultent pour les investissements correspondants : rentabilité pour les entreprises et surplus des consommateurs finaux de transport (déplacements de personnes pour motifs privés). Les résultats sont à peu près du double, soit dans une fourchette 0,10-0,60. Nadiri et Mammuneas aboutissent à une rentabilité moyenne des investissements routiers de 28 %. Morrison et Schwartz (1996) obtiennent des résultats similaires à partir d'une analyse qui combine découpage géographique et sectoriel. Or les rentabilités ex-ante des investissements routiers se situent usuellement dans la fourchette 0,05-0,20. La comparaison des deux fourchettes fait planer une forte présomption en faveur de l'existence d'externalités positives, et incite au minimum à développer les recherches en ce domaine.

1. De la formule (1-3) précédente, on tire une fonction de coût pour l'entreprise i :

$$C_i(y_i, r, w, c, \pi, G) = k_i \min_{x_i} (rk_i + wl_i + ct_i + \pi x_i)$$

avec :

$$y_i = A_i(G) f(k_i, l_i, t_i, x)$$

Un investissement se traduit par une variation (marginale) de G et une variation de c et entraîne une variation du coût C, telle que d'après le théorème de l'enveloppe :

$$\frac{\partial C_i}{\partial c} = t_i \quad \text{et} \quad \frac{\partial C_i}{\partial G} = \frac{\partial C_i}{\partial y_i} y_i \frac{A_i}{A_i}$$

Le calcul de rentabilité usuel ne prend en compte que le premier terme, l'évaluation de la rentabilité ex-post intègre les deux.

3.3.4. Constatations historiques

D'autres conséquences, mal quantifiables, sont fondamentales. Ce sont celles qui se manifestent sur le long terme, à la suite de progrès dans les transports. Il en résulte des transformations dans les modes d'exercice de la vie économique et sociale, non seulement sur le plan quantitatif mais aussi d'un point de vue plus qualitatif.

Quelques exemples, et d'abord la création d'un marché national unique au cours des 19^e et 20^e siècles dans les pays européens, avec comme conséquence la disparition des famines, en raison de la possibilité de transporter les denrées alimentaires d'un point à l'autre du territoire dans des conditions de coût et de délais acceptables.

L'amélioration des transports a été aussi un moteur du développement des échanges internationaux : il y a quelques siècles, seuls les épices, denrée rare et coûteuse, venaient de l'Extrême-Orient ; maintenant ce sont toutes les matières premières, même celles dont la valeur est faible, qui peuvent être importées de contrées lointaines.

Les délocalisations, dont on se plaint si souvent et qui deviennent un outil de stratégie essentiel pour les entreprises, ne présentent leur attrait qu'en raison des progrès des transports. Ces progrès permettent aux entreprises de rechercher des approvisionnements et des clients de plus en plus loin, donc de développer leur production et de bénéficier des économies d'échelles qu'une production accrue entraîne. On peut dire qu'une des causes fondamentales de la mondialisation à laquelle nous assistons réside dans l'amélioration des transports et l'abaissement de leur coût.

Le même genre de constatations pourrait être fait en ce qui concerne les transports de voyageurs, avec le développement du tourisme lointain, l'extension des voyages d'affaires, ou le rapprochement des modes de vie qui résulte de ces contacts de plus en plus fréquents et éloignés, le développement urbain et les phénomènes de métropolisation. L'abolition progressive des distances est une caractéristique de la vie moderne au même titre que la multiplication des biens et services qui résultent de la croissance économique.

4. Les implications

Les rapports des transports avec l'activité économique générale ont de tout temps fourni le support à des mesures de politique économique. On n'a pas attendu Keynes pour soutenir la conjoncture et réduire le chômage par des politiques de grands travaux, même si c'est lui qui, le premier, a popularisé l'effet multiplicateur de l'investissement. Les ateliers nationaux de 1848 correspondaient à cette idée. Le plan Freycinet de 1870, du nom du Ministre des Travaux Publics qui l'avait fait adopter, avait aussi pour but de relancer l'activité de la sidérurgie grâce à la construction d'un réseau de chemin de fer et de voie navigable ; et l'idée que ce réseau de

communications améliorées aurait, une fois construit, un effet bénéfique sur l'activité économique n'était certainement pas absente de l'esprit de ses promoteurs.

L'analyse économique permet de mieux cerner la portée et les limites de ce type d'action pour le soutien de la conjoncture et de l'emploi. Tout d'abord, les effets de la décision de réaliser un investissement mettent du temps à prendre corps : à moins que le projet en soit déjà arrivé au stade des plans d'exécution et que les terrains soient disponibles, il faut un ou deux ans avant les premiers travaux sur le terrain, et ces travaux eux-mêmes prennent plusieurs années. Ces délais de réalisation doivent être soigneusement introduits dans les données exogènes des modèles, qui font alors en général apparaître que les effets de relance sont bien tardifs au regard des exigences de la conjoncture. Ils montrent aussi que ces effets sont éphémères, et qu'une fois l'investissement réalisé, on revient peu à peu à la situation initiale.

En outre les modèles macro-économiques intègrent rarement la totalité des conséquences de la mesure qu'on cherche à évaluer. Ainsi IMPACT et DÉPENSE PUBLIQUE, déjà cités, supposent que les prix sont fixes, une hypothèse peut-être valable dans le très court terme, mais pas à long terme. D'autres modèles se débarrassent partiellement de cette hypothèse, mais fort peu vont jusqu'à endogénéiser les taux d'intérêt et de change, et lorsqu'ils le font c'est au prix d'une simplification de la description de l'appareil productif qui les rend impropres à l'usage recherché, le test des conséquences d'une mesure particulière. Il faut donc accueillir avec prudence les affirmations du type : « la mesure crée tant de milliers d'emplois ».

Aussi leurs résultats ne fournissent qu'une partie des conséquences des actions qu'ils analysent. Les conséquences les plus directes sur le système productif. En outre, dès que pèsent des contraintes macro-économiques, les mesures d'investissements envisagées doivent être accompagnées de contreparties ; par exemple le financement par le budget de l'État d'une infrastructure nouvelle a un effet favorable sur l'emploi ; mais si une contrainte d'équilibre pèse sur ce budget ou sur la dette publique, cette dépense devra être gagée par une économie d'un montant analogue qui, elle, aura un effet déflationniste et le résultat final sera de signe indéterminé. Si on finance la mesure par des impôts supplémentaires, alors un théorème dû à Haavelmo dit que l'augmentation de PIB qui en résulte est juste égale au montant de l'investissement, ceci dans le cadre de prix fixes.

On peut se demander en revanche si la pratique décisionnelle donne une place suffisante aux effets d'externalité positive que peut créer le système des transports. Que les transports aient sur la croissance un effet allant au-delà de ce qu'impliquerait un facteur de production banal, des raisons théoriques et expérimentales permettent de le considérer : les transports et leurs améliorations autorisent le développement des contacts, une meilleure information et communication, l'accroissement des aires de marché, un meilleur appariement des besoins avec les moyens de les satisfaire. Certes la mesure précise de cet effet reste à faire.

Et pourtant, les conséquences pratiques à en tirer peuvent être importantes. En termes d'infrastructures d'abord. Il conviendrait de développer les réseaux au-delà de ce à quoi conduisent les évaluations coût-bénéfice actuelles. Peut-être est-ce l'intuition de ce type de conséquence qui a inspiré certaines décisions de programmation. On a cité plus haut le programme Freycinet lancé en France à la fin du 19^e siècle. Il faut aussi mentionner les débats qui ont eu lieu aux États-Unis sur le retard de ce pays en équipements publics, et sur l'intérêt de lancer un grand programme d'infrastructures qui fut inscrit dans la « plate-forme » de B. Clinton. Citons aussi l'adoption par l'Union Européenne d'un programme de Réseaux Trans Européens destinés à accélérer la construction européenne et à promouvoir le développement à l'intérieur de l'espace européen.

Une autre conséquence à tirer de l'existence d'externalités positives concerne la tarification des transports. Si les résultats statistiques sur les effets du capital d'infrastructure sont liés à des externalités de réseau, ou à la concrétisation de rendements croissants dans les entreprises qui en bénéficient¹, alors une réduction de la tarification serait justifiée. Faut-il penser que la sous-tarification consciente dont nombre de transports sont l'objet de la part de la puissance publique – par exemple les transports urbains, les transports collectifs régionaux – est le fruit d'une intuition de ce résultat ?

On pourrait être tenté de rapprocher de ces constatations et de ces tendances le mouvement général qui pousse chaque zone géographique, région, département ou commune, à réclamer un développement des infrastructures qui la desservent, arguant des effets qu'un désenclavement aurait sur leur activité économique. Certes, dans ces demandes, il y a l'idée que les infrastructures développent la productivité ; mais il faut bien sûr y voir aussi implicitement l'idée que de bonnes dessertes sont un facteur d'attraction des entreprises. C'est alors, explicitement ou non, un argument de localisation favorisée par la qualité des transports, argument dont on va voir maintenant le bien-fondé.

BIBLIOGRAPHIE

- D.A. Aschauer (1989), « Is Public Expenditure Productive ? », *Journal of Monetary Economics*, 23.
 R. Barro (1989), *Economic growth in a cross-section of countries*, Harvard University.
 R.J. Barro et X. Sala-i-Martin (1995), *Economic Growth*, Mac-Graw Hill.

1. Dans ce dernier cas une raison de sous-tarifier le transport résulte de ce que l'entreprise disposant d'un pouvoir de monopole va alors, compte tenu de cette sous-tarification de son facteur de production, pratiquer un prix plus proche du coût marginal réel, donc plus conforme à l'optimum économique.

- E. Bennathan, J. Fraser et L. Thompson (1992), « What determines demand for freight transport ? », *WPS*, 998 World Bank, Washington D.C.
 E.R. Berndt et B. Hansson (1991). « Measuring the contribution of Public Infrastructure Capital in Sweden », *NBER Working Paper* n° 3842.
 J.C. Blain et L. Nguyen (1994). *Modélisation de trafics de voyageurs*, Notes de synthèse OEST, janvier.
 D. Bureau et C. Cipriani (1986), *Impacts macro-économiques du TGV Paris Sud-Est*, Roneoté-Ministère des Finances, Paris.
 CERAS, CERTU, SES (1997), *Le modèle Dépense Publique*, miméo, Paris.
 L. Clément. A. Lauer. J. Maurice et J.P. Taroux (1997), *Le modèle Dépense Publique*, Note Ministère des Transports, Paris.
 C. Charmeil (1968), *Essai sur l'influence macro-économique des investissements*, Annales des Ponts et Chaussées, mai.
 Commissariat Général du Plan (1992), *Transports 2010*, La Documentation Française- Paris.
 Costa, J. da Silva. R.W. Ellson et R.G. Martin (1987), « Public Capital, Regional Output and Développement : Some Empirical Evidence », *Journal of Regional Science*, 27.
 DAEI-SES- INSEE (1996), Paris, *Les Transports en chiffres*, brochure SES Paris, 1996. *Note de conjoncture des Transports*, brochure SES Paris, déc. 1996.
 K.T. Duffy-Deno et R.W. Eberts (1991), « Public Infrastructure and Regional Economic Development : A Simultaneous Equation Approach », *Journal of Urban Economics*, vol. 30.
 K. Eberts (1986), *Estimating the contribution of urban public infrastructure to regional growth*, F.R.B. Cleveland.
 R.M. Fogel (1964), *Railroad and American Economic Growth*, The John Hopkins Press.
 R. Ford et P. Poret (1991). *Infrastructure and private sector productivity*, working paper, OECD, Paris.
 B. Fritsch et R. Prud'homme (1997), « Measuring the Contribution of Road Infrastructure to Economic Development in France », in *The Econometrics of Major Transport infrastructures*, E. Quinet and R. Vickerman editors, Macmillan.
 C. Gerondeau (1996), *Les Transports en Europe*, EDS, Paris.
 M. Girault et J.C. Blain (1997), *La demande de Transport en 2015*, Études du SES, Ministère des Transports, Paris
 M. Girault, J.C. Blain et K. Meyer (1995), *Élasticités de court et de long termes des trafics de marchandises à la croissance économique*, OEST, notes de synthèse, mars.
 M. Girault et J.M. Kail (1997), *Perspectives de la demande de transport et des émissions de polluants à l'horizon 2015*, Actes du Congrès de l'ATEC, Presses de l'École des Ponts, Paris.
 B. Goodwin (1992), « A review of new demand elasticities with special reference to short and long run effects of price changes », *Journal of Transport Economics and Policy*, May.
 J. Hakfoot (1996), « Public Capital, Private Sector productivity and Economic Growth : a macro-economic perspective », in *Infrastructure and the com-*

- plexity of Economic development, D. Batten and C. Karlsson editors, Springer.
- INSEE (1986), *Effets d'entraînement multisectoriels à court et moyen termes des investissements en infrastructures et matériel de transport*, Ronéoté, Paris.
- L. Kechi (1996), *Synthèse de documents siir les élasticités tarifaires*, CETUR note ronéotée.
- R. Looney et P. Frederiksen (1981), « The regional impact of infrastructure investment in Mexico », *Regional Studies*, 15.
- J.L. Madre et T. Lambert (1989), *Prévisions à long terme du trafic automobile*, CREDOC, Paris.
- J.L. Madre (1997), *La croissance des transports de voyageurs est-elle évitable*, 14^e Symposium de la CEMT, Innsbruck.
- K. Méra (1973), « Regional production functions and social overhead capital », *Regional and Urban Economics*, 3.
- K. Meyer (1997), « L'économétrie au service du fret ferroviaire » *Rail et Recherche*, février, Paris.
- M. Morisugi et Y. Hayashiyama (1997), « Post-Evaluation of the Japanese Railway Network 1875-1940 », in E. Quinet and R. Vickerman editors. *The Econometrics of Major Transport infrastructures*, Macmillan, London.
- C. Morrison et A.E. Schwartz (1996), « State infrastructure and productive performance », *The American Economic Review*, dec.
- A.H. Munnell (1992), « Infrastructure Investment and Economic Growth », *Journal of Economic Perspectives*, vol 6, 4.
- M.I. Nadiri et T.P. Mamuenas (1991), *Highway Capital Infrastructure and Industry Productivity Growth*, NBER and Cyprus University.
- T.H. Oum, W.C. Waters II et J.S. Yong (1992), « Concepts of price elasticities of transport demand and recent related estimates », *Journal of Transport Economics and Policy*, May.
- E. Quinet (1993), « Transport between monopoly and competition », in J. Polak and A. Heerje (1993), *European Transport Economics*, Blackwell.
- E. Quinet (1994), *Rapport Route-Air-Fer*, Rapport ronéoté pour le compte du Ministère des Transports.
- R. Prud'homme (1996), « Assessing the role of Infrastructure in France by Means of Regionally Estimated Production Functions », in *Infrastructure and the complexity of Economic development*, D. Batten and C. Karlsson editors, Springer.
- E. Quinet (1992), *Infrastructures de Transport et croissance*, Economica.
- J.B. Ratner (1983), « Government Capital and the Production Function for US Private Output », *Economics Letters*, 13.
- SES (1996), *Les transports en 1995*, Rapport de la Commission des comptes de transport de la Nation, La Documentation Française, Paris.
- E. Selvanathan et S. Selvanathan (1994), « The demand for transport and communication in the United Kingdom and Australia », *Transportation Research*, vol. 28 B, n° 1.
- J.P. Taroux et G. Buchmuller (1980), *PRETRAP : modèle de prévision de trafic de personnes*, SAE, Ministère des Transports.
- J.P. Taroux (1995), *Modèle IMPACT. Rapport d'actualisation*, OEST, Ministère des Transports, Paris.
- M.W. Toen-Gout et M.M. Jongeling (1993), « Investeringen in infrastructuur en economische groei », *ESB*, 12 mai 1993.
- J.C. Toutain (1967), « Les Transports en France », *Archives de l'ISEA*, septembre-octobre.
- J. F. Vacher (1997) « L'évolution des Transports depuis 40 ans », *INSEE première*, n° 522, juin, Paris.
- E. Van de Voorde et H. Meersmann (1997), *La croissance des transports de marchandises est-elle évitable ?*, 14^e Symposium de la CEMT, Innsbruck.

Transports et localisation

L'étude des interactions entre transports et localisation des activités peut être menée à différents niveaux. Au niveau micro-économique, on analysera la manière dont les unités économiques élémentaires prennent leurs décisions de localisation en fonction des conditions de transport. Au niveau macro-économique, on examinera les degrés d'activité des régions en fonction de la qualité des transports dont elles bénéficient. Chacun de ces deux aspects fera l'objet d'une section. La troisième section sera consacrée à une appréciation des conséquences pratiques à tirer des considérations précédentes.

1. Transport et localisation des unités économiques

On se centrera ici sur les unités de production, c'est-à-dire les entreprises. Les liens entre transport et localisation des ménages, de nature surtout urbaine, étant analysés dans le chapitre III « Les transports et la ville ».

La théorie classique (Perreur, 1988) donne au transport un rôle essentiel dans le choix de localisation de l'entreprise : celle-ci se place à l'endroit qui minimise son coût total de transport (acheminement des intrants et expédition des produits). Ce principe conduit à une localisation unique, indépendante du niveau de production, et il aboutit à une maximisation du profit de l'entreprise, mais sous réserve d'hypothèses restrictives : constance des rendements et des coefficients de production, absence d'externalités, niveau et localisation de la demande prédéterminé. Analysons ce qui se passe lorsqu'on se débarrasse de chacune de ces hypothèses.

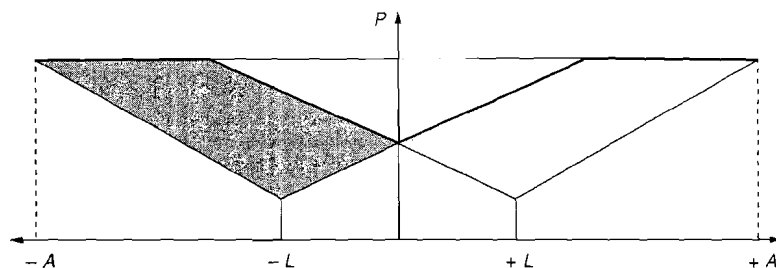
1.1. Demande prédéterminée en volume et en localisation

Cette hypothèse est un peu le pendant en économie spatiale de l'hypothèse classique de concurrence parfaite selon laquelle le prix est une don-

née pour l'entreprise. En fait l'introduction de l'espace induit une situation de concurrence dans laquelle les entreprises ont un pouvoir de marché sur la clientèle qui les entoure en raison de la protection que leur procure le coût du transport. On peut examiner ce point sur un modèle simple de duopole spatial, dans lequel deux entreprises identiques, fixant des prix rendus client, se concurrencent le long d'un espace linéaire sur lequel la clientèle est répartie selon une densité uniforme et achète une unité du bien, pourvu que son prix soit inférieur à une borne supérieure P . On suppose que les deux entreprises ont le même coût marginal de production c et que le coût de transport unitaire est t : en chaque point d'abscisse x , le coût total (production + livraison) pour l'entreprise ($-L$) est : $c + t|x + L|$; il est représenté par l'entonnoir assis au point d'abscisse ($-L$), et même chose pour l'entreprise ($+L$). Si ces entreprises se font, pour chaque client défini par son abscisse x , une concurrence à la Bertrand¹, on peut déterminer :

- les prix pratiqués (segments en traits épais) qui correspondent au coût de l'entreprise la moins performante ;
- les aires de marché ($0 - A$ pour l'entreprise 1 ; $0 + A$ pour l'entreprise 2), chaque entreprise opérant dans l'aire de marché où elle a les coûts les plus faibles ;
- et les profits de chacun de ces entreprises, égaux aux aires grisées.

FIGURE 2-1
Duopole spatial prix rendu



On illustre ainsi le fait que l'espace introduit une différence majeure par rapport à la situation a spatiale, savoir que chaque entreprise dispose d'un pouvoir de marché en son voisinage. La réduction des coûts de transports accroît la compétition entre les entreprises en permettant un élargissement des aires de marché.

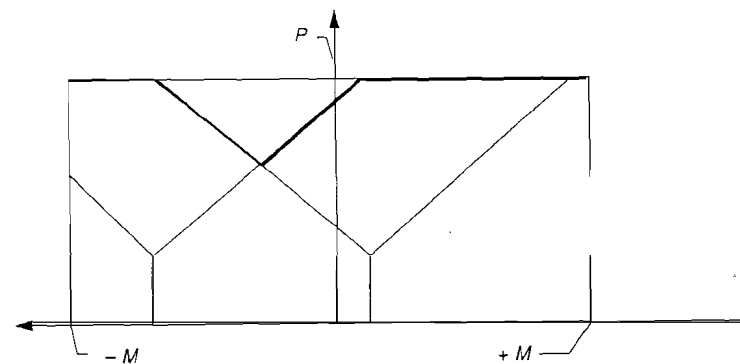
Ceci conduit à des interactions stratégiques en termes de choix de localisation. On peut le voir en modélisant les stratégies des deux entreprises

1. Pour des rappels sur les duopoles et notamment le duopole de Bertrand, voir le chapitre VIII et notamment la section 8-1.

précédentes sous forme d'un jeu à deux étapes, la 1^{re} étant pour chacune le choix d'une localisation, et la seconde le choix d'un prix. On suppose que l'espace est limité en étendue, entre les abscisses M et $+M$ (figure 2-2). On démontre alors que la solution de ce jeu (équilibre de Nash) correspond aux localisations $-M/2$ et $+M/2$ pour chaque entreprise ; c'est, pour chacune d'elles, la localisation qui minimise le coût de transport à aire de demande donnée ($0 - +M$ pour l'entreprise 1, $0 - M$ pour l'entreprise 2), mais l'aire de demande n'est pas prédéterminée, elle résulte des conditions de transports et du jeu stratégique entre les entreprises.

Au duopole spatial avec prix rendu client on peut comparer le duopole spatial avec prix départ usine. Dans ce dernier cas, il n'y a qu'un prix par entreprise, alors que dans le premier il y en avait autant que de localisations des clients. Alors on démontre que, avec des coûts de transport proportionnels à la distance, il peut, à localisation fixée, ne pas exister d'équilibre de prix. L'équilibre de localisation est encore aux points $-M/2$ et $+M/2$, et le prix commun départ usine est $C + 2tM$ (voir Thisse, 1988).

FIGURE 2-2
Localisation des entreprises



On peut aisément repérer les facteurs qui entrent en jeu dans le choix de ces localisations : pour chaque entreprise, se rapprocher de l'autre augmente son aire de marché, mais l'oblige, en raison de la concurrence accrue, à réduire ses prix. Les localisations optimales sont celles où ces deux facteurs antagonistes s'équilibrent. Rien sûr, les résultats obtenus dépendent de la forme de la fonction qui relie la distance au coût de transport. Les résultats précédents sont établis sur la base de coûts de transport linéaires en fonction de la distance. Dans une autre hypothèse souvent évoquée, celle où le coût de transport varie comme le carré de la distance, on aboutit au résultat que les entreprises recherchent la différenciation maximale et se placent aux extrémités de l'espace.

On dispose également du résultat selon lequel, lorsque deux entreprises sont en compétition imparfaite (duopole de Cournot) pour servir des marchés situés aux nœuds d'un réseau de transport, alors les localisations optimales de chacune des deux entreprises se situent sur des nœuds du réseau (Thisse, 1993). Ce résultat est confirmé par l'observation (Plasard 1976), et donne une certaine robustesse à l'idée que la structure des coûts de transport et les conditions de la concurrence conduisent à des localisations aux nœuds de transport et donc à une polarisation de l'espace.

1.2. Constance des rendements et externalités

On pourrait penser que la réduction constante des coûts de transport et de leur part dans le coût de production viendrait atténuer le déterminisme que les conditions de transport font peser sur la localisation, et conduirait à une répartition plus uniforme des activités sur le territoire, aboutissant à la limite à l'archétype du « *backyard capitalism* ». Cette intuition est contredite par les faits : Savy et Veltz (1993) constatent, en France comme dans le reste du monde, une tendance à la métropolisation, marquée selon les auteurs par le passage d'une production de masse à une économie de variété et de services; et par l'importance croissante de l'incertitude, appelant souplesse et flexibilité.

Les mécanismes économiques, analysés notamment par Jayet, Puig et Thisse (1995), conduisent à un renforcement de la polarisation et de la concentration. Ce résultat provient du jeu des effets externes d'agglomération et des rendements croissants. Ces effets sont ceux qui ont été analysés au chapitre précédent (section 3.3.). On s'y préoccupait alors de leurs conséquences macro-économiques, de leur effets sur la production et sur la productivité. On s'intéresse ici à leurs conséquences en termes de localisation des activités.

Les entreprises préfèrent s'installer là où elles trouveront des marchés plus vastes pour leur approvisionnement, des bassins de main-d'œuvre larges, où elles pourront bénéficier d'une plus grande diversité de services spécialisés. Dans le même sens, ces services spécialisés, qui constituent des activités à rendement croissant, ont besoin, pour amortir leurs coûts fixes, d'un large marché. L'agglomération des activités permet à chaque entreprise de bénéficier de la diversité; elle les conduit aussi à développer cette diversité, qui est un moyen de pallier aux inconvénients qu'entraîneraient une proximité accrue des concurrents, due à leur agglomération et à la réduction des coûts de transport. Ceci se traduit par des avantages pour les consommateurs salariés (biens plus diversifiés, salaires plus élevés), qui les incitent à s'installer dans les agglomérations et induit une spirale de croissance de cette agglomération.

Les études de Prud'homme (1998) et Poulit (1994) ont bien mis en évidence la croissance de la productivité et des revenus en fonction de la taille de l'agglomération.

Ainsi la polarisation géographique apparaît comme une des formes de la croissance, et les résultats de l'analyse moderne rejoignent ceux mis en avant dans le passé par exemple par A. Marshall avec la notion de « district industriel », ou par F. Perroux avec celle de « pôle de développement ». Ces tendances de fond résultent évidemment de bien d'autres facteurs que de l'évolution des conditions de transports, mais ces dernières les accompagnent et les accentuent.

On peut objecter que dans les modèles présentés la représentation des transports est très schématique. Peters, Thisse et Thomas (1997) ont étudié un modèle où le réseau de transport, sans être très complexe, est plus élaboré. Deux hypothèses sont considérées : celle d'un réseau à mailles rectangulaires et celle d'un réseau concentrique. On y cherche la localisation optimale d'une entreprise qui minimise ses coûts totaux, sachant qu'elle doit servir une demande uniformément répartie sur tous les nœuds du réseau. C'est bien sûr la présence de coûts fixes qui entraîne qu'il n'y a pas une usine par nœud, et les usines sont d'autant moins nombreuses que les coûts fixes sont élevés. Il ressort que, dans la situation de réseau concentrique, la tendance à la localisation centrale est beaucoup plus forte que dans le réseau rectangulaire, ce qui montre l'influence de la forme du réseau d'infrastructures sur la localisation. On peut, comme le font les auteurs, illustrer ce résultat par la constatation des différences entre les structures françaises, très centralisées autour de Paris et accompagnées d'un développement en étoile des infrastructures autour de la capitale, et les structures allemandes, qui présentent les caractéristiques inverses. La tendance à la centralisation dans le cas radial est atténuée si une rocade est mise en place pas trop loin du centre. Enfin les auteurs étudient la situation de deux réseaux voisins, l'un radial l'autre rectangulaire, d'abord séparés par exemple par une frontière, puis réunis par la suppression de cette frontière (cas des marchés communs). Il apparaît que les localisations dans la situation finale ne sont guère modifiées par cette suppression; il se produit seulement une certaine tendance à la localisation près de l'ancienne frontière.

FIGURE 2-3
Structure réticulaire à deux niveaux

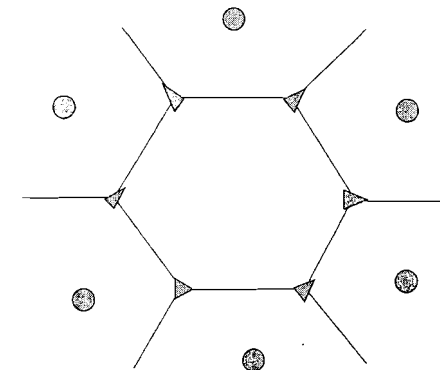
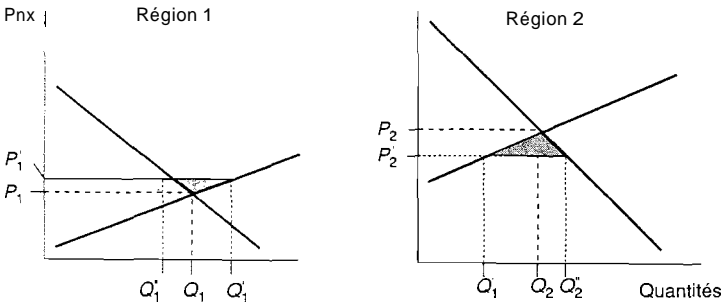


FIGURE 2-4



Cette analyse est à rapprocher de la théorie des hiérarchies urbaines de Losch (1940) et Christaller (1933) selon lesquelles les villes s'organisent en réseaux successifs, de mailles d'autant plus larges que l'on montre dans la hiérarchie et la taille des villes, cette montée dans la hiérarchie s'accompagnant d'une diversité croissante des services offerts ; ils sont de plus en plus spécialisés, et leur marché s'étend sur une aire croissante (figure 2-3).

2. Transport et développement régional

Les considérations qui viennent d'être développées justifient qu'on s'intéresse aux effets des transports sur le développement régional, qu'on cherche à le quantifier, à dégager les lois auxquelles il obéit, ou à le prévoir. Ces préoccupations s'expriment par deux démarches; la modélisation et l'observation des faits.

2.1. Les transports dans les modèles régionaux

Dans la modélisation régionale, les transports jouent un rôle central à travers les deux mécanismes suivants :

- Tout d'abord, à l'intérieur de chaque région, la réduction des coûts de transports entraîne une modification des coûts de production. donc de la demande intérieure à la région, et une transformation du niveau d'activité des branches qu'on peut apprécier à travers les tableaux d'échanges interindustriels. Notons que dans la plupart de ces modèles les entreprises opèrent à rendements constants et qu'il n'y a pas d'effet de diversification des biens, contrairement aux considérations développées dans la section précédente. Les effets de polarisation seront donc mal traduits.

- Ensuite, entre régions, une réduction des coûts de transports entraîne une modification des échanges d'importation et d'exportation, selon un schéma présenté ci-dessous dans le cas particulier où, pour simplifier la présentation, il n'y avait aucun échange (coûts de transport infinis) dans la situation « avant ». Dans la situation « après », des échanges se produisent, et les prix des deux régions se rapprochent ; leur différence devient égale au coût de transport.

		Région 1	Région 2
Avant	Quantités produites	Q1	Q2
	Quantités consommées	Q1	Q2
	Prix	P1	P2
Après	Quantités produites	Q'1	Q'2
	Quantités consommées	Q''1	Q''2
	Prix	P'1	P'2

L'équilibre est tel que $Q_1' + Q_2' = Q_1'' + Q_2''$. La figure fait apparaître des gagnants et des perdants. Les gagnants ce sont les consommateurs de la région 2 et les producteurs de la région 1. Les autres sont perdants (on verrait, en utilisant les outils qui seront définis au chapitre VII que le bilan global intégrant les gagnants et les perdants est exactement égal au surplus engendré par la réalisation de l'infrastructure). Ce résultat conforte l'idée générale que lorsqu'on améliore les relations entre deux régions: il y a spécialisation par branche et que globalement ce sont les producteurs de la région la plus efficace économiquement qui sont bénéficiaires.

Dans la plupart des modèles, les échanges interrégionaux sont modélisés selon ce principe, mais d'une manière moins brutale, et en tenant compte de la multiplicité des régions. Pour chaque produit, le volume d'échange entre deux régions dépend de la différence de prix et des coûts de transports entre les deux régions, selon des formules du type (Rietveld et Nijkamp, 1993) :

$$t_{irs} = k_{is} \exp(-\beta_i(p_{is} + u_{irs})) / \sum_q k_{iq} \exp(-\beta_i(p_{iq} + u_{irq}))$$

formule dans laquelle :

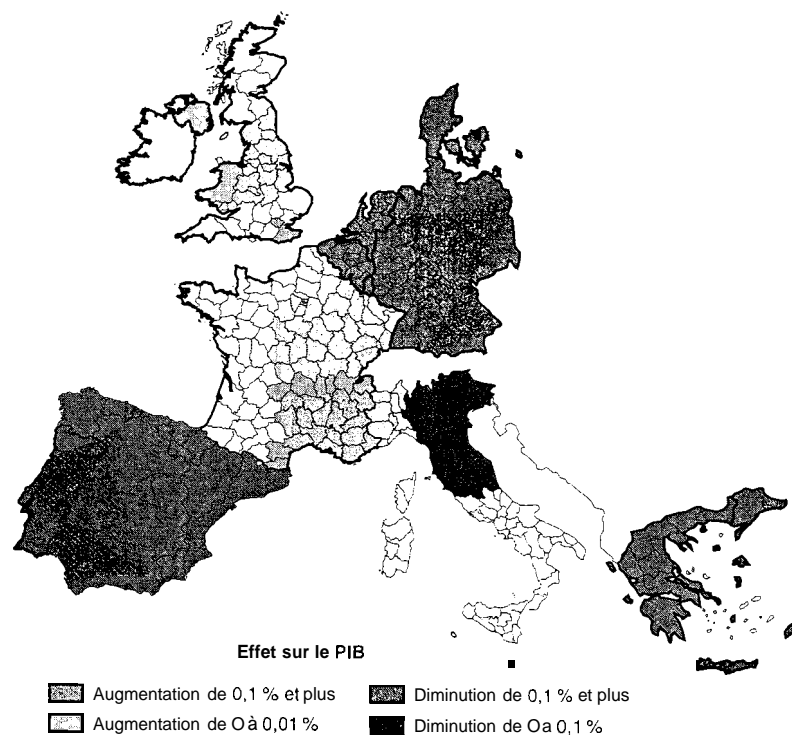
- t_{irs} est la part du trafic en bien i entre r et s,
- k_{iq} est la capacité de la région q en production du bien i,
- p_{iv} est le prix du bien i dans la région v,
- u_{irs} est le coût de transport du bien i de r en s.

Ces modèles font aussi intervenir les conditions de transport dans la localisation des facteurs de production. Par exemple les migrations de population entre régions sont déterminées en fonction des revenus par tête, des niveaux d'emploi et des distances de transports.

2.1.1. Un exemple de modèle multirégional transport-économie : Meplan

Citons à titre d'exemple les caractéristiques du modèle Meplan, qui a été utilisé pour tester l'effet du Tunnel sous la Manche pour l'économie des régions européennes (Rapport Cost 317).

FIGURE 2-5
Les effets du Tunnel sous la Manche
d'après le modèle Meplan



Ce modèle, élaboré par Marcial Echenique and Partners, procède à une évaluation conjointe de l'activité économique par région et des flux de transport, sur la base d'un modèle input-output régionalisé. La demande de transport et le développement régional dépendent bien sûr des caractéristiques de l'offre de transport. Dans l'application du modèle au Tunnel sous la Manche, l'Europe était divisée en 33 régions. Le modèle comporte trois modules :

- Le module économique régional évalue les transformations des structures de production et consommation dans chaque région, sur la base de la croissance attendue des populations et des revenus, et des différences régionales d'accessibilité. Dans chaque région, les liens entre les secteurs sont représentés par des coefficients de matrice *input-output*. La demande

pour chaque bien est satisfaite à la fois par la production de la région et par les importations, dont le volume dépend des écarts de prix avec les autres régions et des coûts de transport.

- Un module d'interface transforme les échanges commerciaux en flux de transports.

- Un module transport déduit de ces flux des coûts de transport qui, réinjectés dans le module économique, doivent, à l'équilibre, être cohérents avec les coûts de transport choisis dans l'étape précédente du processus itératif.

Ce modèle fournit de nombreux résultats. On ne retiendra ici que ceux qui concernent les changements de PIB régionaux résultant de la mise en service du Tunnel sous la Manche.

On voit ainsi nettement apparaître sur cette carte le résultat du modèle théorique présenté dans la figure 2-4 : il y a des gagnants et des perdants dans la réalisation d'une infrastructure ; elle produit des effets positifs sur certaines régions, négatifs sur d'autres. On voit aussi l'effet de polarisation de l'activité créé par l'infrastructure : l'activité continentale augmente en son voisinage, et se réduit dans les zones éloignées.

2.1.2. Autres modèles

Il existe de nombreux autres modèles régionaux ; une référence historique est le modèle d'Amano et Fujita (1970), dont l'organisation générale est figurée sur le schéma joint tiré de l'article des auteurs.

Les transports interviennent en plusieurs endroits :

- pour le choix des modes et des itinéraires qui dépend des caractéristiques de l'offre de transport, c'est-à-dire de la qualité de l'infrastructure et du volume de trafic ;

- la fixation des trafics conduit à définir le coût de transport entre régions ;

- le commerce interrégional est défini à partir d'une répartition des échanges d'une région vers les autres qui dépend de la nature du bien transporté, des prix du marché dans chaque région pour le bien en cause, et des coûts de transport ;

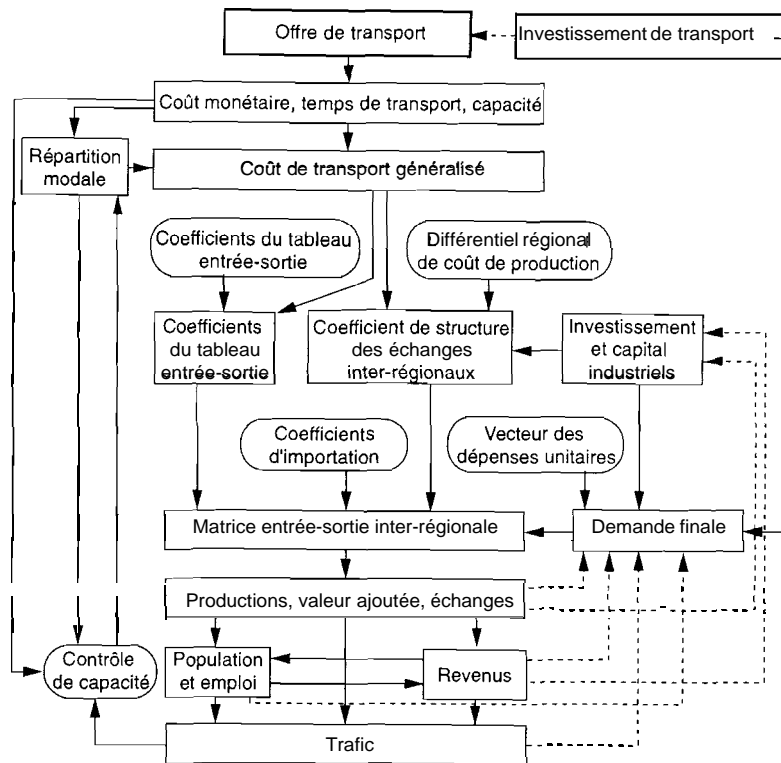
- on tient compte de l'évolution des coûts de transport dans la détermination des coefficients de la matrice input-output et de la part de la valeur ajoutée de chaque bien, ainsi que du produit régional.

- on en déduit la production, par secteur économique, de chaque région, et, également pour chaque région, le revenu, la valeur ajoutée, la demande finale, les exportations et importations vers les autres régions et l'étranger ;

- les migrations de population entre régions sont déterminées en fonction des revenus par tête, des niveaux d'emploi et des distances de transport ;

- l'investissement par région et par secteur est relié aux changements dans les niveaux de production par secteur.

FIGURE 2-6
Organigramme du modèle d'Amano et Fujita



Ce modèle a été utilisé au Japon pour apprécier les effets de la construction d'un pont au Japon entre Hokkaido et Shikoku, deux îles séparées par la mer intérieure, et en particulier pour comparer les conséquences de différentes localisations possibles de ce pont.

À côté de ces modèles assez complètement formalisés et quantifiés, d'autres méthodes sont utilisées pour évaluer les effets attendus d'une modification des conditions de transport.

Une première série de méthodes se fondent sur les variations d'indicateurs d'accessibilité des régions dont la hiérarchie devrait représenter la hiérarchie des effets attendus. Ces indicateurs d'accessibilité sont de la forme :

$$A_i = \sum_j E_j \exp(-\mu c_{ij})$$

expression dans laquelle

- A_i est l'accessibilité de la région i ,
- E_j est le « poids » (PIB, population) de la région j ,
- c_{ij} est le coût ou le temps de transport entre i et j .

L'importance des effets attendu est alors mesuré, en valeur relative, par :

$$\Delta q_i = \Delta \left(\frac{A_i E_i}{\sum_j A_j E_j} \right)$$

et le changement attendu dans la région i sera de la forme :

$$\Delta E_i = \sigma E_i \Delta q_i$$

o étant un paramètre de sensibilité calé sur des études antérieures.

C'est selon des méthodes semblables qu'opèrent Simmonds et Jenkinson (1996) pour évaluer l'impact du Tunnel pour la Manche ou encore Botham (1983) pour évaluer l'impact du programme routier britannique.

D'autres méthodes sont plus qualitatives. Elles visent à décrire de façon raisonnée les conséquences des investissements que l'on cherche à évaluer à travers l'analyse d'informations statistiques et d'interviews d'acteurs dont on recueille l'opinion. C'est ainsi que Spiekerman et Wegener (1996) élaborent des prévisions sur les effets à attendre du Tunnel sous la Manche.

Au total les analyses macro-économiques régionales mettent en évidence un effet de polarisation des infrastructures, en faveur des régions proches et défavorable dans les régions éloignées. Mais ce résultat est en fait contenu dans les hypothèses qui les sous-entendent et qui n'ont souvent d'autres justifications que leur vraisemblance. L'appareillage de modélisation vise surtout à assurer les cohérences d'ensemble, mais ne permet pas de valider ces hypothèses. Aussi, on peut préférer des démarches plus limitées, et davantage fondées sur l'observation des faits.

2.2. Les études empiriques

Elles sont extraordinairement nombreuses, et leur classement ne peut pas ne pas présenter un certain arbitraire. On exposera d'abord les études françaises, puis les études effectuées à l'étranger, avant d'en tirer des conclusions d'ensemble.

2.2.1. Les études françaises

Elles ont surtout porté sur les autoroutes et sur les TGV. Une bibliographie très complète en a été donnée par Denant-Boemont et Gabella (1991).

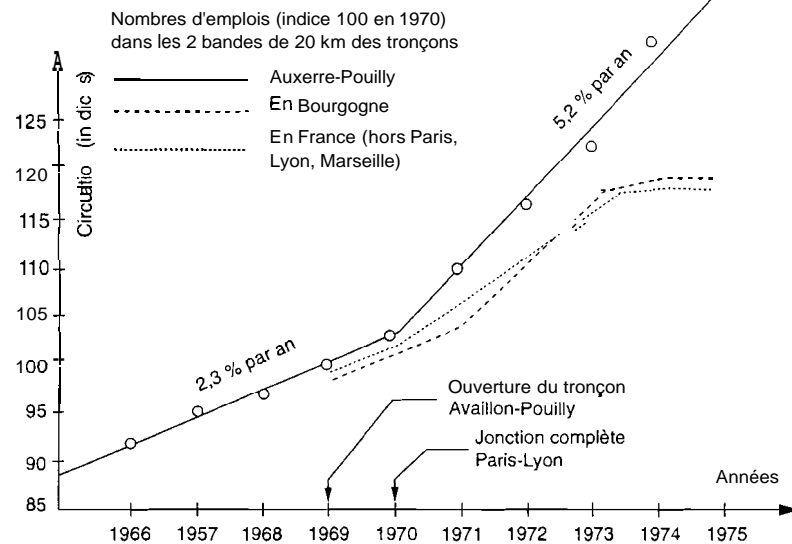
En ce qui concerne le TGV Paris-Sud Est, les études les plus complètes ont été faites par F. Plassard (1990b). Il résulte selon cet auteur que les effets du TGV sur les structures économiques ont été relativement faibles et lents à se manifester. Ce n'est que pour un petit nombre d'entreprises que la présence du TGV aurait constitué un facteur de localisation. Ceci pourrait s'expliquer par la situation de crise qui prévalait durant la période qui a suivi la mise en service du TGV en 1981,

et au cours de laquelle les créations d'entreprises ont été beaucoup plus faibles que par le passé. La décision de déménager ne semble pas résulter de l'existence du TGV, mais une fois qu'elle a été prise, cette existence a alors influé sur la localisation choisie.

La présence d'une ligne à grande vitesse ne peut à elle seule susciter le développement, comme le montre l'exemple du Creusot, ville à longue tradition industrielle qui comptait sur l'arrivée du TGV pour susciter le redémarrage de son activité.

À Lyon, à la suite de l'arrivée du TGV les effets se sont faits sentir, mais de manière plus subtile : c'est le quartier de la Part-Dieu situé à proximité immédiate de la gare qui a connu une forte extension. Le prix des sols y a crû de façon considérable et de nombreuses activités de tertiaire supérieur s'y sont installées. L'effet le plus net du TGV semble avoir été l'extension des aires de marché. Les enquêtes de trafic ont montré la sensibilité des activités de service et surtout de service aux entreprises, qui s'explique par une attitude dynamique des entreprises les plus dynamiques de la zone lyonnaise à l'égard du marché parisien. L'arrivée du TGV a eu pour effet de rattacher le marché parisien à l'aire de marché des entreprises de la région de Lyon. Mais il ne semble pas que, en sens inverse, les entreprises parisiennes aient bénéficié du rapprochement du marché lyonnais. Un autre effet a été le développement du tourisme lié au TGV. C'est ainsi que des compagnies d'autocars redéplient leurs activités en fonction des arrivées aux gares TGV et proposent des circuits de visite organisés complétant le transport par la ligne nouvelle, vers les lieux de visites de la région Bourgogne.

FIGURE 2-7
Évolution des indicateurs socio-économiques
le long de l'autoroute A6

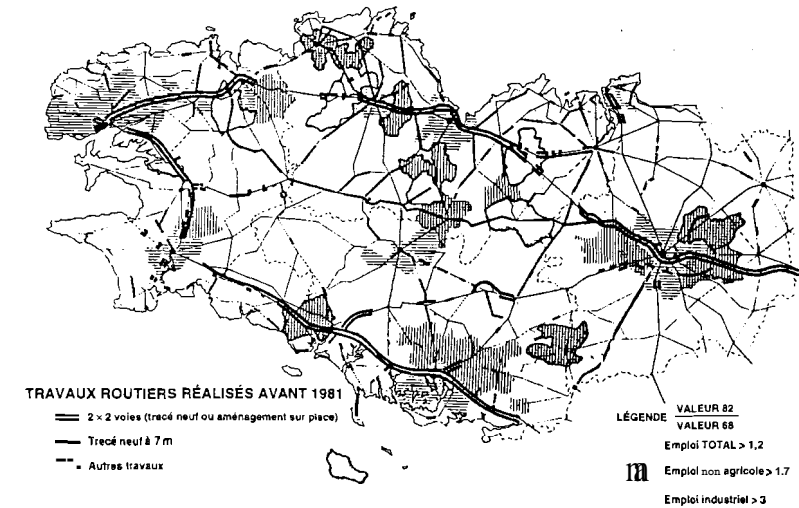


Source : Danzanvilliers, Duchon, Morellet, 1980

En ce qui concerne les autoroutes, la Direction des Routes (SETRA) a mis en place des observatoires statistiques aux alentours des principales infrastructures nouvelles. Elle a pu ainsi évaluer les transformations économiques qui se sont passées à partir de la mise en service de ces infrastructures. Certaines évolutions sont particulièrement significatives, comme celles des emplois dans une bande de 20 km de part et d'autre de l'autoroute A6.

De façon plus générale, la qualité des routes est un facteur du maintien ou de développement de l'emploi. comme le montre, sur la carte jointe, la coïncidence remarquable entre les zones à population croissante et les sections de travaux routiers.

FIGURE 2-8
Localisation des zones en forte croissance de l'emploi



Source : Danzanvilliers, Duchon, Morellet, 1980.

Le tableau ci-dessus met en évidence le même phénomène global, en le différenciant selon les cantons urbains, bénéficiaires, et ruraux, défavorisés.

TABLEAU 2-1
Solde des emplois en fonction du temps d'accès
aux infrastructures routières

	1975-1982				1982-1990			
	< 20 mn	20-40 mn	> 40 mn	Total	< 20 mn	20-40 mn	> 40 mn	Total
Cantons urbains	18 663	3 463	2 933	25 059	12 207	-976	+ 1212	12 443
Cantons ruraux	- 101	1 115	90	1 104	- 4 540	- 3 252	- 162	- 7954
Total	18 562	4 578	3 023	26 163	7 667	- 4228	1050	4 489

Source : ORUS 1995.

Il ressort de ces études que les infrastructures routières sont un facteur du développement économique; mais si le terrain sur lequel elles sont implantées n'est pas convenablement préparé, leurs conséquences peuvent être nulles. L'effet n'est favorable que si des potentialités préexistent, que l'autoroute met alors en valeur. On peut en particulier citer :

- l'existence d'un potentiel touristique. L'infrastructure routière ne crée pas à elle seule le tourisme; elle lui permet simplement de s'exprimer;
- un secteur tertiaire développé, et un tissu urbain équipé permettant l'accueil des entreprises, en général de tertiaire assez élevé, intéressées par les possibilités du transport routier;
- des politiques publiques d'accompagnement, comportant notamment une politique d'accueil aux entreprises, des actions dans le domaine foncier et en matière d'urbanisme (lutte contre l'habitat diffus, création et extension de zones d'activités, préparation de POS cohérents avec l'infrastructure: rsmenibrement des terrains agricoles; valorisation et promotion des zones desservies: coopération intercommunale dans le cas de petites communes aux capacités techniques et financières insuffisantes);
- la disponibilité en main-d'œuvre qualifiée, un tissu industriel sensible au transport, une structure démographique favorable.

Plus récemment Burmeister et Colletis-Wahl (1996) ont analysé les conséquences du TGV Paris-Nord sur les fonctions tertiaires à Lille et à Valenciennes. Les auteurs partaient de l'hypothèse selon laquelle un train à grande vitesse peut surtout agir sur les activités du tertiaire supérieur et sur les entreprises entretenant des relations de réseau (dépendance réciproque des entreprises pour l'accès à des ressources contrôlées par l'autre complémentarité et adaptation). Le résultat des enquêtes menées auprès des entreprises du Nord fait apparaître des effets limités: un certain élargissement des aires de marché, dont les entreprises lilloises ont bénéficié, mais dont elles ont aussi pâti; des effets positifs pour les localisations lilloises, mais plutôt négatifs pour les localisations valenciennes dont la compétitivité et les possibilités de mise en réseau sont réduites. Les auteurs concluent en soulignant l'efficacité des programmes d'accompagnement pour accentuer les effets favorables susceptibles d'être retirés de la nouvelle infrastructure.

Brua (1993) étudie avec une grande rigueur les effets du désenclavement sur les zones à faible densité de population (zones de montagne, zones rurales désertifiées) en France. La conclusion à laquelle arrive le groupe de travail qu'il présidait est que le développement économique ne peut pas venir de la seule réalisation d'infrastructures à haute performance telles les TGV ou les autoroutes, mais d'une mise en œuvre équilibrée et complète de moyens de transport diversifiés, à longue distance mais aussi et surtout régionaux et locaux, sans parler des actions extérieures aux transports.

Plassard (1990a) résume dans un article l'expérience de l'effet des autoroutes sur le développement des régions. Rappelant certains points déjà soulignés (l'absence d'effets automatiques, la nécessité pour les collectivités publiques d'une stratégie de développement fondée sur des objectifs clairs), il souligne que l'effet des infrastructures économiques dépend de la stratégie des acteurs, et distingue à cet égard les entreprises qui cherchent à tirer profit d'une politique de transports qu'ils n'ont pas suscité et cela sans modifier leur structure de production, et celles qui au contraire établissent une synergie entre cette structure de production et le développement des infrastructures, éventuellement en influant sur ce développement (cas des houillères de Blanzac, Plassard, 1985). Battiau (1990) souligne l'effet

de concentration des grandes infrastructures qui aboutit à une massification et à une internationalisation des flux, et développe l'idée de rétrécissement de l'espace avec accumulation de plusieurs infrastructures sur le même axe, avec le développement des effets de pôle (gares de chemin de fer, échangeurs autoroutiers) au détriment des effets de traversée qui prédominent aux temps du réseau routier ordinaire. Quinet (1988) évoque la dualisation de l'espace, qui se subdivise en zones modernes, séparées géographiquement mais reliées par des moyens de transport rapides, qui vivent au même rythme et où les progrès et changements se diffusent très vite, et le reste du territoire, composé de zones séparées et isolées entre elles, à la croissance plus faible.

2.2.2. Études étrangères

Les études empiriques sont très nombreuses aussi à l'étranger. On ne traitera pas celles qui concernent les pays en développement, et on ne retiendra des autres que celles dont les résultats sont les plus sûrs.

Et d'abord les effets du *Shinkansen* (Nakamura et Ueda, 1987). L'étude menée par les auteurs est du type avant-après, tout à fait comparable à celles effectuées en France. Des résultats, on tirera d'abord le tableau illustratif suivant, qui fournit le taux de croissance du revenu par tête selon la zone et la manière dont elle est desservie par le *Shinkansen* et l'autoroute qui le double.

TABLEAU 2-2
Effet du Shinkansen
Taux d'augmentation du revenu par tête (1981-1985)

Dotation en infrastructure	Shinkansen et autoroute	Autoroute seule	Sans l'un ni l'autre	Toutes communes
Commune dont la population accru	23,7	18,2	19,3	19,7
Communes dont la population a décré	10,9	11,6	4,4	5,5
Toutes communes	15,7	14,7	5,7	8,2

Les auteurs se livrent ensuite à une analyse discriminante, qui fait apparaître globalement les principales conditions pour que le passage du *Shinkansen* occasionne une croissance de la population. Il faut que :

- les industries tertiaires fondées sur les échanges d'information tiennent une place importante dans la zone;
- des possibilités d'enseignement supérieur suffisantes existent;
- l'accessibilité à la gare *Shinkansen* soit bonne.

En revanche les conditions suivantes ont un effet négatif sur le développement régional :

- une proportion importante d'industries manufacturières;
- une proportion importante de personnes âgées;
- un temps d'accès élevé à l'autoroute.

Briggs (1982) étudie l'effet du système autoroutier inter-États des USA sur le développement des zones non métropolitaines. Il en déduit que les comtés munis

d'une autoroute, pris dans leur ensemble, ont un taux de croissance de la population et de l'emploi) entre 1950 et 1975 supérieur à la moyenne. Mais l'accès à l'autoroute est loin de suffire à assurer la croissance. L'accès ne conduit pas à la certitude du déclin. Si le tourisme est fortement associé à l'industrie manufacturière et le commerce de gros ne semblent pas lui être nettement associés. En outre son rôle fut d'augmenter les facilités d'accès sur l'ensemble du territoire. Une zone semble pas avoir bénéficié uniquement aux comtés situés à sa proximité.

Mills et Carlino (1989) ont étudié la dynamique de la croissance de la population et de l'emploi des comtés aux USA au cours de la décennie 70. Leur analyse fait intervenir un grand nombre de facteurs tels que le niveau initial de la population et de l'emploi, les taxes, les revenus des ménages, le taux de criminalité, etc., et pour ce qui nous concerne la densité du réseau autoroutier inter-Etats dans le comté. Cette dernière variable apparaît comme très significative pour l'explication des taux de croissance en cause. Comme en concluent les auteurs : « le système d'autoroutes inter-Etats a eu un effet significatif sur la croissance à la fois de la population et des emplois ».

Dodgson (1974) avait mené à bien la même étude pour le Royaume-Uni. Pour un découpage en 66 zones de ce pays, ces auteurs ont mesuré le taux de croissance de l'emploi aux coûts de transport et à d'autres variables telles que l'importance de la structure industrielle, ou le taux naturel de croissance de la population - venue dans chaque zone. L'auteur conclut de son analyse de régression que le signe du coefficient est conforme aux attentes (la croissance s'accroît) mais que les coûts de transport sont faibles) et que ce coefficient est significatif au degré de confiance de 97,5 %. Mais si l'effet est assuré en moyenne, sa valeur dans un cas particulier connaît une forte dispersion car le coefficient de corrélation obtenu est relativement peu élevé ($r = 0,7$).

Simon (1987) analyse les effets de la réalisation du Humber Bridge en Grande-Bretagne. À partir d'une enquête auprès des chefs d'entreprise régionaux, se sont manifestés une gamme d'adaptations et de comportements aux possibilités provenant de la construction du pont. Mais l'essor attendu de la région ne s'est pas produit ; l'impact global a été nul. Les effets sont négatifs du fait que la région était très fortement touchée par la crise pétrolière. Une fois de plus ce résultat que la réalisation d'une infrastructure est une condition nécessaire, mais non suffisante, du développement économique.

Dans le même ordre d'idée, Botham (1986) présente les résultats d'une analyse de régression menée sur un découpage du Royaume-Uni en une trentaine de régions, dans laquelle les taux de variation de l'emploi sur la période 1960-1985 ont été reliés à un certain nombre de variables. Certaines de ces variables - l'accessibilité et représentent par exemple la politique régionale - sont des variables indépendantes. Les autres variables représentent directement la qualité des transports. L'accessibilité est mesurée soit par un indice d'accessibilité, soit par un indicateur du coût de transport. L'accessibilité est spécifiée sous différentes formes, par exemple l'accessibilité aux marchés ou bien l'accessibilité aux marchés non compétitifs, l'accessibilité aux sources d'approvisionnement. Il apparaît que l'accessibilité aux marchés est la variable la plus significative et joue un rôle important dans le processus de l'emploi.

Miller, Cooner et Jameson (1986) ont étudié les effets de l'infrastructure sur le développement des aires urbaines aux USA. Ils ont analysé des expériences telles que le BART (métro de l'aire urbaine de San Francisco) ou le *Country Metrorail* et le *Metromover*. Ces infrastructures ont eu des effets marqués pour

polariser les constructions de bureaux et à un moindre degré les constructions de logements. *Metrorail* et *Metromover* ont donné lieu à des opérations d'urbanisation coordonnées avec la mise en service du système. En conclusion de leur étude, les auteurs insistent sur la nécessité de cette coordination pour obtenir le maximum d'efficacité aux infrastructures en matière de développement local et d'utilisation des sols.

Au total, les effets de localisation sont certes difficiles à quantifier et d'importance variable, mais néanmoins incontestables. Les possibilités de perception de changements sont limitées et limitent l'étendue de la zone où ces effets sont appréciables ; elle est de l'ordre de 10 à 20 km autour des autoroutes, de taille plus limitée autour des gares ferroviaires l'infrastructure. Les modèles macro-économiques régionaux permettent d'évaluer les effets sur une aire plus étendue, mais leurs résultats sont plus fragiles car, comme tous les modèles, ils dépendent étroitement des hypothèses faites au départ.

3. Conclusions

L'analyse de l'effet des transports sur la localisation permet de mieux comprendre les voies de leur influence sur la croissance, vue au chapitre I : l'amélioration des transports suscite la concentration et la polarisation des activités et engendre une croissance accrue par l'effet des externalités positives et des rendements croissants.

Toutefois - les vérifications empiriques permettent ainsi de contrôler et préciser les conclusions des analyses théoriques, sur ce point - les transports constituent un facteur important, mais non unique bien sûr, de localisation des activités. On exprime parfois cette conclusion en disant que les transports sont un facteur nécessaire mais non suffisant de localisation. Mais on peut être plus précis. D'abord, si les transports contribuent à augmenter l'activité dans certaines zones, il y a bien sûr des zones où elle se réduit. On l'a vu à travers les résultats de modèles régionaux. Ce que l'on constate aussi, c'est que dans l'ensemble les effets positifs sont supérieurs aux effets négatifs ; et ceci non plus n'est guère étonnant puisque, pour un investissement rentable, l'effet global toutes régions confondues, lié à sa rentabilité, est positif.

Cette discrimination opérée par les transports, créant des zones avantagées et des zones qui ne le sont pas, se traduit par une polarisation, c'est-à-dire par un renforcement des zones fortes. Mais cette polarisation est sélective. Elle joue à distance modérée, et pour différencier des zones déjà très inégales : les TGV n'ont pas réduit l'influence des métropoles régionales, surtout lorsqu'elles sont loin de Paris. En revanche les réseaux autoroutiers ont développé les métropoles régionales au détriment des pays qui les entourent. Il semble que la polarisation transversale à courte distance ait été plus forte que la polarisation longitudinale à longue distance : le réseau d'autoroutes bretonnes n'a pas vidé la Bretagne de ses activités, mais a contribué à les localiser en son voisinage. Elle joue aussi différemment selon les secteurs d'activité économique. Il semble que les activités tertiaires y soient plus sensibles que les activités secondaires et surtout tertiaires.

Les effets de localisation sont également différents selon les modes de transport. Les effets du chemin de fer, et plus précisément du chemin de fer à grande vitesse sont plus limités géographiquement semble-t-il ; ils sont concentrés autour des gares, et concernent surtout les activités tertiaires (services). Les effets des autoroutes sont plus étendus, ne serait-ce qu'en raison du plus grand nombre des lieux d'échanges, et aussi parce que l'automobile est un « adaptateur universel » selon la formule de G. Dupuy. Ces effets s'étendent à un éventail d'activités plus grand ; en revanche, il semble qu'ils touchent moins profondément l'organisation propre des entreprises, le degré de décentralisation entre les établissements et la structure des rapports avec la clientèle et avec les fournisseurs.

Ces résultats d'ensemble constituent une appréciation statistique des observations les plus fréquentes et non l'énoncé de lois intangibles. Ils fournissent néanmoins une certaine justification a posteriori aux nombreuses revendications des élus locaux qui réclament une amélioration des conditions de transport pour leur circonscription, un « désenclavement » selon le terme consacré. Il n'en reste pas moins que les effets bénéfiques qui pourraient ainsi être obtenus pour leur circonscription, et dont l'apparition n'a rien de systématique, le seraient au détriment d'autres zones géographiques, et c'est en ce sens que les effets structurants des transports peuvent être considérés comme un mythe politique (Offner, 1993).

Car alors les revendications présentées sous l'argument d'efficacité sont en fait des arguments de répartition en l'absence de zones géographiques, qui jouent à l'encontre des autres zones. Toutes satisfaites, les demandes d'investissement de ce type annuleraient leurs effets, au détriment du contribuable qui aurait à payer des dépenses inutiles, et de l'intérêt général qui se trouverait privé d'autres investissements plus utiles. Bien prendre en compte les effets de localisation dans les décisions de transports est donc une tâche difficile sur le double plan de la connaissance et du processus décisionnel. Au plan de la connaissance, il convient d'apprécier avec exactitude ces effets, et on a vu que les données sur cela étaient relativement pauvres. Au plan du processus décisionnel, il faut d'abord se fixer des priorités sélectives et, notamment, ne pas se laisser aller à des tâches redoutables pour des pouvoirs publics, en réponse aux demandes réitérées et insistantes des élus locaux de leur circonscription, de relations continues et réciproques.

BIBLIOGRAPHIE

- K. Amano et M. Fujita (1970). A long run economic growth and transportation facility plans. *Journal of Economic Geography*, 1, 1-10.
- M. Battiau (1991), « Effets des évolutions spatiales de la région parisienne : une approche spatiale », *L'Information Géographique*, 45, 1, 1-10.
- R. Botham (1983). « The road program and regional development », in *Transport, location and spatial theory*, edited by K. Button and D. Gillingwater, Gower.
- R. Briggs (1982). « Interstate highway system and development in non-metropolitan areas », *Transportation Research Record* 812.
- J. Brua (1993) *L'accessibilité des zones à faible densité*, rapport commun CGPC-CGGREF.
- F. Bruinsma, P. Nijkamp et P. Rietveld (1989). « Employment impact of infrastructure investment », in Peschel. *Infrastructure and the space economy*, Springer, Verlag.
- A. Burmeister et K. Colletis-Wahl (1996), « TGV et fonctions tertiaires : grande vitesse et entreprises se service à Lille et à Valenciennes », *Transports Urbains*, oct-déc.
- A. Burmeister et G. Joignaux (1996). *Infrastructures de Transport et territoires*, l'Harmattan.
- CETE Ouest (1990). *Évaluation de la RN137 Rennes-St-Malo*.
- W. Christaller (1933). *Central places in Southern Germany* (traduction de 1966). Prentice Hall.
- COST 317 (1995). *Les effets sociaux économiques du Tunnel sous la Manche*, Commission Européenne, Bruxelles.
- P. Danzanvilliers, C. Duchon et O. Morellet (1980), *Les effets indirects des transports interurbains*, SETRA, Paris.
- L. Denan-Boemont et C. Gabella (1991), *Les effets structurants. Analyse bibliographique*, LET, Lyon.
- J.S. Dodgson (1974). « Motorway investment and sub-regional growth. A case study of the M 62 », *Regional Studies*, 8.
- A. Gaegler et J. March (1978). « Dynamic social and economic effects of the Connecticut Turnpike », *Transportation Research Record*.
- B. Guitard et A. Plaud (1991). *Impacts et mise en valeur des voies rapides*, DDE Morbihan.
- C. Harris (1980). « New developments and extensions of the MRMI forecasting model », *Journal of Regional Science*, 20.
- H. Jayet, J.P. Puig et J. Thisse (1995), *Enjeux Économiques de l'organisation du Territoire*. Document CERAS n° 95-13.
- C. Liew (1985). « Measuring the impact of a transportation system : a simplified approach », *Journal of Regional Science*, 25.
- A. Losch (1940). *The economics of location* (traduction de 1954), Yale University Press.
- K. Mera (1975), *Income distribution and regional development*, University of Tokyo Press.
- J.R. Meyer (1982), *Techniques of transport planning*, The Brooking Institution.
- C. Miller, D. Cooner et R. Jameson (1985). « Role and function of transit in growth management », *Transportation Research Record*.
- E. Mills et G. Carlino (1989). « Dynamics of country growth », in Andersson, Batten, Johansson, Nijkamp, *Advances in spatial theory and dynamics*, Amsterdam.
- H. Nakamura (1985). *Second model for Tokyo Bay Bridge*, miméo, Tokyo.
- H. Nakamura et T. Ueda (1989), *The impacts of the Sinkansen on regional development*. ronéoté, Tokyo.

Les transports et la ville

- J.M. Offner (1993), « Les " effets structurants " des transports : mythe politique, mystification économique », *l'Espace géographique*, n° 3.
- J.P. Orus (1995), « Les conséquences économiques des grandes infrastructures routières », in *Grandes infrastructures de transport et territoires*. O. Heddebaud edit., Actes INRETS, n° 60.
- J. Perreur (1988), « La localisation des unités de production », in *Analyse Économique Spatiale*, sous la direction de C. Ponsard. PUF.
- F. Plassard (1990a), « Axes autoroutiers et développement des régions », *les cahiers Scientifiques du Transport*, Caen.
- F. Plassard (1990b), « Transport et distribution spatiale des activités ». CEMT, Paris.
- A. Politano, C. Roadifer (1989), « Regional economic impact model for highway systems », *Transportation Research Record*, 1229.
- E. Quinet (1988), « Transport et espace économique ». Communication à l'Académie des Sciences Morales et Politiques.
- J. Poulit (1994), *Évaluation de l'utilité économique et environnementale des infrastructures de transport. Influence de la taille de l'agglomération*, mimeo, DREIF.
- R. Prud'homme (1998), *Size, Sprawl, speed and efficiency of cities*. l'Ceil.
- E. Quinet (1992), *Infrastructures de transport et croissance*. Economica. Paris.
- P. Rietveld et P. Nijkamp (1993), « Transport and Regional development », in J. Polak and A. Heertje (1993), *European Transport Economics*.
- M. Savy et P. Veltz (1993), *Les nouveaux espaces de l'entreprise*. DATAR. Éditions de l'Aube, Paris.
- D. Simmons et N. Jenkinson (1996), « Les impacts économiques régionaux du tunnel sous la Manche », in *Infrastructures de transport et territoire*. sous la direction de A. Burmeister et G. Joignaux, l'Harmattan.
- D. Simon (1987), « Spanning muddy waters : the Humber bridge and regional development », *Regional Studies*.
- Spiekerman, Wegener (1996), « The Channel Tunnel and Regional Development : Combining qualitative and quantitative methods », in Quinet and Vickerman, *Econometrics of major transport infrastructures*, Mac Millan.
- Y. Stephanides et D. Eagle (1986), « Highway expenditures and non-metropolitan employment », *Journal of Advanced Transportation*.
- J. Thisse (1988), « La concurrence spatiale », in *Analyse Économique Spatiale*, sous la direction de C. Ponsard, PUF.
- J. Thisse (1993), « Oligopoly and the polarization of space », *European Economic Review*, n° 37.
- G. Treyz (1980), « Design of a multiregional policy analysis model », *Regional Science*, 20.
- D. Peters., J.F. Thisse et I. Thomas (1997), *Transportation networks and the location of human activities*, mimeo, CORE Louvain, Belgique.
- H. Stabenau (1997), « Les nouvelles tendances de la logistique en Europe », rapport pour la CEMT, *Table Ronde*, n° 104, Paris.
- F. Wilson, A. Stevens et T. Holyoke (1982), « Impact of transportation on regional development », *Transportation Research Record*, 851.

Les transports urbains concernent très directement le citoyen. 80 % de nos compatriotes vivent en ville et y effectuent l'écrasante majorité de leurs déplacements. Les problèmes les plus fréquemment cités sont de différents ordres ; les encombrements et les atteintes à l'environnement viennent en tête ; suivent les problèmes de la gestion des transports publics, leurs coûts et la qualité des services qu'ils offrent ; enfin l'harmonie entre les transports et le développement urbain. Les deux premiers points seront examinés en détail dans le cadre général des chapitres V (le coût de transports), VII (les décisions optimales) et IX (la gestion des monopoles). Aussi c'est surtout le dernier aspect qui sera traité ici, celui qui concerne l'insertion des transports dans le développement urbain et leurs relations à la structure urbaine ; il ne préoccupe guère l'opinion publique, à tort car c'est à ce niveau que les autres problèmes prennent leur source.

Pour le comprendre, il faut faire un retour aux caractéristiques économiques de la ville, qui vont à l'encontre des hypothèses usuelles de l'équilibre concurrentiel classique ; elle seront l'objet de la 1^{re} section. La 2^e et la 3^e seront consacrées à la place des transports dans ce cadre, et notamment à la dualité entre les transports et l'utilisation des sols. La quatrième section traitera de quelques problèmes de politique des transports urbains, en particulier le rôle des transports dans la péri-urbanisation et les conditions de la planification urbaine.

1. Les transports et les tendances à l'agglomération

Si l'économie répondait rigoureusement aux hypothèses de l'équilibre concurrentiel classique, il n'y aurait pas de raison que les agents s'agglomèrent. En présence de rendements constants, en l'absence de biens publics et d'externalités, la meilleure organisation est celle de l'étalement uniforme sur le territoire de toutes les activités ; c'est ainsi qu'on minimiserait les coûts de transport et les coûts tout court (« *Backyard capita-*

ism ». L'existence des villes prouve bien qu'il n'en est pas ainsi. En suivant Fujita (1989), Fujita et Thisse (1995) et Duranton (1997), on peut faire le tour des forces qui poussent à l'agglomération des individus et des entreprises, donc à la création des villes, et on verra que leur intensité est liée à la qualité des transports.

1.1. Localisation de ressources naturelles, indivisibilités

La ville peut provenir d'abord de l'existence de ressources naturelles concentrées, une mine par exemple, autour de laquelle s'agglomèrent les industries de transformation. Elle peut résulter d'indivisibilités et d'économie d'échelle dans les transports : un port, un pont, un relais de poste.

1.2. Externalités

De façon plus profonde, la ville est le siège d'externalités. Pour les analyser, il est nécessaire de distinguer entre externalité technologiques et externalités pécuniaires.

ENCADRE 3-1

Externalités technologiques et Externalités pécuniaires

Les premières ne passent pas par le mécanisme des prix : elles se produisent lorsqu'un agent ou une firme influencent directement la fonction d'utilité d'un autre agent ou le profit d'une autre firme : par exemple, en rejetant des déchets dans une rivière, une usine la rend impropre à la pêche et diminue la satisfaction des pêcheurs, ou augmente le coût de production d'une autre usine obligée de dépolluer l'eau qu'elle utilise.

Les externalités pécuniaires correspondent toujours à l'idée qu'un agent bénéficie (ou souffre) de l'action d'un autre agent sans en payer le coût (ou sans en être dédommagé), mais ici le coût ou l'avantage passent par les prix. Par exemple, le coiffeur voit son salaire augmenter, non parce que sa productivité s'est accrue, mais parce que celle des autres secteurs a augmenté et que le niveau général des salaires a cru. Un autre exemple est celui de la rente foncière dont bénéficie un propriétaire du fait de l'implantation de services collectifs en son voisinage.

Les deux se rencontrent dans les agglomérations. Les externalités technologiques comprennent les effets sur l'environnement et les effets de congestion, dont les transports sont de forts générateurs. Les externalités pécuniaires sont par exemple les effets de rentes, et notamment de rentes foncières, dont les mécanismes seront exposés à la section 2.

Les externalités de communication ou d'agglomération sont technologiques ; elles résident dans le fait que la productivité des entreprises est accrue par le voisinage d'autres entreprises, permettant les communica-

tions et l'échange d'informations, selon des mécanismes analogues à ceux qui ont été présentés au chapitre I (section 3-3). Le rassemblement d'entreprises dans un même lieu accroît ces possibilités d'échanges entre entreprises et la productivité de chacune d'elles. Des modèles théoriques présentés dans Fujita et Thisse (1995) font l'hypothèse que la productivité de chaque entreprise est liée à son accessibilité¹ à l'ensemble des autres entreprises, et montrent que plusieurs types d'équilibres peuvent s'établir, aboutissant à un ou plusieurs centres urbains selon les valeurs des paramètres, en particulier le coût des transports.

Le développement des communications entre les départements d'une même firme, développement permis par la télématique moderne, peut d'ailleurs avoir un effet centrifuge : les entreprises maintiennent en zone agglomérée l'implantation des services bénéficiant des effets d'agglomération, rejetant dans des zones où le coût du sol est plus faible ceux de leurs services qui ne tirent pas de bénéfice particulier de la proximité d'autres entreprises.

Les externalités de communication s'appliquent également aux rapports entre consommateurs et traduisent leur valorisation des rapports sociaux.

1.3. Économies d'échelle

La formation des villes peut aussi s'expliquer par l'existence d'économies d'échelles, qui conduisent à concentrer en une même localisation la production d'un bien. Un cas particulier des économies d'échelle est celui des indivisibilités liées par exemple aux équipements publics ou à certaines infrastructures collectives : écoles, universités, équipements hospitaliers spécialisés... En leur présence, des modalités d'organisation de ces activités peuvent permettre de réduire, sans les annuler, les effets négatifs de la dispersion géographique : ramassage scolaire, hôpitaux à domicile ; le développement des télécommunications va dans le même sens pour certaines activités : télé-enseignement, télédiagnostic. Une forme d'économie d'échelle génératrice d'agglomération des activités rejoint les sources de croissance endogène exposées au chapitre I, et peut également recouvrir les externalités d'agglomération décrites au paragraphe 1.2. C'est celle dans laquelle la fonction de production de chaque entreprise f (on suppose que toutes les entreprises sont identiques quant à leurs technologies pour simplifier l'écriture), dont l'effectif est N_j , est de la forme :

$$f(N_j) = N_j^\beta g(N)$$

1. Mesurée par exemple par l'expression

$$A(x) = \int a(x, y) f(y) dy$$

dans laquelle $A(x)$ est l'accessibilité au point x

$f(y)$ est la densité d'entreprise au point y

$a(x, y)$ est la distance entre les points x et y .

expression dans laquelle :

N est la population active de la ville,

$g(N)$ est une fonction croissante de N .

L'hypothèse sous-jacente est que, au niveau de chaque entreprise, les rendements sont constants ; mais la productivité de chaque entreprise dépend d'un facteur qui n'est pas sous son contrôle, à savoir la taille de l'agglomération, ce facteur représentant une externalité d'agglomération. Cependant la fonction de production au niveau de la ville, F , est de la forme :

$$F(N) = N^{\alpha} g(N)$$

Cette formalisation concilie la constance des rendements au niveau de la firme et leur croissance au niveau de l'agglomération, et explique donc, dans un marché d'équilibre concurrentiel pour chaque firme, la tendance à la concentration des activités résultant des externalités positives que chaque nouvelle firme procure aux firmes déjà installées.

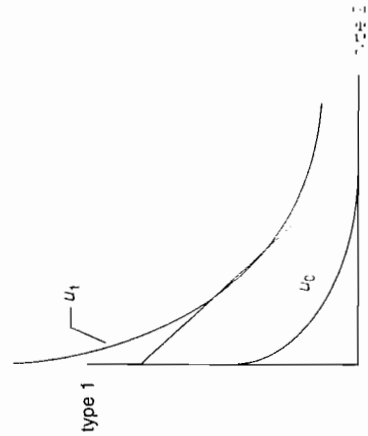
1.4. L'intérêt pour la diversité

La ville permet de satisfaire l'intérêt pour la diversité que ressentent les firmes et les consommateurs.

ENCADRE 3-2

Le goût pour la diversité

Le goût des consommateurs pour la diversité peut s'exprimer et s'illustrer simplement par le schéma suivant, où le consommateur consacre un budget prédéterminé à un type de dépense, la boisson par exemple. S'il y a deux types de boisson le consommateur atteindra le niveau d'utilité u_1 ; s'il n'y en a qu'un, il devra se contenter de u_0 .



On voit que ce goût pour la diversité résulte de la concavité des courbes d'utilité. Si les courbes étaient convexes, le consommateur ne consommerait qu'un bien. Le goût pour la diversité s'exprime commodément à travers une fonction d'utilité du type :

$$U = \left(\sum_{i=1}^n x_i^{\rho} \right)^{\frac{1}{\rho}}$$

L'individu est par ailleurs soumis à une contrainte de revenu :

$$\sum_{i=1}^n p_i x_i = r$$

On voit aisément que si la concavité des courbes U est positive (cas du schéma), alors l'utilité maximale que l'individu peut obtenir augmente avec le nombre de biens : $\rho < 1$.

Si $\rho = 1$ ce goût est nul ; il devient de plus en plus fort à mesure que $\rho \rightarrow 0$. Au niveau de l'entreprise, l'intérêt pour la diversité des intrants s'exprime souvent par l'expression suivante de la fonction de production :

$$F = N^{\alpha} \left(\sum_{i=1}^n y_i^{\rho} \right)^{\frac{1-\alpha}{\rho}}$$

Dans cette expression, α et ρ sont des paramètres, N est l'effectif de la firme, et y_i sont les différents services que l'entreprise utilise dans sa production. L'intérêt de la diversité est mesuré par l'opposé du paramètre ρ . Si $\rho < 1$, alors il vaut mieux, pour une quantité totale d'intrants donnée, que cette quantité soit divisée en une pluralité de biens plutôt que concentrée sur un seul.

Les avantages de la diversité sont simples à concevoir en ce qui concerne le consommateur et l'éventail des biens à sa disposition. En ce qui concerne les entreprises, la diversité des facteurs de production permet une meilleure adaptation aux besoins de l'activité de l'entreprise ; qu'on pense en particulier à la diversité des qualités de service des acheteurs ; à l'intérêt que présente pour une entreprise une panoplie d'avocats couvrant chacun une spécialisation particulière : fiscalité, droit du travail. Bien souvent ces spécialités obéissent à des rendements d'échelle (par exemple d'indivisibilité des personnes dans le cas des services spécialisés), et leur rentabilité n'est obtenue qu'avec une clientèle de taille suffisante.

Une amélioration des transports, permettant d'élargir les aires d'approvisionnement pour les consommateurs et pour les entreprises, accroît leur accès à la diversité et leur utilité ou profit.

1. Si $\rho > 1$ l'effet positif de la diversité proviendra de ce que, contrairement à l'hypothèse du modèle, les utilités diffèrent d'un individu à l'autre. L'ajout d'un bien nouveau aux possibilités offertes n'apportera rien à Mr 1, mais sera apprécié par Mr 2 qui déplacera sa consommation vers ce bien nouveau et en retirera un supplément d'utilité.

1.5. Un processus cumulatif d'agglomération

La combinaison des externalités d'agglomération, des économies d'échelle dans certains secteurs de l'économie et du goût pour la diversité entraîne un processus cumulatif d'agglomération des activités, décrit et modélisé en particulier par Fujita et Thisse (1995). Ces auteurs considèrent une économie qui comporte :

- Des entreprises industrielles à rendement constant dont la productivité croît avec la diversité des services auxquels elles peuvent accéder. Leur fonction de production est par exemple de la forme :

$$F = N^{\alpha} \sum_{i=1}^k x_i^{\beta}$$

N est l'effectif de l'entreprise.

x_i représentent les services auxquels elle peut faire appel.

- Des services à rendement croissant, dû à la présence d'un coût fixe qui fixent leur prix à partir d'un « *mark-up* » (écart relatif entre le prix et le coût marginal) inversement proportionnel à l'élasticité de la demande (comportement d'oligopole de Cournot).

• Des consommateurs dont la fonction d'utilité est croissante avec la diversité des biens consommés :

$$U = \sum_{i=1}^k x_i^{\gamma}$$

Alors le développement du modèle montre qu'un accroissement de la taille de l'agglomération entraîne une augmentation de la productivité et du nombre des services, donc de la productivité et du nombre des entreprises industrielles, donc une diversification et une baisse du prix des biens augmentant la satisfaction des consommateurs : d'où un appel de population, augmentation du besoin de main-d'œuvre et meilleur appariement sur le marché du travail. Un marché du travail plus ample permet une plus grande diversité des offres et demandes d'emploi, donc un meilleur appariement entre offre et demande d'emploi, entraîne une satisfaction accrue des salariés et une productivité plus forte des entreprises, deux facteurs qui s'amplifient l'un l'autre.

Ces effets peuvent être obtenus par une réduction des coûts de transports, qui permet aux entreprises d'accéder à davantage de services, à ces services d'accroître leur clientèle donc de s'installer plus nombreux ; corrélativement les aires de marché s'élargissant, les entreprises trouvent des approvisionnement et des salariés en plus grande diversité et donc mieux adaptés, les consommateurs des biens plus diversifiés.

Les phénomènes d'augmentation des coûts de transport et de congestion des services à coût fixe sont les freins au processus cumulatif ainsi

déclenché. Ce processus aboutit, selon les cas, à plusieurs équilibres : concentration en une seule ville, apparition de plusieurs centres (Fujita et Thisse 1995). On démontre qu'il est équivalent à celui qui résulterait au niveau de l'ensemble de l'agglomération du jeu d'une fonction de production à rendements croissants au niveau de la ville, du type de celle décrite à la section 1-3 (Fujita 1989).

1.6. Les effets de ségrégation

S'il existe de nombreuses forces qui poussent à l'agglomération, il en existe aussi de nombreuses qui poussent à la ségrégation des populations à l'intérieur des zones urbaines.

D'abord les externalités positives d'agglomération, celles qui résultent du voisinage d'entreprises, s'exercent souvent préférentiellement pour des entreprises de mêmes types qui trouvent la même main-d'œuvre qualifiée et les mêmes sous-traitants : l'industrie spatiale française est concentrée à Toulouse, la Silicon Valley est le lieu de développement de l'informatique de pointe ; en Région Parisienne, la banlieue sud connaît une concentration anormale d'universitaires et de chercheurs, qui trouvent un avantage à vivre à proximité les uns des autres, et profitent aussi de l'accessibilité dans cette zone à de nombreux centres de recherche et d'enseignement supérieur.

Cette dernière forme de ségrégation s'apparente à celle qui résulte d'affinités ou de répulsions entre groupes sociaux, le cas le plus typique étant celui des affinités ou répulsions raciales. Plusieurs modèles ont été bâtis pour étudier ces mécanismes raciaux ; leur hypothèse est que la société est composée de 2 groupes : et que l'un d'eux ressent une aversion pour l'autre ; le résultat général est celui d'une séparation des populations par race. La ségrégation géographique liée à l'éducation et aux revenus a été étudiée en particulier par Benabou (1996). À cette ségrégation par les goûts, par les capacités et spécialités professionnelles et par les caractéristiques des populations, s'ajoute une ségrégation par les revenus. Ainsi la création d'une voie nouvelle, qui détériore l'environnement de part et d'autre de cette voie nouvelle, va entraîner une tendance au départ des habitants et une baisse des prix fonciers ; les habitants nouveaux auront dans l'ensemble des revenus plus faibles ; constatant, et éventuellement anticipant cette évolution, les promoteurs vont développer des logements de moindre qualité, et ceux qui s'en porteront acquéreurs ou locataires auront aussi des revenus plus faibles ; un effet de ségrégation par les revenus va se développer. La ségrégation peut venir aussi de similitudes de situation, par exemple d'origine historique. Ainsi, M. Wiel (1997) observe à Brest une tendance des actifs d'âge moyen à aller rechercher des logements individuels en zone périurbaine et à se retrouver ensemble dans ces zones, alors que les personnes âgées qui n'ont ni le goût ni le moyen d'en faire autant, se retrouvent en centre ville en forte proportion. Les habitudes de déplacement et les zones de chalandises des groupes sont également marquées par leur histoire : ceux qui habitaient depuis longtemps

dans une zone boudent les centres d'attraction nouveaux qui sont davantage fréquentés par les nouveaux arrivants.

Le paradoxe de la ségrégation est quel résultat de l'existence des coûts de transport, elle peut être accentuée par la réduction de ces coûts : c'est ainsi l'amélioration des transports urbains qui a provoqué et accentué la partition en zones de logement et zones d'activités.

Les modèles de rente foncière qui ont été obtenus dans la section suivante permettront de préciser certains de ces mécanismes et de mettre en évidence l'effet du coût de transport sur la structure urbaine.

2. Les transports, la rente foncière et la localisation

2.1. Rente foncière et coût de transport

L'analyse du rôle des transports dans la ville passe par le mécanisme de la rente foncière. Celui-ci peut être modélisé sur le plan théorique de multiples façons. On en fera ici une présentation très simple renvoyant à l'annexe 2 pour une formalisation plus poussée, puis on en tirera quelques règles concernant l'effet des transports sur la structure urbaine et enfin on examinera la manière dont l'observation des faits correspond à ces modèles. Ces présentations théoriques seront faites dans le cas schématique où, par suite des effets d'agglomération, les activités sont rassemblées dans un espace restreint unique, le surplus réduit à un point, le reste de la ville étant occupé par des logements. La ville est établie sur une terre agricole dont le prix (loyer par unité de temps et de surface) est R_A . Les habitants situés à une distance x du centre paient un coût de transport $T(x)$ par unité de temps, occupent un logement de superficie $S(x)$ qu'ils paient $R(x)$ par unité de temps et de surface à un propriétaire qui loue le logement selon un mécanisme concurrentiel au plus offrant. Bien sûr,

$$R(x) \geq R_A$$

car sinon les propriétaires auraient intérêt à louer le terrain à un exploitant agricole. On supposera en outre que l'agglomération contient un nombre N d'habitants qu'on supposera, sans précision contraire, fixé. Par ailleurs, chaque habitant consomme une quantité Z du « bien commun » somme agrégée de l'ensemble des autres biens disponibles, et son utilité a pour expression $U(Z, S)$ ou la fonction U convexe, est la même pour tous. Quelles sont les relations entre $T(x)$, $R(x)$ et $S(x)$?

On supposera d'abord que la superficie de logement de chaque habitant est une quantité fixée :

$$S = \bar{S}$$

Alors la densité est constante et égale à \bar{S} et la ville s'étend jusqu'à une distance x_M telle que :

$$\int_0^{x_M} \frac{2\pi}{S_0} x dx = N$$

et l'utilité de chaque individu ne dépend que de Z , le choix entre les différentes localisations, se réduit alors à un arbitrage entre le coût du transport $T(x)$ et le coût du logement $SR(x)$. Chaque individu cherchera à minimiser la somme des deux, c'est-à-dire que pour la localisation qu'il a choisi, on doit avoir :

$$T'(x) + S_0 R'(x) = 0$$

On en déduit :

$$R'(x) = -\frac{T'(x)}{S_0} \quad (3-1)$$

ou encore :

$$R(x) = -\frac{1}{S_0} T(x) + \text{Constante}$$

La constante se détermine en remarquant qu'à la frontière x_M , la rente est R_A :

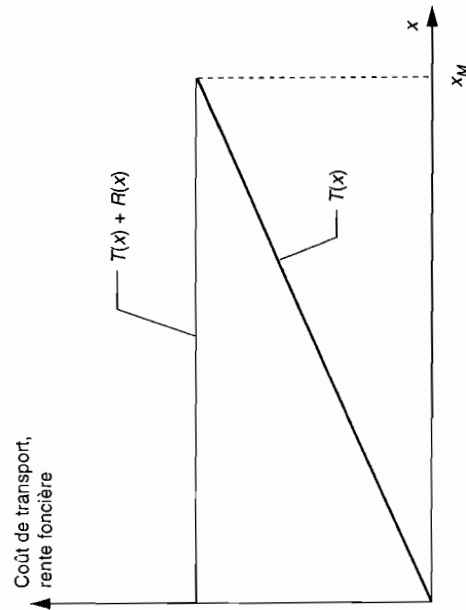
$$\text{Constante} = R_A + \frac{T(x_M)}{S_0}$$

D'où la valeur de la rente foncière :

$$R(x) = -\frac{T(x)}{S_0} + R_A + \frac{T(x_M)}{S_0}$$

FIGURE 3-1

Rente foncière et coût de transport



1. Le CBD (« Center business district ») de la littérature anglo-saxonne.

On remarquera que dans cette situation, les individus sont indifférents quant à leur localisation ; toutes les localisations possibles leur coûtent la même somme ; seul le partage de cette somme entre coût des transports et coût foncier change d'un individu à l'autre.

En chaque point la rente foncière résulte d'un mécanisme d'enchères : c'est la somme la plus élevée que le propriétaire peut recueillir - et que les locataires peuvent offrir - compte tenu de la situation dans les autres points de l'espace urbain. La rente foncière décroît du centre, où elle est maximale, à la périphérie, où elle est égale à R_A .

On peut voir facilement que, si

$$T(x) = ax^\alpha$$

alors

$$RDT = \frac{\alpha}{\alpha + 1} CTT \quad (3-2)$$

expression dans laquelle :

- RDT représente la rente différentielle totale :

$$RDT = 2\pi \int_0^{x_M} (R(t) - R_A) t \frac{dt}{S_0}$$

- CTT représente le coût total de transport :

$$CTT = 2\pi \int_0^{x_M} \frac{t}{S_0} T(t) dt$$

En particulier, si le coût de transport $T(x)$ est proportionnel à la distance ($\alpha = 1$), alors :

$$RDT = \frac{1}{2} CTT$$

On voit apparaître, sur ce modèle simple, les liens entre coût des transports et rente foncière. On peut aussi en déduire les conséquences d'une réduction des coûts de transport, visibles sur la figure 3-1 : la rente foncière diminue également ; la part du revenu des habitants disponible pour d'autres usages que les transports et le logement augmente et donc leur utilité aussi. On peut également en déduire que la ville va s'étendre : d'une part, si des migrations sont possibles, des individus extérieurs à la ville vont venir s'y installer, attirés par la croissance des revenus et des utilités ; d'autre part, en sollicitant le modèle, on peut supposer que les habitants, possesseurs de revenus accrus, vont en partie les utiliser à des logements plus spacieux (effet de revenu).

L'annexe 2 présente une démonstration plus rigoureuse de ces résultats à partir de modèles où la surface du logement est variable et constitue un élément de décision pour chaque habitant et où il existe des catégories différentes d'habitants par leur revenu et leur fonction d'utilité. Les résultats généralisent ceux qui viennent d'être obtenus. En particulier la rela-

tion (3-1) reste valable, sous la réserve que la surface du logement est maintenant variable avec la distance au centre :

$$R'(x) = -\frac{T'(x)}{S(x)}$$

On démontre aussi que la surface des logements augmente quand on s'éloigne du centre. La relation (3-2) reste vraie.

Enfin les résultats concernant les liens entre coût des transports et occupation du sol sont confirmés. Si le coût des transports diminue :

- les limites de la ville s'éloignent
- la rente foncière va se réduire au centre et augmenter en périphérie
- l'utilité des agents augmente
- la taille des logements augmente.

Ces résultats sont aussi ceux qui résultent d'une augmentation des revenus, au moins lorsqu'il s'agit de revenus fixes. Les choses sont un peu plus compliquées lorsque les revenus sont salariaux et que le coût du transport comprend le coût du temps. Alors une augmentation du salaire entraîne à la fois une augmentation du revenu et une augmentation du coût de transport. La taille des logements évolue sous l'influence de deux facteurs contradictoires : elle augmente sous l'influence de l'augmentation du revenu, liée à l'augmentation du salaire horaire, mais elle se réduit en raison de l'attrait pour le temps de loisir dont la valorisation augmente en proportion de l'augmentation du salaire horaire. La résultante dépend de l'importance relative de ces deux effets.

2.2. Les équipements collectifs et les transports de proximité

Dans les modèles examinés jusqu'ici, les infrastructures de transport avaient pour seul rôle de permettre aux agents d'aller de leur domicile à leur travail, au centre de la ville. La réalité est plus complexe et à côté de ces infrastructures radiales, existent des voiries de proximité, qui ont un caractère de bien public et qui présentent pour les habitants de la ville une utilité plus générale liée à la facilitation des rapports sociaux et des activités de proximité. Si, à la suite de Fujita (1989), l'on appelle K le niveau général d'équipement en transports de proximité de la ville (on pourrait différencier géographiquement, par zones, ce niveau d'équipement) et $E(K)$ l'utilité que chaque agent (les agents sont supposés tous identiques) recueille de cet équipement, et qui constitue, à côté des biens Z et S , un des arguments de la fonction d'utilité de l'agent, celle-ci s'écrit :

$$U = U(Z, S, E(K))$$

Dans le cadre de cette formalisation, l'analyse économique s'est attachée à étudier le comportement d'un promoteur qui aurait comme objectif de maximiser le profit qu'il est susceptible de retirer de l'activité économique de la ville, et à le comparer à l'optimum collectif. L'analyse sera présentée dans le cas d'une ville ouverte, le promoteur étant le proprié-

faire des sols. On peut envisager ce promoteur comme un planificateur qui achète le terrain agricole au prix RA , paie les équipements publics K , fait travailler les N habitants en recevant leur produit YN et en leur procurant un niveau d'utilité U donné. Son problème est alors à la fois :

- De minimiser son coût :

$$C(U, N, K) = \min_{X_M, s(r)} \int_0^{X_M} (T(x) + Z(S, U, E(K)) + R_A s(x)) dx$$

avec

$$\int_0^{X_M} \frac{2\pi x}{s(x)} dx = N$$

- Et de maximiser son profit :

$$\max_{N, K} \Pi = \max_{N, K} YN - C(U, N, K) - K$$

On démontre que ce problème de planification aboutit à la même solution qu'un mécanisme décentralisé où le promoteur achèterait la surface de terre optimale, établirait l'équipement public au niveau optimal K^* et attribuerait les terres selon un processus concurrentiel d'enchères. La population se fixerait au niveau optimal N^* , et le promoteur recueillerait la rente $RDT(N^*)$. La concurrence entre les villes amènerait ce profit à s'annuler, et donc à l'équilibre :

$$Y(N^*) - C(U, N^*, K^*) = RDT(N^*) = K^*$$

c'est-à-dire que les développeurs financent le coût du bien public grâce à la rente foncière obtenue par le jeu de la concurrence entre les habitants sur le marché foncier ou par une taxe sur le prix du sol fixée au bon niveau par le développeur. Ce résultat, qui se retrouve dans de nombreux autres modèles, est connu sous le nom de théorème d'Henry George, du nom de celui qui l'a énoncé à la fin du 19^e siècle.

L'équilibre ainsi obtenu correspond à une situation de concurrence entre les villes pour attirer les agents ; alors les villes offrent toute le même niveau d'utilité à l'équilibre, qui ont la même population. C'est aussi un optimum, puisque le coût d'obtention du niveau d'utilité U donné est minimisé. En subventionnant — ou taxant — de manière adaptée les développeurs, on peut obtenir un niveau d'utilité fixé à l'avance.

On a supposé jusqu'ici que les avantages procurés par le bien public qu'est l'infrastructure de proximité étaient indépendants du niveau de fréquentation. S'il n'en est pas ainsi, si l'on est en présence de congestion, alors il faut moduler la taxe en fonction du coût de congestion que la présence des habitants occasionne. On retrouve, par cette approche, le principe général de tarification de la congestion qui sera examiné plus en détail au chapitre VII.

Le modèle précédent peut être appliqué au cas où le paramètre K représente, non pas des services de transport, mais n'importe quel équipement public, ou encore lorsque l'équipement public n'étend pas ses effets sur l'ensemble de la ville, mais seulement sur une zone. Dans ce dernier

cas l'égalité trouvée entre coût de l'équipement K et la rente différentielle doit alors être restreint à la zone.

3. Les modèles et leur adéquation à la réalité

Comment les modèles qui viennent d'être présentés se concilient-ils avec la réalité ? Il ne faut pas de longues analyses statistiques pour constater que la forme des villes n'est pas circulaire, ou encore que les valeurs foncières ne décroissent pas régulièrement quand on va du centre à la périphérie.

Cette inadéquation s'explique par de nombreuses raisons. Ainsi, le sol n'est pas isotrope et uniforme comme le postule la théorie. Il y a des obstacles naturels, des zones où la construction est impossible, du relief, les fleuves... De même, il est clairement irréaliste de supposer que le centre est réduit à un point où se concentrent toutes les activités. Il y a des activités de proximité dont il est évident qu'elles se répartissent sur l'étendue de la ville (commerces de détail par exemple). D'une manière générale le modèle monocentrique ne comporte pas de théorie de localisation des activités.

Ce modèle, tout au moins dans la version qui en a été présentée, exclut l'existence de biens publics locaux (crèches, parcs de loisirs, écoles) qui modifient l'attrait des localisations et donc la structure des valeurs foncières. Enfin, le modèle est un modèle d'équilibre, c'est-à-dire qu'il suppose que tous les ajustements engendrés par les paramètres exogènes ont eu le temps de se réaliser. C'est une situation qui n'est jamais réalisée ; et les analyses théoriques font apparaître que les cheminement aboutissant à l'équilibre sont rien moins qu'assurés ; il peu y avoir plusieurs équilibres, ils ne sont pas forcément stables. Devant cette situation on peut adopter différentes attitudes. La première consiste à construire des modèles qui prennent en compte les caractéristiques précédentes, notamment l'anisotropie de l'espace, les phénomènes dynamiques d'ajustement temporel, l'explication des mécanismes de localisation des activités productrices. Ces modèles, qui visent à expliquer à la fois l'utilisation des sols et les flux de transports, peuvent être calés sur des situations réelles. Ils seront présentés au chapitre IV, relatif à la demande. Disons toutefois dès maintenant que leur multiplicité est une manifestation de ce que la perfection n'est pas encore atteinte, et de loin. Il est difficile de choisir entre les jeux d'hypothèses théoriques qui caractérisent chacun d'eux ; leur calibrage nécessite des données pas toujours disponibles.

La deuxième est de retenir la conclusion qualitative fondamentale que dégage le modèle monocentrique, la liaison entre coût de transport et rente foncière, et d'examiner si cette conclusion est vérifiée au moins en sens. Ceci devrait se traduire d'abord par des transformations dans la valeur et l'utilisation des sols résultant de transformations dans le système de transport, et ensuite par une relation positive à l'intérieur d'une agglomération entre la valeur d'un terrain et la qualité de sa desserte, définie par exemple par son accessibilité (cf. note n° 1 de la page 65), les analyses

statistiques s'efforçant de vérifier ces grandes conclusions fournissent des résultats mitigés.

Citons à cet égard quelques études récentes :

Mills (1993) et Gravel *et al.* (1996) ont utilisé la méthode des prix hédoniques¹ pour analyser les facteurs du prix des terrains.

Mills étudie les prix de location des bureaux dans Chicago. Les variables qu'il fait intervenir concernent à la fois la présence de services divers à proximité immédiate des bureaux, et la qualité de l'accessibilité des bureaux. Il en résulte que cette accessibilité intervient certes, mais qu'elle n'est qu'un des facteurs de formation des prix.

Gravel *et al.* recherchent les mécanismes de formation des loyers moyens dans différentes agglomérations du département du Val d'Oise. Il apparaît là aussi que la bonne desserte de l'agglomération n'est qu'un des éléments du loyer, les autres étant relatifs notamment à la criminalité dans la ville, ou au taux de réussite scolaire.

Y. Huang (1997) étudie le lien entre densité et accessibilité, sur une cinquantaine de quartiers de Washington. Il trouve une relation très significative, aussi bien pour la densité de logements que pour la densité des emplois. Mais les coefficients de corrélation sont relativement faibles (de 0,5 à 0,7), montrant que d'autres facteurs interviennent. S. Song (1996) procède à des études similaires sur des zones urbaines du Nevada et aboutit à des résultats analogues.

D'autres études sont relatives aux effets d'une modification du système de transports sur le prix et le mode d'occupation des sols ; l'analyse de ces effets peut s'effectuer à travers soit des études avant-après, soit des études comparatives entre zones différentes d'une même agglomération.

Des études du premier type ont été effectuées par le Certu (Certu 1996) pour analyser les conséquences d'une création de transport en commun en site propre (T.C.S.P.) une dizaine de créations de ce type sont examinées, et on y observe à la fois les prix, changement et intervention et sur l'occupation des sols. Les résultats sont flous. Comme le résume l'auteur, la réalisation d'un T.C.S.P. ouvre la possibilité d'une modification dans l'utilisation des sols, et d'une augmentation de leur valeur, c'est un révélateur, un accélérateur et un amplificateur de tendances préexistantes. Mais cet effet est limité dans l'espace au voisinage des arrêts ; il n'est pas systématique, et son apparition dépend en particulier des actions d'accompagnement menées par les autorités locales en matière d'urbanisme.

Deboulet et Renard (1992) ont effectué en région Ile-de-France des analyses comparatives des prix des sols en fonction de leur commodité d'accès à la zone centrale. Là aussi les résultats sont mitigés. Il y a une relation entre l'accessibilité d'un logement (défini comme le temps de déplacement en T.C. du centre de la commune à l'Opéra) et le prix du m² ; mais cette relation n'est pas stricte, l'image sociale de la commune concerné est un facteur important pour les prix du sol les plus élevés correspondant aux communes dont les habitants utilisent surtout l'automobile, l'accessibilité T.C. est à remplacer par la distance kilométrique, meilleur reflet des conditions du transport automobile².

Pour achever ce tour d'horizon rapide de quelques études relatives à l'effet des transports sur l'urbanisation, rapportons les dires d'experts et la pratique du marché immobilier, qui mentionnent quelques règles simples, et en particulier le fait

que la réalisation d'une nouvelle autoroute entraîne, surtout au voisinage des échangeurs, une appréciation des terrains. Par ailleurs, le niveau des prix fonciers détermine l'utilisation des terrains (commerces, logements...) selon le mécanisme suivant : pour chaque utilisation possible, il y a un prix maximum qui peut être payé ; l'utilisation retenue est celle qui procure le prix le plus élevé.

Au total, la vérification expérimentale de l'enseignement majeur du modèle présenté, la relation entre valeur du sol et coût de transport, donne des résultats beaucoup moins nets que ce que le modèle impliquerait. Il faut semble-t-il y voir deux raisons. La première tient à ce que le prix du sol dépend de nombreux facteurs, et le transport n'est que l'un d'eux. Sont également importants : le voisinage, les hasards historiques, les effets de mode.

L'autre raison tient à ce que le marché de concurrence que stipule le modèle monocentrique n'est pas réalisé. Les pouvoirs publics interviennent, non seulement en tant que régulateur comme on vient de le voir, mais aussi en tant qu'acteur, en suscitant des opérations d'urbanisme ou en y participant ; dans ces cas, la raison d'être de l'action est précisément de contrarier le jeu du marché. Ces dernières considérations conduisent à examiner quelques problèmes de la politique des transports en ville.

4. Politique des transports et politique urbaine

Les transports sont un des facteurs importants qui modèlent l'urbanisme, à travers leur influence sur les valeurs foncières. Cette constatation emporte de nombreuses conséquences en termes de politique urbaine, et on va dans ce qui suit en analyser quelques unes.

4.1. La planification urbaine

La première concerne les modalités de la planification urbaine. Celle-ci est souvent conçue comme une démarche séquentielle, qui part d'un parti d'urbanisme souhaité, figuré par exemple dans un schéma directeur. Puis une fois ce parti arrêté, on en déduit les flux de transports résultant des localisations figurant dans ce parti, et les infrastructures nécessaires à l'écoulement des flux de transport ; ensuite on programme ces infrastructures au fil des années en fonction de leur urgence mesurée par exemple par leur taux de rentabilité. L'analyse économique précédente montre que cette procédure est insuffisante ; chaque fois qu'en réalise une infrastructure, elle réagit sur les prix fonciers et sur les localisations et va entraîner un développement urbain dont on ne peut assurer qu'il sera conforme au Schéma que l'on veut réaliser.

Certes, on peut imposer de force ce Schéma Directeur à travers la réglementation de l'urbanisme. Mais il faut alors être bien conscient de ce que, si cette réglementation est très contraignante – c'est-à-dire si elle veut imposer un urbanisme très différent de l'urbanisme naturel – alors les forces

1. Pour des rappels sur les prix hédoniques, voir le chapitre V, section 5.1.

2. On y retrouve là le thème développé par G. Dupuy autour de l'automobile « Univer-sal adapter » (Dupuy, 1995).

poussant à l'enfreindre ou à la modifier seront très fortes, et peuvent l'emporter. La démarche de planification urbaine devrait ainsi être itérative et prendre en compte les deux sens de la causalité transport \leftrightarrow urbanisme, et non pas le seul sens urbanisme \rightarrow transport. Les outils permettent de le faire sont imparfaits mais en train de se développer, comme on l'a indiqué plus haut en citant les modèles conjoint trafic-localisation qui seront présentés au chapitre IV, ou les résultats des études hédoniques de prix du sol.

4.2. La péri-urbanisation

La taille et la structure des villes connaissent selon les pays et les époques de nombreuses transformations (Gannon, 1996). Le phénomène auquel on assiste actuellement dans de nombreuses villes françaises, à commencer par la région Ile-de-France, est celui de la péri-urbanisation, dans lequel, à population de l'agglomération à peu près stable, on constate un étalement de la ville au détriment des zones rurales, avec urbanisation des zones périphériques, uniformisation des densités résidentielles et dispersion des zones d'emploi et d'activités (Le Seannic, 1997).

F. Gannon (1996) et M. Wiel (1997) ont analysé les raisons de ce mouvement qui sont dominées par la dynamique du marché de l'habitat, marqué lui-même par l'attrait de la maison individuelle et l'abondance de l'offre de terrains périurbains. Le résultat de cette péri-urbanisation est un accroissement des trafics et un allongement des parcours entraînant la réalisation d'infrastructures, qui elles-mêmes poussent au développement de la péri-urbanisation qu'elles étaient chargées d'accompagner. Cette péri-urbanisation a de nombreuses conséquences d'ordre sociétal, dont l'analyse et la description sortent du champs de l'économie. Cette discipline peut toutefois évaluer les conséquences en termes d'efficacité et de distribution. De ce point de vue, il faut remarquer que la péri-urbanisation, c'est-à-dire l'étalement de la ville, est accentuée par la faiblesse des coûts du transport urbain. Comme on le verra au chapitre VII, ceux-ci sont tarifés bien en dessous du coût marginal social, non seulement pour les transports en commun, mais aussi et surtout pour l'automobile.

4.3. Problèmes de répartition

L'efficacité est un critère important pour l'économiste ; mais dans les transports urbains plus que dans les transports interurbains, les effets de répartition ne peuvent être éludés. D'abord parce qu'ils sont massifs : le passage d'une autoroute entraîne des nuisances élevées et permanentes et le choix du tracé va donc concerner fortement les riverains. Ensuite parce que les choix et priorités concernant la politique d'urbanisme (développer l'emploi, porter remède aux quartiers en difficulté...) échappent à l'économie ; c'est aux responsables politiques d'une agglomération de définir l'ensemble de valeurs et d'objectifs sociétaux propres à leur agglomération, qui s'imposent à l'analyste et ont leur traduction en termes de redistribution et de répartition. L'économiste doit alors aider les décideurs à évaluer les conséquences

de leurs choix. Ainsi Wiel (1997) fait apparaître les effets de la péri-urbanisation en termes de ségrégation sociale et générationnelle. La ségrégation sociale résulte de la concentration de l'habitat collectif locatif au voisinage du centre, alors que la périphérie voit se développer un habitat individuel de proximité. La ségrégation générationnelle, résultant de l'attrait de la péri-urbanisation pour les actifs, entraînant un vieillissement autour du centre.

Un autre exemple d'effets de redistribution réside dans les conséquences de la création d'une nouvelle infrastructure en zone urbaine. D'abord, autour des échangeurs, vont se mettre en place des centres commerciaux, qui entraînent un déclin des commerces traditionnels du centre, avec soit un changement dans la nature des activités du centre, soit un déclin de celui-ci. Ensuite le long de la voie en dehors des échangeurs, la dégradation de la qualité de vie entraînera un exode des ménages les plus fortunés, et une paupérisation du quartier, accrue par le fait que les futurs logements qu'on construira sur cette zone seront adaptés aux revenus de ces habitants, donc de qualité moindre.

4.4. Le fédéralisme fiscal

Les impôts sont levés à certains niveaux de fédération des collectivités territoriales (en France on rencontre l'empilement commune-département-région) et utilisés à la réalisation directe d'équipements collectifs ou au subventionnement d'autres collectivités territoriales, selon des règles de nature constitutionnelle. L'ensemble peut induire des déviations dans le comportement des autorités publiques qui accentuent les imperfections de la politique des transports. Wiel (1997), à titre empirique, et Henderson (1988) sur le plan théorique, ont étudié ce phénomène.

Par exemple, la rivalité des communes pour attirer les entreprises les conduit à réduire les zones offertes aux logements au profit des zones d'activités, qui rapportent davantage de ressources fiscales, et leur concurrence sur ce plan entraîne un éloignement de l'habitat, un facteur de péri-urbanisation. Cette péri-urbanisation peut aussi être accrue si par exemple les infrastructures dont profitent les migrants quotidiens pour leurs déplacements domicile-travail sont financés par une collectivité autre que les communes qui profitent de leur présence, par exemple le Département ou la Région. Alors, l'agglomération ne paie plus le coût des infrastructures et les communes qui la constituent n'ont plus de frein à leur étalement. Une autre situation de sous-optimalité est celle où le désenclavement d'une commune passe par des voies situées sur le territoire d'une autre commune ne présentant pas d'intérêt pour elle. À l'inverse, lorsque les populations se déplacent du centre, bien équipé en services publics (écoles, crèches...) devenus inutiles, vers la périphérie où les équipements n'existent pas, il en résulte un coût collectif non intégré dans les décisions individuelles, qui conduit à une perte d'efficacité économique.

Ceci ne se produirait pas en présence d'une autorité unique qui assurerait le développement de l'ensemble de l'agglomération en ayant en

charge la totalité des équipements publics, comme l'ont montré les analyses présents en 2.4. Ces résultats plaident pour l'instauration d'une autorité d'agglomération qui ait compétence sur l'ensemble des équipements publics, dont bien sûr les transports, et sur l'urbanisme.

BIBLIOGRAPHIE

- R. Benabou (1996), « Equity and Efficiency in Human Capital Investment : the local connection », *Review of Economic Studies*, 63.
- A. Bonnafous, « Prospective de la mobilité urbaine », dans Y. Bussière et A. Bonnafous, *Transport et étalement urbain : les enjeux*, Éditions du Programme Rhône-Alpes de Recherche en Sciences Humaines.
- Certu (1996), *Évaluation des Transports en commun en site propre*, Lyon, étude réalisée par O. Cormier.
- A. Deboutet et V. Renard (1992), « Les plus values induites par les infrastructures de transport », *Rapport de Recherche*.
- G. Dupuy (1995), « The automobile system : a territorial adapter », *Flux* n° 21, July-Sept. 1995.
- G. Duranton (1997), « La nouvelle économie géographique », *Économie et Prévision*, n° 131.
- M. Fujita (1989), *Urban economic theory*, Cambridge University Press, Mass.
- M. Fujita et J.F. Thisse (1995), « Economics of agglomeration », *Rapport de recherche CERAS*, n° 95-15.
- F. Gannon (1996), *La péri-urbanisation*, Communication au séminaire d'analyse économique MELTT Paris (ronéotypé).
- N. Gravel, M. Martinez et A. Trannoy (1996), « Évaluation des prix hédoniques du logement dans les communes du Val-d'Oise de plus de 10 000 habitants », Rapport du centre THEMA Université de Cergy-Pontoise.
- J. V. Henderson (1988), *Urban Development. Theory, fact and illusion*, Oxford University Press, 242 p.
- J. V. Henderson (1985), *Economic theory and the cities*, Academic Press, Inc. Orlando, Floride, 274 p.
- Y. Huang (1997), « Identifying the link between transportation and Land use Density with Accessibility », Communication à la 76^e réunion annuelle du Transportation Research Board.
- T. Le Jeannic (1997), « Trente ans de péri-urbanisation : extension et dilution des villes », *Économie et Statistique*, n° 307.
- E.S. Mills (1993), « The spatial pattern of Office asking rents in the Chicago Metropolitan Area », in *Does Economic Space Matter*, edited by H. Ohta et J.F. Thisse, St Martin's Press.
- S. Song (1996), « Alternative Accessibility Measures », *Land Economics*, November 474-482.
- M. Wiel (1997), « Comportements de mobilité et évolution de l'organisation urbaine », Rapport de Recherche GERME, Brest.
- H. Zoller (1988), « L'espace résiduel et le prix du logement, in *Analyse économique spatiale*, sous la direction de C. Ponsard, PUF.

DEUXIÈME PARTIE

La demande et les coûts

La demande de transport

La demande de transport et ses déterminants peuvent être analysés à un niveau global, par exemple pour l'ensemble d'un pays tous modes confondus ou par mode de transport ; cet angle de vue a été présenté au chapitre I. Elle peut aussi être analysés à un niveau spatial et plus microscopique, celui d'un réseau géographique donné ; elle aura alors pour but de déterminer les niveaux de trafic par mode et pour chaque mode, sur chacun des arcs du réseau, en fonction des conditions de l'offre de transport, c'est-à-dire de la qualité de service et des prix. C'est le point de vue qui sera adopté ici.

L'analyse de la demande est fondée sur la théorie du consommateur, qu'on rappellera dans une première section. Puis on présentera la manière dont cette théorie est utilisée dans les méthodes de prévision des trafics voyageurs, fondées sur le modèle à quatre étapes, et on analysera quelques autres voies de modélisation en cours de développement. Une section sera ensuite consacrée à la prévision des trafics marchandises, qui a été moins étudiée. Enfin, en conclusion on prendra une vue de synthèse sur la portée et les limites des méthodes, ainsi que leurs implications en terme de politique des transports.

1. Les bases théoriques

1.1. Le comportement du consommateur

L'annexe 3 rappelle les fondements de la théorie du comportement du consommateur. On en résumera ici les principes et les résultats utiles à la compréhension des modèles de trafic.

On fait l'hypothèse que chaque consommateur est muni d'une fonction d'utilité qui dépend de la quantité qu'il consomme de chaque bien. Cette fonction n'est définie qu'à une fonction monotone croissante près. Elle est croissante : l'utilité augmente avec les quantités consommées ; elle est

concave. Le comportement du consommateur peut se traduire par le programme suivant :

$$\max U(x_1, x_2, \dots, x_n) \\ \sum p_i x_i \leq r$$

avec

Le consommateur cherche à maximiser son utilité sous la contrainte de son revenu r .

La solution de ce programme (voir Annexe 3) aboutit aux lois de demande, qui ont une expression particulièrement simple lorsque U est de forme dite « quasi linéaire », par exemple dans le cas de trois biens :

$$U = x_0 + u(x_1, x_2)$$

On utilise souvent cette expression en mettant sous x_0 un composé de l'ensemble des biens sauf ceux sur lesquels se concentre l'attention, et on normalise son prix à l'unité. Alors on démontre que les lois de demande de x_1 et x_2 ne dépendent pas du revenu r . Par ailleurs la valeur maximale atteinte alors par U , qui est une fonction des paramètres p_i et r et qu'on appelle l'utilité indirecte, a alors une expression simple :

$$V(p_1, p_2, r) = r + v(p_1, p_2)$$

La théorie classique du choix du consommateur dont les grands principes viennent d'être rappelés doit être développée dans deux directions pour son application aux transports.

1.2. Extension aux choix discrets

Il s'agit des situations dans lesquelles, pour reprendre l'exemple des trois biens ci-dessus, le consommateur ne peut consommer que l'un des deux biens 1 et 2 ; c'est le cas du choix du mode de transport. Alors, tout se passe comme si l'individu avait deux fonctions d'utilité :

$$U_1 = U(x_0, x_1, 0) \\ U_2 = U(x_0, 0, x_2)$$

et les fonctions d'utilité indirecte associées :

$$V_1(p_1, r) \text{ et } V_2(p_2, r)$$

Si les utilités U_1 et U_2 dépendent aussi de paramètres de qualité q_1 et q_2 afférents aux biens 1 et 2, alors ces paramètres se retrouvent dans les fonctions d'utilité indirectes.

Un cas particulier de modèle de choix discret très utilisé dans les transports est celui où l'utilité de chacun des biens 1 et 2 est aléatoire¹. Alors les utilités indirectes retirées de chacun des biens 1 et 2 sont :

$$V_1 = r + v_1(p_1, q_1) + \varepsilon_1 \\ V_2 = r + v_2(p_2, q_2) + \varepsilon_2 \quad (4-1)$$

1. Ceci peut se produire soit parce que les goûts du consommateur sont changeants, soit parce qu'ils dépendent de paramètres qu'on n'a pas réussi à identifier et dont on connaît seulement la répartition à travers la population considérée.

On ne peut plus déterminer *a priori* le choix précis de l'individu entre 1 et 2, mais simplement la probabilité de son choix, ou la proportion dans la population concernée des choix en faveur de 1 ou 2. Par exemple :

$$\Pr(1) = \Pr(V_1 \geq V_2) = \Pr(\varepsilon_1 - \varepsilon_2 \geq v_2(p_2, q_2) - v_1(p_1, q_1))$$

La connaissance des lois que suivent ε_1 et ε_2 (ou simplement de celle que suit leur différence) permet de déterminer $\Pr(1)$ si l'on connaît v_1 et v_2 . Deux lois sont spécialement utilisées pour les variables aléatoires ε (Benakiva et Lerman, 1991) :

- la loi normale centrée, qui aboutit au modèle « probit » :

$$\Pr(\varepsilon \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2\sigma^2}\right) dt$$

Si ε_1 et ε_2 suivent chacune une loi normale de moyenne nulle et ont des variances σ_1^2 et σ_2^2 et une covariance σ_{12} , alors $(\varepsilon_1 - \varepsilon_2)$ suit une loi normale centrée de variance :

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}$$

et :

$$\Pr(1) = \Pr(\varepsilon \leq v_2 - v_1) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{v_1 - v_2} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt$$

- La loi de Gumbel¹ de paramètres (μ, m) qui aboutit au modèle « logit » :

$$\Pr(\varepsilon \leq x) = \exp(-\exp(-\mu(x - m)))$$

La moyenne de cette loi est :

$$\left(m + \frac{\gamma}{\mu}\right)$$

γ étant la constante d'Euler ($\gamma = 0,577 \dots$). Sa variance est :

$$\frac{\pi^2}{6\mu^2}$$

1. La loi de Gumbel est un cas particulier de la loi des valeurs extrême généralisée

$$F(\varepsilon_1, \dots, \varepsilon_n) = \exp(-G(\exp(-\varepsilon_1), \dots, \exp(-\varepsilon_n)))$$

où G est définie par les propriétés suivantes :

- G est positive ;
- G est homogène de degré μ ;
- La N -ième dérivée partielle de G est non négative si N est impair, non positive si N est pair, alors, si les aléas sont indépendants et suivent des lois de ce type :

$$\Pr(1) = \frac{\exp(v_1 G_1'(\exp(v_1), \dots, \exp(v_n)))}{\mu G(\exp(v_1), \dots, \exp(v_n))}$$

Cette dernière expression est connue sous le nom de « *logsum* ». Elle représente, à une constante $\gamma\mu$ près, l'utilité moyenne retirée du choix.

Une propriété particulière de cette loi, dite IIA dans la littérature anglo-saxonne, (Independence from Irrelevant Alternatives), est que :

$$\frac{\Pr(1)}{\Pr(2)}$$

est indépendant des autres choix possibles.

Ce résultat contredit certaines intuitions, illustrées dans le domaine des transports par le classique paradoxe « bus bleu-bus rouge » : supposons qu'un usager obéissant à un modèle logit ait le choix entre deux moyens de transport, l'automobile et le bus, qui présentent pour lui la même utilité certaine : $v_A = v_B$.

Alors ses probabilités de choisir l'auto ou le bus seront toutes deux égales à 1/2. Supposons maintenant qu'on subdivise le mode « bus » en deux, en peignant la moitié des bus en rouge, l'autre en bleu – ou de façon plus réaliste, qu'on introduise un autre mode, très semblable au bus, par exemple un trolley présentant la même utilité pour l'usager. Alors l'usager se trouve en présence de trois modes dont les utilités sont pour lui égales, et la propriété IIA indique que la probabilité de l'automobile tombera à 1/3, résultat peu vraisemblable. Pour pallier cet inconvénient, il faut abandonner l'hypothèse d'indépendance des aléas. On peut le faire sans retomber dans les difficultés de calcul impliquées par le retour au modèle probit général ou à un modèle des valeurs extrêmes généralisées quelconque, dans l'hypothèse de décisions emboîtées.

Pour le voir, considérons le choix à la fois d'une destination d parmi D possibilités et d'un mode de transport m parmi M possibilités, et supposons que l'utilité totale soit de la forme

$$U_{dm} = v_d + v_m + v_{dm} + \varepsilon_d + \varepsilon_m + \varepsilon_{dm}$$

alors il y a corrélation entre les U_{dm} :

$$\text{cov}(U_{dm}, U_{d'm'}) = \begin{cases} \text{var}(\varepsilon_d) & d = d' \\ \text{var}(\varepsilon_m) & m = m' \\ 0 & \text{autrement} \end{cases}$$

Supposons que :

$$\text{var}(\varepsilon_m) = 0$$

À destinations différentes les aléas ne sont pas corrélés, mais ils le sont à destinations identiques. On peut grouper les possibilités en partitions indépendantes définies par la destination.

Supposons que ε_{dm} suive une loi de Gumbel $G(\mu m, 0)$ et que ε_d soit tel que :

$$\max_m (\varepsilon_d + \varepsilon_{dm})$$

suive une loi de Gumbel $G(\mu d, 0)$. Cette hypothèse *ad-hoc* permet de mener les calculs jusqu'au bout de manière aisée (elle est rendu admissible par le résultat que le maximum de plusieurs variables aléatoires suivant une loi quelconque est une variable aléatoire dont la loi de probabilité tend vers une loi de Gumbel

Si ε_1 et ε_2 sont indépendants et suivent chacun une loi de Gumbel de paramètres $(\mu, 0)$, alors $\varepsilon = \varepsilon_1 - \varepsilon_2$ suit une loi logistique définie par :

$$\Pr(\varepsilon \leq x) = \frac{1}{1 + \exp(-\mu x)}$$

Alors dans le cas particulier du choix discret précédent défini par (4-1) :

$$\Pr(1) = \Pr(\varepsilon \leq v_2 - v_1) = \frac{1}{1 + \exp(\mu(v_2 - v_1))} = \frac{\exp(\mu v_1)}{\exp(\mu(v_1) + \exp(\mu v_2))}$$

Toujours si ε_1 et ε_2 sont indépendants et suivent chacun une loi de Gumbel de paramètres $(\mu, 0)$, alors : $\max(V_1, V_2)$ suit également une loi de Gumbel dont les paramètres sont :

$$\mu \quad \text{et} \quad \frac{1}{\mu} \log(\exp(\mu v_1) + \exp(\mu v_2))$$

On peut étendre la situation au cas où il y a plus de deux alternatives. Soit n le nombre des alternatives, dont les utilités sont, en omettant le terme r de la formulation précédente :

$$V_i = v_i(p_i, q_i) + \varepsilon_i$$

Si l'on connaît la fonction de répartition des ε_i , on peut déterminer la probabilité de choix de l'alternative 1. Ce sera :

$$\Pr(1) = \Pr(U_1 \geq U_2 \quad \text{et} \quad U_1 \geq U_n) \\ = \int_{\varepsilon_1 = -\infty}^{+\infty} \int_{\varepsilon_2 = -\infty}^{v_1 - v_2 + \varepsilon_1} \dots \int_{\varepsilon_n = -\infty}^{v_1 - v_n + \varepsilon_1} f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) d\varepsilon_1 d\varepsilon_2, \dots, d\varepsilon_n$$

• Si les ε_i suivent une loi normale de moyenne nulle et dont on connaît la matrice des variances et covariances, le calcul de l'expression précédente est numériquement possible mais difficile même si les ε_i sont indépendants

• Si ils suivent une loi de Gumbel, Pr (1) a une expression simple à condition que les aléas soient indépendants, et que leurs lois aient les mêmes paramètres $(\mu, 0)$; alors

$$\Pr(1) = \Pr \left(v_1 + \varepsilon_1 \geq \max_{j \in \{2, \dots, n\}} (v_j + \varepsilon_j) \right) = \frac{\exp(\mu v_1)}{\sum_{j=1}^{j=n} \exp(\mu v_j)}$$

En outre :

$$\max(V_1, \dots, V_n)$$

suit une loi de Gumbel dont les paramètres sont :

$$\mu \quad \text{et} \quad \frac{1}{\mu} \log \left(\sum_{j=1}^{j=n} \exp(\mu v_j) \right)$$

lorsque le nombre des variables aléatoires tend vers l'infini). Alors :

$$\begin{aligned} \Pr(d) &= \Pr\left(\max_{m \in Ddm} U_{dm} \geq \max_{m \in Dd'm} U_{d'm}\right) \\ &= \Pr\left[(v_d + \varepsilon_d + \max_{m \in Ddm} (v_m + v_{dm} + \varepsilon_{dm})) \right. \\ &\quad \left. \geq (v_{d'} + \varepsilon_{d'} + \max_{m \in Dd'm} (v_m + v_{d'm} + \varepsilon_{d'm})), \forall d' \in Ddm, d' \neq d\right] \end{aligned}$$

Or :

$$\varepsilon_d = \max_{m \in Ddm} (v_m + v_{dm} + \varepsilon_{dm})$$

suit une loi

$$G\left(\mu_m, \frac{1}{\mu_m} \log\left(\sum_{m \in Ddm} \exp[\mu_m (v_m + v_{dm})]\right)\right)$$

ou

$$G(\mu_m, v_d) \quad \text{avec} \quad v_d = \frac{1}{\mu_m} \log\left(\sum_{m \in Ddm} \exp \mu_m (v_m + v_{dm})\right)$$

Donc

$$\Pr(d) = \Pr\left[v_d + v_d + \varepsilon_d + \varepsilon_d \geq v_{d'} + v_{d'} + \varepsilon_{d'} + \varepsilon_{d'}, \forall d' \neq d\right]$$

En raison des lois suivies par ε_d et $(\varepsilon_d + \varepsilon_d)$ il résulte que :

$$\Pr(d_1) = \frac{\exp(v_{d_1} + v_{d_1}) \mu_{d_1}}{\sum_{d_i} \exp(v_{d_i} + v_{d_i}) \mu_{d_i}}$$

avec :

$$v_{d_i} = \frac{1}{\mu_{d_i}} \log\left(\sum_{m_j} \exp(v_{m_j} + v_{m_j d_i}) \mu_{m_j}\right)$$

et que :

$$\Pr(m_1 \text{ si } d_1) = \frac{\exp(v_{d_1} + v_{m_1 d_1}) \mu_{m_1}}{\sum_j \exp(v_{d_1} + v_{m_j d_1}) \mu_{m_j}} = \frac{\exp(v_{m_1 d_1} \mu_{m_1})}{\sum_j \exp(v_{m_j d_1} \mu_{m_j})}$$

Dans ce modèle, tout se passe comme si, dans un premier temps, l'utilisateur choisissait, pour chaque destination, son mode à travers un modèle logit propre ; puis qu'ensuite il choisissait sa destination à partir d'un autre modèle logit où l'utilité est la somme de l'utilité propre de la destination et de l'utilité qu'il retire du mode choisi par la première étape. D'où l'expression de logit emboîté donnée à ce modèle. L'expression du « *logsum* », déjà vue pour les modèles simples, s'applique aussi aux modèles emboîtés.

1.3. Extension au temps. Le coût généralisé de transport

La deuxième extension de la théorie classique du consommateur consiste à prendre en considération la contrainte de temps qui s'exerce sur l'utilisateur, ce qui conduit à définir la valeur du temps. Plusieurs modèles permettent d'introduire cette notion. On en présentera trois qu'on baptisera des noms de modèle simple, modèle d'activité, modèle d'opportunité (voir Gonzales, 1997 pour une présentation théorique générale).

1.3.1. Modèle simple

Le plus simple consiste à considérer que le consommateur, dans l'optimisation de son comportement, est soumis non seulement à la contrainte de revenu, mais aussi à une contrainte de temps, résultant de ce que pour consommer une unité du bien i , il faut un temps t_i , et que le temps total disponible est contraint. Il doit donc maximiser :

$$U(x_1, \dots, x_p, \dots, x_n)$$

avec non seulement :

$$\sum_i p_i x_i \leq r$$

mais aussi :

$$\sum_i t_i x_i \leq T$$

En appelant λ et μ les variables duales correspondant à ces deux contraintes, on obtient :

$$\frac{\partial U}{\partial x_i} \frac{1}{\lambda p_i + \mu t_i} = \text{Constante}$$

U^* , valeur optimale de l'utilité, dépend des p_i , t_i , de r et de T . De même que λ représente, par le théorème de l'enveloppe, l'utilité marginale du revenu r , de même μ représente l'utilité marginale du temps T , et μ/λ peut être appelé la valeur du temps : augmenter le temps disponible T d'une unité est équivalent à augmenter le revenu de : μ/λ .

1.3.2. Modèle d'activité

On sent bien que ce modèle est naïf. Le temps passé n'est pas proportionnel aux quantités consommées, il n'a pas la même valeur selon les activités auxquelles il est consacré. Une modélisation plus élaborée fait dépendre l'utilité d'un individu de deux catégories d'arguments ; la quantité des biens consommés x_i et le temps passé aux activités offertes t_j , soit :

$$U = U(x_1, \dots, x_n, t_1, \dots, t_m)$$

avec les contraintes :

$$\sum_{i=1}^{i=n} p_i x_i - (w t_w + r) \leq 0$$

$$t_w + \sum_{j=1}^{j=m} t_j \leq T$$

$$\bar{t}_j - t_j \leq 0$$

expression dans laquelle w est le salaire horaire, t_w le temps de travail, \bar{t}_j le temps minimum pour se livrer à l'activité j et T est le temps total disponible (24 h par jour). Avec les variables duales λ, μ, v_j la dérivation de Lagrangien donne en particulier :

$$\frac{\partial U}{\partial t_j} - \mu - v_j = 0$$

$$-\frac{\partial U}{\partial t_w} + \lambda w - \mu = 0$$

$$(t_j - \bar{t}_j) v_j = 0$$

et aboutit à la fonction d'utilité indirecte :

$$V(p_i, \bar{t}_j, w, T, r)$$

La valeur du temps passé dans l'activité j est :

$$h_{\bar{t}_j} = \frac{\partial V}{\partial \bar{t}_j} / \frac{\partial V}{\partial r} = \frac{v_j}{\lambda} = \frac{\frac{\partial U}{\partial t_j} - \mu}{\lambda} \quad (4.2)$$

ou :

$$h_{t_j} = w + \frac{1}{\lambda} \left(\frac{\partial U}{\partial t_j} - \frac{\partial U}{\partial t_w} \right)$$

Deux cas se présentent :

- Si $t_j > \bar{t}_j$, alors

$$v_j = 0 = \frac{\partial U}{\partial t_j} - \mu$$

et la valeur marginale du temps passé à l'activité j est nulle. Il s'agit d'activités de loisir pur.

- Si $t_j = \bar{t}_j$, alors $v_j \neq 0$.

Il s'agit d'activités intermédiaires, auxquelles appartiennent les transports. La formule (4.2) dit alors que la valeur du temps marginal passé dans cette activité est égale au salaire horaire, auquel s'ajoute la différence entre la pénibilité marginale du travail et la pénibilité marginale de l'activité j .

On voit qu'ici la valeur du temps

$$\frac{\partial V}{\partial \bar{t}_j} / \frac{\partial V}{\partial r}$$

est dépendante de l'activité. Elle dépend aussi, bien sûr, des caractéristiques de l'utilisateur. Ainsi beaucoup de modèles mettent en jeu des valeurs du temps qui dépendent du revenu et des motifs de déplacement de l'utilisateur, ainsi que du degré de confort du mode.

1.3.3. Modèle d'opportunité

Ce type de raisonnement ne vaut que pour les activités de consommation, c'est-à-dire en matière de transport, pour les déplacements non liés au travail. Henscher (1997) procède à une analyse du coût d'opportunité des déplacements liés au travail, qui est :

$$h = (1 - r - pq) P_m + \frac{1-r}{1-t} VW + \frac{r}{1-t} VL + \Delta PM$$

formule dans laquelle :

r est la proportion du temps économisé utilisé pour des activités de loisir,

p est la proportion de temps de trajet économisée qui aurait été utilisée à travailler,

q est le rapport entre la productivité quand on travaille durant le transport et la productivité normale,

PM est la productivité marginale du travail,

VW est la différence entre la pénibilité du transport et la pénibilité du travail pour le travailleur,

VL est la pénibilité du transport pour le travailleur,

ΔPM représente le produit supplémentaire permis par la diminution de la fatigue, t est le taux de l'impôt sur le revenu.

1.3.4. Le coût généralisé de transport

À partir de là on peut introduire le concept de coût généralisé de transport. Ainsi, dans le cas du modèle simple va en 1.3.1., si x_i représente la quantité de transport utilisé, une variation d'une unité de cette quantité conduit à une variation d'utilité donnée par la formule :

$$\frac{\partial U}{\partial x_i} = \frac{\partial V}{\partial r} p_i + \frac{\partial V}{\partial T} t_i$$

et

$$\left(p_i + \frac{\partial V}{\partial T} / \frac{\partial V}{\partial r} t_i \right)$$

apparaît jouer le même rôle que le prix dans la formulation classique,

$$\left(\frac{\partial V}{\partial T} / \frac{\partial V}{\partial r} \right)$$

étant la valeur du temps.

De même, dans le modèle d'activité, si le transport représente l'activité j dont le coût monétaire est p_j on voit que :

$$\frac{\partial u}{\partial x_j} = \frac{\partial V}{\partial r} p_j + \frac{\partial V}{\partial t_j} t_j = \frac{\partial V}{\partial r} \left(p_j + \frac{\partial V}{\partial t_j} \frac{t_j}{r} \right)$$

Là aussi, on peut définir un coût généralisé ; ce coût généralisé a pour expression :

$$p_i + \frac{\partial V}{\partial t_j} \frac{t_j}{r}$$

On peut aussi introduire d'autres facteurs qui influencent l'utilité du transport pour le consommateur : sécurité, fiabilité, confort...

2. L'application aux transports : les modèles de prévision de trafic

Les premiers modèles de trafic ont été élaborés par des géographes et par des ingénieurs. On retrouve la trace du passage de ces derniers à travers les noms dont été baptisés ces modèles et les formules qu'ils utilisent : « impédance », « formule gravitaire », « entropie ». Ce sont ainsi des économistes qui sont à l'origine du modèle en 4 étapes universellement utilisé :

- génération : on y modélise le nombre de déplacements effectué à partir d'un centre d'émission,
- distribution : on y détermine comment ces déplacements se répartissent entre les destinations possibles,
- choix modal : pour chaque ensemble de trajets entre une origine et une destination données, répartition de ces trajets entre les modes de transports possibles,
- choix d'itinéraires. Cette étape concerne essentiellement le mode routier et vise à répartir les usagers de ce mode entre les différents itinéraires joignant l'origine à la destination en cause.

À ces quatre étapes, traditionnelles depuis plusieurs décennies, il devient de plus en plus nécessaire d'ajouter une étape de choix de l'heure de départ et, en zone urbaine de recherche de stationnement.

Avant d'examiner successivement chacune de ces étapes, notons que, du point de vue de l'économiste, leur indépendance est une hypothèse restrictive non-naturelle. Si par exemple les destinations sont plus nombreuses, plus variées, ou moins coûteuses d'accès, il est *a priori* normal que le nombre de déplacements augmentent, c'est-à-dire que la génération en soit affectée. De même, lorsqu'un mode améliore ses performances en temps et en coût entre deux villes, la répartition des destinations, c'est-à-dire la distribution, change au profit des deux villes en cause. On verra comment il est possible d'introduire les liens entre les 4 étapes précédées.

2.1. Les modèles de génération

Il en existe deux catégories. Dans la première, qu'on pourrait appeler agrégée, on détermine le nombre moyen de déplacements d'une population prise dans son ensemble. En général, on reliera cette variable à plusieurs autres concernant le niveau d'activité ou de développement économique et les autres caractéristiques de la population ; par exemple :

$$T_i = k r_i^a p_i^b$$

T_i est le nombre moyen de déplacements par unité de temps de la population de la ville ou du quartier i ,

r_i est le revenu moyen par habitant de cette population,

p_i est le taux de possession d'une automobile,

k, a et b sont des paramètres.

Dans la seconde, qu'on peut qualifier de désagrégée, on subdivise la population et ses déplacements en catégories homogènes, à l'intérieur desquelles une certaine stabilité de la génération de déplacements peut être envisagée. Par exemple on subdivisera les déplacements selon leurs motifs :

- domicile-travail,
- déplacements professionnels,
- déplacements privés pour affaires,
- déplacements de loisirs de week-end,
- déplacements de loisirs de grandes vacances.

Les voyageurs seront classés suivant plusieurs critères :

- revenu,
- structure et taille du ménage,
- âge,
- possession ou non d'une voiture,
- catégorie socioprofessionnelle,
- caractéristiques de localisation.

On peut penser qu'à l'intérieur de chacune de ces catégories les comportements de mobilité présentent une certaine constance, ce qui facilite la transposition du modèle dans le temps et dans l'espace : soient Π_{ijk}^t les proportions d'utilisateurs présentant à l'instant t (ou dans la ville i) les caractéristiques :

- i de revenu,
- j de structure et taille du ménage,
- k d'âge,

et soit g_{kij} le taux d'engendrement des déplacements de cette catégorie d'usagers. Alors le taux moyen de déplacement de la population t est :

$$G^t = \sum_{ijk} \Pi_{ijk}^t g_{kij}$$

et dans une autre ville t' (ou à un autre instant t') il sera :

$$G^{t'} = \sum_{ijk} \Pi_{ijk}^{t'} g_{kij}$$

On constate des changements dans les destinations, donc dans la distribution des trafics, marqués par l'allongement des parcours moyen des déplacements : 0,85 % par an en région Ile-de-France, près de 3 % par an pour les agglomérations de province. Ce dernier phénomène est lié à l'étalement des villes ; mais, comme les vitesses de déplacement augmentent, le temps total moyen passé en transports reste à peu près constant ; ce résultat statistique approché, fruit de l'approximative compensation de phénomènes contraires, est parfois dénommé « loi de Zahavi » du nom du chercheur qui l'a mis en évidence il y a quelques décennies : autour de 29 minutes par jour en région Ile-de-France, de 17 ailleurs.

Les modèles de génération et les résultats d'enquêtes analysés ne font pas dépendre la mobilité des conditions de l'offre. En bonne logique, une réduction des coûts de transport ou de la qualité de service qu'ils offrent devrait se traduire par une augmentation des déplacements. C'est bien ce que traduisent les courbes de la figure 4-1 précédente. Mais on n'a jamais tenté de traduire cet effet en faisant dépendre la génération issue d'un centre de l'accessibilité de ce centre⁴. Il est en effet compréhensible que cet effet soit faible au regard de facteurs tels que le revenu ou la possession d'une automobile. En outre les évolutions de l'accessibilité sont lentes et corrélées avec d'autres facteurs, donc difficiles à déceler statistiquement.

2.2. Les modèles de distribution

La distribution, phase qui répartit le trafic issu d'un centroïde entre les différentes destinations possibles, est presque totalement fondée sur le modèle gravitaire. Celui-ci connaît de nombreuses justifications et revêt de multiples formes. La plus ancienne, d'où le modèle tire son nom, s'écrit :

$$T_{ij} = k \frac{(P_i P_j)^a}{d_{ij}^b}$$

dans cette formule :

- T_{ij} est le trafic de i à j ,
- P_i et P_j sont les populations de i et j ,
- d_{ij} est la distance entre i et j ,
- a et b sont les coefficients.

Comme les estimations de a tournaient autour de 1 et de b autour de 2, l'analogie avec la formule de la gravitation universelle était patente. Une formulation plus générale du modèle est :

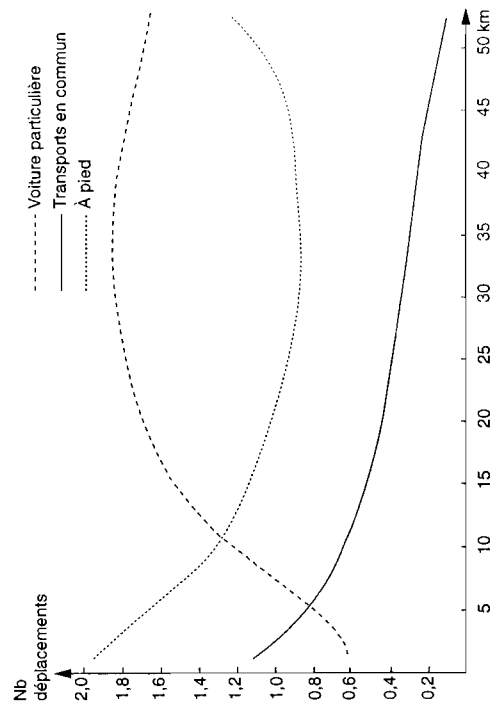
$$T_{ij} = k O_i P_j f(c_{ij})$$

4. Définie par exemple comme une moyenne pondérée des coûts généralisés de transport entre ce centre et les autres centres de la zone étudiée.

Que fournit l'observation quant à ces lois de génération ? De nombreuses enquêtes statistiques sont faites à ce sujet. Les résultats les plus intéressants sont relatifs aux déplacements quotidiens urbains et suburbains, qui présentent une grande répétitivité et pour lesquels donc la mise en évidence de régularités est plus aisée. Ils concernent la mobilité, c'est-à-dire le nombre de déplacements par jour, en la séparant selon ses motifs.

Les analyses font apparaître une grande constance du total des déplacements (Coindet, 1996 et Guider, 1996). Ceux-ci s'élèvent en moyenne à environ 3,5 déplacements par personne et par jour (y compris marche à pied) en région Ile-de-France ; ce chiffre a à peine évolué de quelques pour cent au cours des 20 dernières années. En province, le chiffre est un peu plus faible, mais ne varie guère ni dans le temps depuis une quinzaine d'années ni d'une agglomération à l'autre, et se situe entre 3,2 et 3,4 déplacements par jour. La génération des déplacements quotidiens est donc assez stable. Mais cette constance temporelle au niveau d'une agglomération, et spatiale entre agglomérations, recouvre en fait de grandes différences au niveau individuel. Ainsi, Lefèvre et Offner (1990), commentant une enquête sur les déplacements des franciliens, dont certains résultats sont présentés dans la figure 4-1, mettent en évidence les modifications du partage modal des déplacements selon la distance au centre de la ville : la marche à pied et les transports collectifs, prépondérants au centre, cèdent la place à l'automobile en périphérie ; le nombre total des déplacements par personnes, et notamment des déplacements motorisés, est plus stable, mais varie quand même entre 1,7 et 2,1 déplacements par jour, en fonction de l'éloignement au centre.

FIGURE 4-1
Mobilité des franciliens selon la distance au centre



Source : Lefèvre et Offner.

formule dans laquelle :

- O_i est un facteur d'émission de la zone origine i ,
- D_j est un facteur d'attraction de la zone destination j ,
- f est une fonction décroissante parfois appelée fonction d'impédance,
- C_{ij} est le coût généralisé de circulation entre i et j .

Selon les contraintes qu'on ajoute à la formule précédente, on obtient le modèle non contraint (absence de contrainte supplémentaire sur les T_{ij}), le modèle simplement contraint (la somme des trafics sortant de chaque zone est fixée de façon exogène), ou le modèle doublement contraint (ce sont à la fois les trafics qui sortent de chaque zone et ceux qui entrent dans chaque zone qui sont fixés).

Cette formule peut se déduire de plusieurs principes. Le plus ancien constitue une analogie avec la notion d'entropie. On y considère les flux issus de chaque origine comme fixés, ainsi que les flux arrivant dans chaque destination. Alors le nombre d'arrangements permettant d'aboutir à une matrice donnée est (Williamson et Ortuzar, 1990) :

$$W(\|T_{ij}\|) = \frac{T!}{\prod_{ij} (T_{ij}!)}$$

Par la formule de Stirling on peut écrire sous forme approchée :

$$\log(W) = \log T! - \sum_{ij} (T_{ij} \log T_{ij} - T_{ij})$$

On cherche à maximiser l'entropie avec les contraintes :

$$\sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = O_i$$

et une contrainte additionnelle qui fixe à un niveau donné le coût total de transport :

$$\sum_{ij} T_{ij} C_{ij} = C$$

La maximisation du Lagrangien, avec les variables duales λ_i, μ_j, v , aboutit à

$$T_{ij} = \exp(-\lambda_i) \exp(-\mu_j) \exp(-\beta C_{ij})$$

ce que l'on peut écrire :

$$T_{ij} = O_i a_i D_j b_j \exp(-\beta C_{ij})$$

Si l'on n'avait pas introduit la contrainte de coût total, on aurait obtenu :

$$T_{ij} = O_i a_i D_j b_j$$

Le modèle simplement contraint (seules subsistent la 3^e et, par exemple, la 2^e contraintes) aboutit à :

$$T_{ij} = O_i \frac{D_j \exp(-\beta C_{ij})}{\sum_k D_k \exp(-\beta C_{ik})}$$

On peut trouver une utilisation du modèle sans contrainte de coût total lorsque l'on veut mettre à jour une matrice $O \cdot D$ qu'on connaissait à une date antérieure, où elle était $\|t_{ij}\|$ et où les données disponibles actuellement concernent les nouvelles marges O_i et D_j . Alors il peut paraître normal de chercher à maximiser (Williamson et Ortuzar, 1990)

$$W = - \sum_{ij} \left(T_{ij} \log \frac{T_{ij} - T_{ij} + t_{ij}}{t_{ij}} \right)$$

expression dont on démontre qu'elle vaut à peu près :

$$W' = -0,5 \sum_{ij} \frac{(T_{ij} - t_{ij})^2}{t_{ij}}$$

le résultat est de la forme

$$T_{ij} = a_i O_i b_j D_j t_{ij}$$

On calcule les T_{ij} par itérations sur la matrice t_{ij} dont on multiplie alternativement les lignes et les colonnes par les coefficients permettant d'ajuster les marges et cela successivement jusqu'à convergence de la procédure.

Ces présentations du modèle gravitaire sont peu convaincantes, car elles procèdent par analogie et similitude et ne résultent pas d'un raisonnement issu d'hypothèses de comportement. Des fondements économiques peuvent lui être trouvés à partir de la théorie des choix discrets présentée en section 1.2. Présentons le raisonnement dans le cas des déplacements domicile-travail, en supposant qu'un individu issu de la zone i cherche à maximiser l'utilité qu'il retire de son travail, égale à $s_k + c_{ij}$.

Dans cette expression s_k est le salaire obtenu par l'emploi k , j est la localisation de l'emploi k . Faisons l'hypothèse que les salaires sont distribués selon une loi de répartition indépendante de la localisation, de la forme :

$$s_k = \bar{s} + \varepsilon_k$$

les ε suivant une loi de Gumbel de moyenne nulle et de paramètre μ .

Supposons aussi qu'il y ait D_j emplois dans la zone j . Alors l'utilité maximale nette des coûts de transport susceptible d'être obtenue dans la localisation j sera une variable aléatoire u_j égale au maximum de $s_k - c_{ij}$ qui suivra une loi de Gumbel de paramètres μ et :

$$\frac{1}{\mu} \log D_j \exp(\bar{s} - c_{ij}) \mu$$

Sa moyenne sera :

$$m_j = \frac{1}{\mu} (\mu(\bar{s} - c_{ij}) + \log D_j)$$

la probabilité que la destination j soit choisie sera :

$$\Pr(j) = \frac{\exp(\mu m_j)}{\sum_k \exp(\mu m_k)} = \frac{D_j \exp \mu(\bar{s} - c_{ij})}{\sum_k D_k \exp \mu(\bar{s} - c_{ik})} = \frac{D_j \exp(-\mu c_{ij})}{\sum_k D_k \exp(-\mu c_{ik})}$$

et pour le trafic total

$$T_{ij} = O_i \frac{D_j \exp(-\mu c_{ij})}{\sum_k D_k \exp(-\mu c_{ik})}$$

C'est le modèle gravitaire simplement contraint où les destinations ne sont pas mutuellement exclusives. On peut de ce modèle tirer une formule pour le modèle doublement contraint (Cochrane, 1975 ; Koenig, 1974), avec l'hypothèse que les individus retenus dans la localisation j sont ceux qui en retirent la plus forte utilité. On trouve alors :

$$\Pr(j) = O_i \frac{D_j \exp(-\mu c_{ij}) \exp(-\beta_j)}{\sum_k D_k \exp(-\mu c_{ik}) \exp(-\beta_k)}$$

Koenig (1974) montre que cette formule reste valable au premier ordre si les aléas suivent une loi quelconque, pas forcément doublement exponentielle, pourvu que les nombres d'emploi à chaque destination D_j soient grands.

Enfin, le modèle peut être étendu au cas non contraint : l'individu se déplace dès que l'utilité qu'il retire du déplacement est positive. Alors, on démontre (Cochrane, 1975) que :

$$T_{ij} = k O_i D_j \exp(-\mu c_{ij})$$

On retrouve alors le modèle gravitaire sous sa forme initiale.

L'utilisation de chacune de ces catégories de modèles, non contraint, simplement ou doublement contraint, est affaire de situation. Le modèle non contraint admet une élasticité au coût de transport, et en fait, il traite à la fois la génération et la distribution. C'est lui qui est en général utilisé dans les études de trafic interurbain. Le modèle doublement contraint est bien adapté aux déplacements urbains et suburbains domicile-travail. Quant au modèle simplement contraint, il cadre mieux avec les autres motifs de déplacements urbains.

2.3. Le partage modal

La modélisation du partage modal est fondée sur l'idée que ce partage dépend de la différence, ou parfois du rapport, entre les coûts de transport. La forme générale, dans le cas du choix entre deux modes, est celle d'une courbe logistique, de la forme :

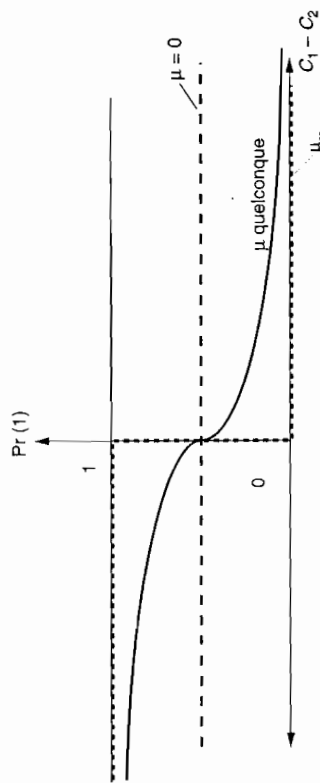
$$\Pr(1) = \frac{1}{1 + \exp \mu(C_1 - C_2)}$$

expression dans laquelle :

- $\Pr(1)$ est la proportion d'utilisation du mode 1,
- C_1 et C_2 sont les coûts généralisés de chacun des deux modes,
- μ est un paramètre positif.

Selon la valeur de μ , le choix est plus ou moins sensible à la différence des coûts (figure 4-2).

FIGURE 4-2
Courbe logistique



Le choix par tout ou rien correspond à la valeur infinie de μ . Une situation dans laquelle le partage serait insensible aux coûts correspond à la valeur 0. Bien sûr d'autres formules aboutissent à la même forme de courbe. Toutefois celle-ci présente la propriété de pouvoir être déduite d'un modèle logit de choix discret.

La mise en œuvre de ce principe général peut être effectuée soit au niveau agrégé, soit au niveau désagrégé.

2.3.1. Modèles agrégés

Au niveau agrégé, l'unité d'observation est constituée par un couple origine-destination et $\Pr(1)$ représente la proportion d'utilisateurs utilisant le mode 1 dans la population totale : par exemple la proportion d'utilisateurs utilisant le chemin de fer pour aller de Paris à Lyon, ou encore la proportion d'utilisateurs du chemin de fer utilisant le train de nuit pour aller de Paris à Nice. L'utilisation d'une formule du type logistique peut dans ce cas être justifiée par le raisonnement suivant :

Chaque usager i a un coût généralisé de transport pour chaque mode 1 et 2 qui est de la forme :

$$C_1^i = C_1 + \varepsilon_1^i$$

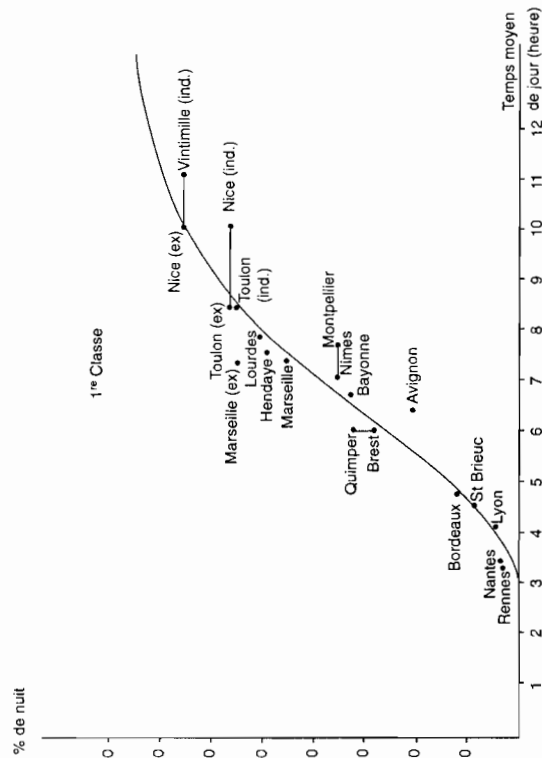
$$C_2^i = C_2 + \varepsilon_2^i$$

Les ε sont des variables aléatoires centrées et indépendantes ; elles représentent les différences entre les modes. Elles peuvent concerner soit des paramètres non pris en compte dans les coûts généralisés C_1 et C_2 (dans ce cas, elles ne seront plus forcément de moyenne nulle : si par exemple un mode présente un avantage de commodité qui n'est pas pris

en compte dans les coûts moyen C_1 et C_2 les aléas ne sont plus de moyenne nulle, ce qui se traduit par le fait que l'égalité des coûts C_1 et C_2 n'entraîne plus l'égalité des parts modales), soit des caractéristiques particulières à chaque mode de transport en cause introduisant des différences sur les coûts C_1 et C_2 (par exemple différences dans l'appréciation de la durée du trajet). On est alors en présence d'un modèle de choix discret aléatoire, et on retrouve les formulations logit ou probit développées dans la section 1.

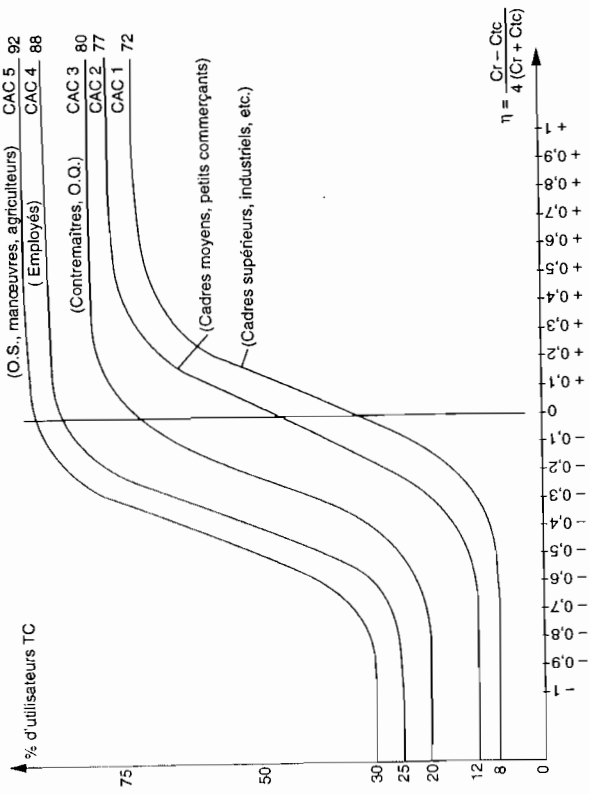
Ce type de modèle est utilisé pour représenter le partage entre trafic de jour et trafic de nuit à la SNCF, ou encore pour modéliser le partage de trafic entre transports publics et individuels à la RATP (figures 4-3 et 4-4).

FIGURE 4-3
Pourcentage de trafic de nuit en fonction du temps moyen « jour »



Source : SNCF, 1993.

FIGURE 4-4
Affectation VP/TC par catégories sociales



Source : Hainault et Karam, 1998.

2.3.2. Modèles désagrégés

Au niveau désagrégé, l'unité d'observation est constituée par le choix d'un individu, et les utilités de chaque mode figurant en argument de la fonction logit, qui sont des utilités indirectes au sens de la théorie du consommateur, dépendent du prix du mode, du temps passé, et éventuellement d'autres caractéristiques de l'individu et du mode :

$$U_{ij} = \sum_k \beta_{ij}^k X_{ij}^k + \varepsilon_{ij}$$

formule dans laquelle :

- i : représente l'individu,
- j : représente le mode,
- k : représente chacun des k attributs qui caractérisent le couple (i, j) ,
- β_{ij}^k : sont des paramètres,
- ε_{ij} : représentent les aléas affectant l'utilité.

Selon les hypothèses sur la forme de la loi suivie par les aléas, on obtient différents modèles. Les plus utilisés sont les modèles logit, en raison des commodités de calcul qu'ils offrent. Mais certaines de leurs conséquences sont paradoxales et peu vraisemblables. On a déjà vu la propriété

« IIA », et le moyen d'y remédier grâce aux modèles emboîtés. Mais d'autres anomalies se présentent. Ainsi, de la formule :

$$\frac{P_i(j)}{P_i(l)} = \frac{\exp(\mu V_{ij})}{\exp(\mu V_{il})}$$

on tire :

$$\Delta \frac{P_i(j)}{P_i(l)} = \frac{P_i(j)}{P_i(l)} \exp(\mu \Delta V_{ij})$$

On en déduit que par exemple une variation de un quart d'heure d'un des modes a le même effet sur le partage entre deux modes que le trajet dure une 1/2 heure ou 10 heures. De la même manière, on verrait facilement que les élasticités directes et croisées sont :

$$E\left(\frac{P_i(j)}{P_i(l)} / X_{ij}^k\right) = (1 - P_i(j)) \beta_{ij}^k X_{ij}^k$$

$$E\left(\frac{P_i(j)}{P_i(l)} / X_{il}^k\right) = -P_i(l) \beta_{il}^k X_{il}^k$$

D'après ces formules, les modifications sur un mode affectent de façon similaire la part des autres modes, un résultat très restrictif. Enfin la spécification linéaire ne permet pas de rendre compte d'effet de seuil sur les variables X_{ij} . Ces inconvénients peuvent être évités par la transformation Box-Cox, qui fait intervenir les variables transformées :

$$Y_{ij}^k = \begin{cases} \frac{(X_{ij}^k)^{\lambda_k} - 1}{\lambda_k} & \text{si } \lambda_k \neq 0 \\ \log X_{ij}^k & \text{si } \lambda_k = 0 \end{cases}$$

Gaudry (notamment Gaudry et Wills, 1992) a spécialement étudié les propriétés de cette transformation et on trouvera dans Mandel, Gaudry et Rothengather (1996) un exemple d'application de cette méthode à la prévision des trafics en Allemagne, et la mise en lumière des progrès qu'elle permet.

La caractéristique essentielle des modèles désagrégés réside dans l'unité d'observation qu'ils utilisent, qui correspond au choix d'un usager, alors que l'unité d'observation des modèles agrégés correspond au choix moyen de l'ensemble des usagers d'une relation origine-destination. Cela crée plusieurs différences à la fois dans l'élaboration et dans l'utilisation des modèles. En effet, même si les variables explicatives sont les mêmes, il suffira dans un modèle agrégé de connaître leurs valeurs moyennes, qu'on peut souvent observer sans enquête nouvelle ; dans un modèle désagrégé il faudra connaître les valeurs de ces variables pour chaque usager et le choix de l'usager. Cela implique donc des enquêtes auprès des usagers, administrées par questionnaires, donc plus coûteuses.

Le calibrage des modèles passe dans les deux cas par des méthodes du maximum de vraisemblance. On dispose en général d'un plus grand nombre d'observations par les enquêtes auprès des usagers nécessaires

pour la mise en œuvre des modèles désagrégés comparées à celles qui résultent de l'observation des choix moyens sur des origines-destinations données ; ainsi les graphiques précédents, correspondant à des modèles agrégés, avaient été établis à partir d'une trentaine d'observations, alors que les points d'observations pour les modèles désagrégés dépassent couramment la centaine. Ceci ouvre la possibilité de faire intervenir dans ce dernier cas un plus grand nombre de variables explicatives de l'utilité du mode, alors que dans les modèles agrégés on n'utilise guère que le coût monétaire et le temps passé, arguments de l'utilité négative que constitue le coût généralisé de transport.

Le modèle désagrégé, une fois calibré, permet de déterminer les probabilités de choix de chaque individu de caractéristiques données ; il ne fournit pas directement la fréquentation totale de chaque mode, qui constitue en général le résultat que l'on vise. Il faut pour cela procéder à l'agrégation et différentes méthodes le permettent (Ben Akiva et Lerman, 1985) :

- la classification des individus en catégories définies par des valeurs particulières des paramètres explicatifs. À l'intérieur de chaque classe on raisonne sur un représentant moyen auquel on affecte la proportion de la classe dans l'effectif total ;

- l'intégration, limite de la classification lorsque le nombre des catégories augmente, la taille de chacune tendant vers zéro ;
- l'énumération d'échantillon, consistant à tirer au hasard un certain nombre d'individus dans la population visée, puis à déterminer les choix de chacun des individus de l'échantillon, et à les agréger ensuite.

Notons que si, dans la méthode de classification, on ne prend qu'une seule catégorie, on applique en fait le modèle désagrégé à l'individu moyen de la population, ce qui revient à utiliser un modèle agrégé. L'erreur qui en résulte peut être importante comme le montre l'exemple figurant dans l'encadré 4-1.

On pourrait penser que les modèles désagrégés, allant plus avant que les modèles agrégés dans l'explication du comportement, seraient plus valablement transférables d'une situation à l'autre, qu'on pourrait par exemple utiliser pour une ville B les modèles calibrés sur une ville A. Ces espoirs ont été en partie déçus, et la transférabilité reste un idéal non atteint.

Les modèles désagrégés ont été relativement peu utilisés en France. Citons l'usage qui en a été fait par l'INRETS pour la prévision du trafic d'Orly-Val (Hivert, Orfeuil, Troulay, 1988) ou par la RATP pour la modélisation du trafic en Région Ile-de-France (Rousseau et Saut, 1991 et 1997). Les chiffres du tableau 4-1 sont tirés de cette dernière étude : ils représentent la liste des coefficients intervenant dans la formule donnant les utilités de chaque mode.

Coefficients du partage modal en Ile-de-France (domicile-travail)

Mode	Paramètre
Constante MAP	2,67
Constante 2R	- 0,45
Constante VP destination Paris	- 2,00
Constante VP destination Banlieue	- 1,03
Constante Bus pur	- 0,07
Constante Mixte	- 0,32
Sexe masculin	1,27
Chef de ménage	1,19
Nombre de VP/Nombre d'actifs	1,64
Temps de trajet MAP	- 0,16
Temps de trajet 2R	- 0,095
Temps de trajet VP	- 0,067
Temps de parking CSP 5	- 0,34
Temps de parking CSP 6-7-8	- 0,16
Temps de roulement bus Paris	- 0,075
Temps de roulement TC sauf bus Paris	- 0,032
Temps de rabattement < 10 minutes	- 0,071
Temps de rabattement > 10 minutes	- 0,161
Temps d'attente -- rupture de charge	- 0,115
Coût 2R	- 2,19
Coût VP/Revenu	- 9,05
Coût Parking/Revenu CSP 1-3	- 26,8
Coût Parking/Revenu CSP 2-4	- 13,8
Coût TC/(Revenu x distance)	- 18,4

Source : Rousseau et Saut, 1991.

2.3.3. Le modèle prix-temps

La faible utilisation des modèles désagrégés en France résulte, pour ce qui concerne les études de trafic interurbains, des bonnes performances du modèle prix-temps. Développé par Abraham et Blanchet (1973) il est utilisé conjointement par la SNCF et par l'aviation civile pour modéliser la concurrence air-fer. Ce modèle, analogue aux modèles de discrimination des produits utilisés en économie industrielle, est intermédiaire entre les modèles agrégés et les modèles désagrégés. Il est fondé sur les principes suivants, exposés dans le cas de la concurrence fer-avion :

On suppose que chaque usager i prend ses décisions en fonctions des coûts généralisés de chaque mode :

$$C_f^i = P_f + h_i t_f$$

$$C_a^i = P_a + h_i t_a$$

ENCADRÉ 4-1

Le biais d'agrégation (tiré de Small, 1992)

Considérons le choix entre auto et bus, modélisé par un modèle logit de la forme :

$$P_b^i = \frac{1}{1 + \exp(\beta_1 (c_a^i - c_b^i) + \beta_0)}$$

expression dans laquelle :

P_b^i est la probabilité du choix du bus,

c_a^i, c_b^i sont les coûts monétaires de l'automobile et du bus,

i étant l'individu,

les β sont des constantes.

Les valeurs des paramètres et des variables sont telles que la moitié des usagers a une probabilité de 0,9 de prendre le bus et l'autre moitié une probabilité de 0,1. Le partage moyen est donc moitié-moitié entre bus et auto.

Si le coût par autobus augmente de dc , la probabilité de prendre le bus change, selon la formule :

$$\frac{dP_b^i}{dc} = P_b^i (1 - P_b^i) \beta_1$$

Le modèle agrégé donnerait un taux de changement moyen :

$$0,25\beta_1$$

L'application du modèle désagrégé à chacune des deux catégories de population donne un taux de changement moyen de :

$$0,9 \times 0,1\beta_1 = 0,09\beta_1$$

On trouve un résultat qui est confirmé par d'autres modèles, à savoir que la diversité des situations adoucit la sensibilité à l'égard des variables.

expressions dans lesquels :

- C_f^i et C_a^i sont les coûts généralisés du fer et de l'avion.
- P_a et P_f sont les prix des 2 modes,
- t_a et t_f sont les temps de trajet,
- h_i est la valeur du temps de l'individu i .

Alors l'individu i choisit le fer si :

$$C_f^i \leq C_a^i$$

soit :

$$h_i \leq \frac{P_a - P_f}{t_f - t_a}$$

On peut déterminer par les conditions de l'offre la quantité :

$$h_b = \frac{P_a - P_f}{t_f - t_a}$$

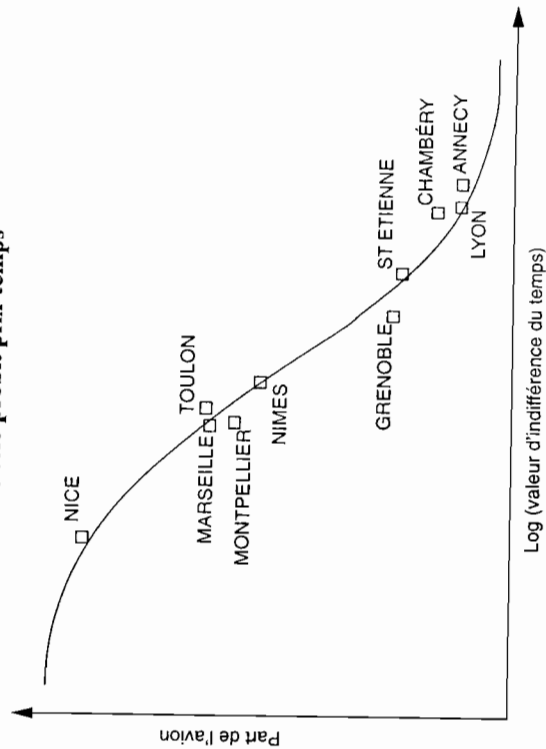
qu'on appelle « valeur de basculement ». On fait en général l'hypothèse que les valeurs du temps sont réparties comme les revenus. Ceux-ci suivent une loi log-normale (Arduin, Ni, Picq, 1994) dont la connaissance permet de déterminer la proportion d'usagers prenant le fer :

$$\Pr(f) = \Pr(h \leq h_b)$$

Ce modèle est en un sens désagrégé, puisqu'il analyse les comportements individuels et non un comportement moyen ; mais, en ce qui

FIGURE 4-5

Modèle probit prix-temps



Source : SNCF, 1993.

concerne les données nécessaires à sa calibration et à son utilisation, il présente les mêmes caractéristiques qu'un modèle agrégé. Il a été étendu dans de nombreuses directions.

Ainsi, la SNCF a construit à partir de lui un modèle probit prix-temps, en ajoutant aux coûts généralisés précédents un aléas normal centré. La figure ci-dessus, tirée du document SNCF 1993, montre la qualité de l'ajustement.

Ce modèle est également utilisé pour la modélisation des choix d'itinéraires, comme on le verra à la section suivante.

Avant d'aborder cette phase de la modélisation des trafics, il est intéressant de procéder à une revue des valeurs numériques trouvées pour les valeurs du temps, car c'est en général à travers les modèles de partage modal que ces valeurs sont déterminées.

2.3.4. Les valeurs du temps

Les valeurs du temps se déduisent des modèles de trafic, comme le taux marginal de substitution entre temps et prix, soit :

$$-\frac{\partial V}{\partial t} / \frac{\partial V}{\partial c}$$

TABLEAU 4-2

Valeurs horaires du temps en France

Auteur	Zone géographique	Année	Valeur en F 1990	Remarques
INRETS	Région parisienne (LASER)	1988	60	valeur par véhicule
			70	motif personnel
			185	motif domicile-travail
CETUR	Marseille et Grenoble	1993	de 30 F à 60 F	motif professionnel
			47	variable selon le motif
Direction des Routes	France entière	1986	31	valeur tutélaire
			49	valeurs dites révélées :
			51	valeur sans bonus
SNCF	France entière	1989	130	en incorporant le bonus
			71	utilisateurs de 2 ^e classe
			63	utilisateurs de 1 ^{re} classe
MATISSE	France entière	1988	188	voiture particulière
			310	train 2 ^e classe
				train 1 ^{re} classe
				avion
SNCF-Air Inter	France entière	1988	230	Nota : il s'agit de valeurs moyennes. La vraie valeur dépendant du revenu, du motif, de la distance... valeur très variable d'une liaison à l'autre

Source : Boiteux, 1994.

Segonne a enfin comparé les résultats d'une enquête de préférences déclarées à ceux obtenus à partir des comportements révélés cités plus haut. La différence est significative ; ainsi, avec un modèle logit à valeurs du temps distribuées selon une loi log-normale, la médiane pour les valeurs du temps aux heures moyennes est de 35 F/h, contre 47,1 F selon les comportements réels.

Henscher (1995) cite pour l'Australie les valeurs du temps suivantes :

	Valeur du temps en US dollar par heure	
	tirée d'un modèle comportement	coût d'opportunité
Déplacement domicile-travail en automobile privée	4,65	4,08
en automobile de société par train	7,87	6,90
par bus	2,43	2,13
	4,00	3,51
Déplacement de longue distance pour motif personnel	6,52	5,72
pour motif professionnel	12,50-25,00	21-25
Déplacement urbain pour motif professionnel	10,10	20,20

Small (1992) conclut de la recension à laquelle il procède que la valeur du temps domicile-travail est d'environ la moitié du salaire brut, que la valeur du temps en trajet professionnel est nettement plus élevée, de l'ordre du salaire brut, alors que la valeur du temps de déplacement personnel se situe aux alentours de celle de déplacement domicile-travail. Dans les modèles prix-temps la valeur du temps est en général voisine du revenu horaire. Merlin (1991) et Small (1992) font apparaître que les valeurs du temps diffèrent selon que le temps est passé en déplacement ou en correspondance, et dans ce dernier cas selon que la correspondance est à l'air libre ou non. Les coefficients pour passer de la valeur du temps en mouvement à la valeur du temps à l'arrêt sont de l'ordre de 2 à 3. Enfin il semble y avoir quelques raisons de penser que la valeur du temps varie avec le temps de trajet (Massiani, 1997).

À côté du temps de transport, on considère aussi, surtout dans l'analyse des situations de congestion, le coût du temps d'arrivée en avance et le coût du temps d'arrivée en retard ; quelques études expérimentales ont été faites sur le sujet ; les résultats, cités dans de Palma et Rochat (1996), en sont rassemblés dans le tableau 4-3.

L'étude la plus complète sur la valeur du temps est celle de MVA (1987), et les travaux qui suivirent ont eu surtout pour fonction d'en préciser les conclusions.

Comme on l'a vu, les modèles, même désagrégés, ne sont pas parfaitement transférables, il faut donc s'attendre à ce que les valeurs du temps soient hétérogènes. Ainsi, pour la France, le rapport Boiteux (1994) présente quelques valeurs du temps issues des modèles les plus fréquemment utilisés.

La plupart des études urbaines distinguent selon les motifs de déplacements. Ainsi, le projet MUSE utilisait en 1990 les valeurs du temps suivantes, valables pour un usager routier en Ile-de-France :

Domicile-travail	60 F
Déplacement personnel	70 F
Déplacement professionnel	267 F

La mise en service du tunnel à péage Prado-Carénage à Marseille a fourni une mine d'informations concernant les arbitrages prix-temps en zone urbaine ; et cette mine a été exploitée par de nombreux chercheurs. On reproduira ici les caractéristiques des distributions de valeurs du temps trouvées :

Auteurs	Modèle	Valeur du temps en F/km			Observations
		Moyenne	Médiane		
Morellet (1997)	Matisse	90	47		moyennes pour un jour
		101	55		moyen annuel trajets de 3 à 10 km
Leurent (1997)	Logit à valeur du temps fixe	63,0	63,0		
	Probit à valeur du temps fixe	60,2	60,2		
	Logit à valeur du temps distribuée selon une loi log-normale	70,3	56,5		
Segonne (1998)	prix-temps	144,5	78,1		tous motifs
		113,1	58,0		· heure de pointe du matin
		127,5	62,5		· heure moyenne
	Logit à valeur du temps distribuée selon une loi log-normale	64,6	57,4		· heure de pointe du soir
		64,2	47,1		· heure de pointe du matin
		69,2	50,0		· heure moyenne

Segonne (1997) a également exploité les résultats d'enquêtes auprès des automobilistes fournissant de précieuses indications sur les temps de parcours estimés et leur comparaison avec les temps de parcours réels. Il en ressort que l'ensemble des usagers surestime le gain de temps de la nouvelle infrastructure : 16,5 minutes estimé contre zéro en réalité ; pour ceux des usagers qui ont effectivement pris la nouvelle infrastructure, le gain de temps estimé est de 23 minutes, alors que le modèle donne un gain de temps de 8 minutes.

TABLEAU 4-3

Auteur	Pays ou ville	Rapport au coût du temps de transport	
		du coût d'arrivée en avance	du coût d'arrivée en retard
Small (1992)	USA	0,64	2,39
Khattak <i>et al.</i> (1995) de Palma et Rochat (1996)	Bruxelles Genève	0,38 0,327	1,03 2,69

2.4. Le choix d'itinéraires

Cette phase de la modélisation est surtout utilisée pour le mode routier. Le choix d'itinéraires présente de nombreux points communs avec le choix du mode ; il s'agit dans les deux cas d'un choix entre un petit nombre d'alternatives. La différence essentielle se présente lorsque les caractéristiques de l'offre de chaque itinéraire, et donc le coût généralisé de transport, dépendent du trafic, différence particulièrement sensible lorsque le réseau des arcs de transport est maillé et fortement fréquenté.

2.4.1. La formule d'Abraham

Cet effet est souvent négligeable pour les trafic interurbains, lorsque les débits restent faibles et qu'on se situe dans la partie plate de la courbe vitesse-débit. Alors les coûts de transport sont peu influencés par le niveau de trafic et les méthodes de choix modal peuvent se transporter au problème du choix d'itinéraire. Ainsi en France, le choix entre itinéraires routiers est effectué par la Direction des Routes (SETRA) depuis plusieurs décennies en utilisant la loi d'Abraham (du nom de son inventeur, Claude Abraham) :

$$\frac{T_1}{T_2} = \left(\frac{C_2}{C_1} \right)^a$$

formule dans laquelle :

- T_1 et T_2 sont les trafics des deux itinéraires concurrents
- C_1 et C_2 sont les coûts généralisés sur ces itinéraires.
- a est un coefficient auquel l'analyse statistique donne des valeurs comprises en général entre 8 et 10.

Cette formule fournit des résultats voisins de ceux de la courbe logistique. On peut d'ailleurs la déduire comme relation approchée d'un modèle de choix discret dans lequel les coûts généralisés ressentis par chaque usager i sont :

$$C_i^j = C_j + \varepsilon_i^j$$

$$C_j^i = C_j + \varepsilon_j^i$$

les ε_j^i et ε_j^i étant des variables indépendantes réparties uniformément.

2.4.2. Le premier principe de Wardrop

Cette procédure n'est plus valable pour le trafic urbain, car on s'y trouve fréquemment en situation proche de la saturation, et les coûts généralisés dépendent des débits sur chaque arc. Le traitement est en outre compliqué par la structure de réseau des arcs de transport. Le problème est résolu par la mise en œuvre du premier principe de Wardrop, qui s'exprime ainsi : « À l'équilibre, les coûts généralisés sur chaque itinéraire utilisé entre une origine et une destination données sont égaux et inférieurs aux coûts généralisés sur les itinéraires non utilisés ».

Pour appliquer ce principe (Leurant, 1996 ; Florian, 1991), posons les notations suivantes :

- $a \in A$: arc,
- v_a : flux sur l'arc a ,
- $i \in I$: couple origine-destination,
- $k \in K_i$: itinéraire, l'ensemble des itinéraires reliant le couple O, D i étant K_i ,
- h_k : le flux sur chaque itinéraire K ,
- δ_{ak} : l'indice d'appartenance de l'arc a à l'itinéraire k ,
- g_i : le flux total concernant le couple O, D i .

On a les relations de compatibilité :

$$v_a = \sum_{i \in I} \sum_{k \in K_i} \delta_{ak} h_k$$

$$g_i = \sum_{k \in K_i} h_k$$

$$h_k \geq 0$$

et la relation qui définit le coût de transport sur l'arc a :

$$s_a = s_a(v_a)$$

le coût de l'itinéraire k sera

$$s_k = \sum_a \delta_{ak} s_a(v_a)$$

Soit u_i défini par :

$$u_i = \min_{k \in K_i} s_k$$

On fait l'hypothèse que

$$g_i = g_i(u_i)$$

On suppose que les fonctions $s_a(v_a)$ sont croissantes, et les fonctions $g_i(u_i)$ sont décroissantes. Le premier principe de Wardrop dit alors que pour les valeurs d'équilibre notées par une * :

$$s_k^* - u_i^* = 0 \quad \text{si } h_k^* > 0$$

$$\geq 0 \quad \text{si } h_k^* = 0$$

Les valeurs d'équilibre vérifient en outre les contraintes définies plus haut. La relation précédente peut aussi s'écrire :

$$(s_k^* - u_i^*)(h_k - h_k^*) \geq 0$$

h_k étant une répartition quelconque vérifiant les contraintes. En sommant ces relations sur tous les itinéraires et tous les couples O, D , on trouve :

$$\sum_k \sum_i s_k^* (h_k - h_k^*) \geq \sum_i u_i^* (h_k - h_k^*)$$

ou encore

$$\sum_a s_a^* (v_a - v_a^*) \geq \sum_i u_i^* (g_i - g_i^*)$$

Cette inégalité s'interprète ainsi : toute variation au voisinage de l'équilibre coûte (membre de gauche) plus que ce que les usagers sont prêts à payer (membre de droite). On démontre que, avec les hypothèses de monotonie et de continuité des fonctions s_a et g_i , l'inégalité précédente a une solution unique. On démontre aussi que l'inégalité précédente est équivalente au problème d'optimisation suivant (Beckmann, 1956) :

$$\min Z = \sum_a \int_0^{v_a} s_a(x) dx - \sum_i \int_0^{g_i} u_i(x) dx$$

avec les mêmes contraintes que précédemment. La plupart des logiciels de recherche d'équilibre utilisent cette équivalence avec un problème d'optimisation, pour lequel existent des algorithmes de résolution.

On peut prendre une vue plus intuitive de ces résultats dans le cas simple où il n'y a qu'un couple O, D et 2 itinéraires 1 et 2, et où le trafic total est fixe

$$g(u) = T$$

On peut alors procéder à une représentation graphique, en comptant le trafic sur l'itinéraire (et arc 1) de gauche à droite à partir de 0, et le trafic sur l'itinéraire (et arc 2) de droite à gauche à partir de T puisque :

$$v_1 + v_2 = T$$

Les ordonnées représentent les coûts. Le point d'équilibre de Wardrop est défini par :

$$v_1^*, v_2^* \text{ et } s_1^* = s_2^*$$

C'est en effet le seul qui vérifie :

$$\min \left(\int_0^{v_1} s_1(x) dx + \int_0^{v_2} s_2(y) dx \right)$$

(formulation d'optimum), ou :

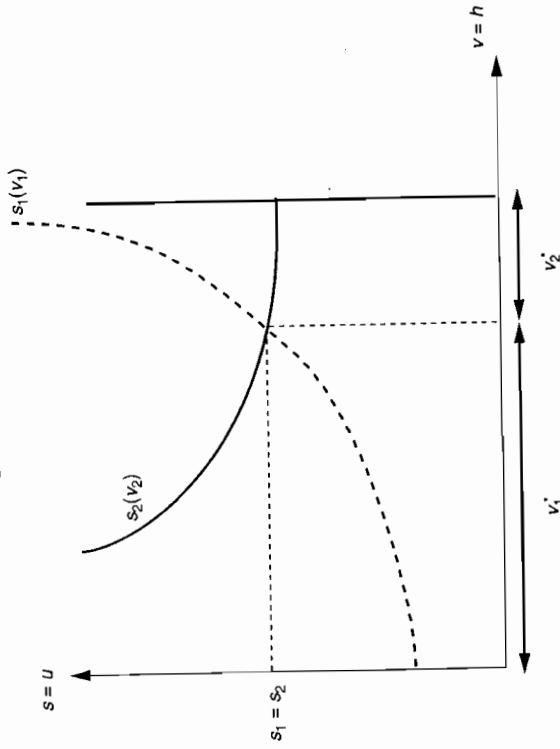
$$s_1^*(v_1^*) (v_1 - v_1^*) + s_2^*(v_2^*) (v_2 - v_2^*) \geq u(T - T) = 0$$

(formulation d'inégalité), avec :

$$v_1 + v_2 = T$$

FIGURE 4-6

Équilibre de Wardrop



2.4.3. Le deuxième principe de Wardrop

On remarque sur cet exemple que l'équilibre obtenu par le premier principe de Wardrop, celui d'optimisation pour l'usager, est en général différent de l'optimum collectif. Ce dernier correspond au deuxième principe de Wardrop ; dans le cas de l'exemple présenté, il sera obtenu par le programme :

$$\min_{v_1, v_2} (v_1 s_1(v_1) + (v_2 s_2(v_2)))$$

et aboutit à la valeur v_1^* définie par :

$$s_1(v_1^*) = s_2(T - v_1^*) - v_1^* \frac{ds_1}{dv_1} + (T - v_1^*) \frac{ds_2}{dv_1}$$

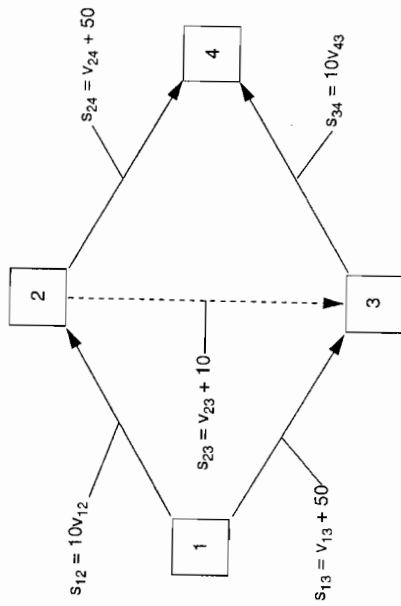
Dans le cas général l'optimum collectif sera défini par :

$$\min Z' = \sum_a v_a s_a(v_a) - \sum_i \int_0^{g_i} u_i(y) dy$$

Ce programme ne conduit pas, en général, au même résultat que le programme de minimisation précédent. L'écart entre l'optimum pour l'usager et l'optimum collectif est la source de paradoxes, dont la méconnaissance peut conduire à nombre d'idées fausses en termes de gestion de la circulation et de politiques d'investissements.

Le fameux paradoxe de Braess (1968) montre qu'il peut être antiéconomique de construire un nouveau barreau dans un réseau ; soit le réseau d'arcs à sens unique défini par la figure 4-7, pour lequel existe un seul couple O. D composé de 6 usagers allant de 1 à 4, et de deux itinéraires.

FIGURE 4-7
Le paradoxe de Braess



On vérifie que le premier comme le deuxième principe de Wardrop conduisent à une répartition de 3 usagers sur l'itinéraire ① ② ④ et 3 sur l'itinéraire ① ③ ④.

Ajoutons maintenant à ce réseau un nouveau barreau orienté de ② vers ③ et pour lequel :

$$s_{23} = v_{23} + 10$$

On vérifie que l'équilibre (1^{er} principe) est obtenu en répartissant également les 6 usagers entre les 3 itinéraires ① ② ④, ① ③ ④ et ① ② ③ ④, alors que l'optimum collectif (2^e principe) est encore obtenu par un partage égal entre les deux premiers itinéraires, comme c'était le cas dans la première configuration de réseau, où le barreau vertical n'existait pas.

Un autre paradoxe connu dans la littérature a été développé notamment par Villé (1970) : deux itinéraires assurent le même déplacement entre une origine et une destination urbaines. Le premier est constitué par exemple par une rocade longue et peu encombrée, et le coût généralisé y est donc indépendant du trafic. Le deuxième est un itinéraire urbain court et encombré. Le partage des usagers se fait entre les deux de telle manière que les coûts généralisés sur chacun d'eux soient égaux. Une analyse « myope » de l'encombrement du deuxième itinéraire peut inciter à l'élargir pour réduire la congestion. Après l'élargissement, la répartition du trafic va changer, et les utilisateurs de l'itinéraire urbain vont être plus nombreux ; mais leur coût de transport ne va pas changer puisqu'il est toujours

égal au coût de l'itinéraire long, qui est constant. L'élargissement n'aura fait que déplacer des usagers d'un itinéraire sur l'autre, sans gain de temps. L'exemple est encore plus probant si l'itinéraire alternatif de coût stable est en fait constitué par une ligne de transport public. Alors la baisse du trafic sur cette ligne conduit logiquement l'opérateur à baisser ses fréquences, et le coût généralisé sur le transport public augmente, donc le coût sur l'itinéraire routier urbain aussi, et finalement la situation s'est globalement détériorée.

L'équilibre jusqu'ici analysé est déterministe, en ce sens que le coût généralisé de chaque itinéraire n'est affecté d'aucune incertitude. Il correspond aux affectations par tout ou rien dans le problème du choix modal. Il existe aussi des modèles stochastiques, dans lesquels les coûts sur chaque itinéraires sont affectés d'un aléas ; cet aléas peut représenter des différences de perception par rapport à la réalité, ou des différences de goûts entre individus. Le coût généralisé est alors de la forme :

$$s_h^j = s_h + \varepsilon_h^j$$

j représentant les individus. Selon les hypothèses faites sur les ε_h^j , on obtient des modèles logit et probit avec ou non corrélation des aléas entre eux. Ainsi, avec l'hypothèse d'aléas indépendants et identiquement distribués selon une loi de Gumbel, on aboutit à la loi de répartition entre itinéraires :

$$\Pr(h) = \frac{\exp(-\mu s_h)}{\sum_k \exp(-\mu s_k)}$$

Toutefois, l'hypothèse d'aléas identiquement distribués est peu vraisemblable ; il est plus raisonnable de supposer que leur variance est liée au coût généralisé :

$$\text{var}(\varepsilon_h^j) = k\sigma_h^2$$

L'expérience confirme ce que l'intuition laisse prévoir, à savoir que les modèles stochastiques sont plus adaptés lorsque les réseaux sont simples et peu congestionnés et les coûts peu dépendants du trafic, auquel cas les modèles déterministes donnent une affectation par tout ou rien trop brutale, en ne sélectionnant qu'un itinéraire, un comportement qui traduit mal la réalité des choix, toujours entachés d'aléas (Florian, 1991 ; Kawakami, 1989).

3. La modélisation en pratique

Les raffinements théoriques qui ont fait l'objet des sections précédentes contrastent avec les imperfections des données utilisées pour leur mise en œuvre.

3.1. L'information statistique

Les recueils de données sont long et coûteux alors qu'en général les études ont des délais courts et des budgets limités¹. Aussi le choix des modèles est bien souvent dicté par la disponibilité des données et par le coût de leur obtention plus que par l'adéquation théorique au problème posé.

ENCADRÉ 4-2

Arbitrage entre qualité du modèle et qualité des données

La problématique de l'arbitrage entre qualité du modèle et qualité des variables explicatives peut être illustré par l'exemple suivant, inspiré de Williamsen et Ortuzar (1990). Soit deux spécifications pour expliquer une variable y . La première, le modèle parfait sans erreur, fait intervenir deux variables x_1 et x_2 connues chacune avec erreur (les erreurs sont indépendantes)

$$y = a_1 x_1 + a_2 x_2$$

La seconde spécification est approchée ; elle ne fait intervenir qu'une variable x connue parfaitement. Mais le modèle comporte une erreur bien sûr :

$$y = cx + \varepsilon$$

Avec le premier modèle, la variance de la prévision est :

$$(\sigma_y^1)^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2$$

Avec la seconde, elle vaut :

$$(\sigma_y^2)^2 = \sigma_\varepsilon^2$$

Si les variables x_1 et x_2 sont connues avec une trop forte erreur, il devient avantageux d'utiliser le modèle fruste. Par ailleurs, les efforts d'amélioration de l'information doivent, à coût égal, porter en priorité sur la variable explicative la moins bien connue (ou plus précisément celle pour laquelle $a\sigma$ est le plus élevé).

Les informations relatives au trafic peuvent venir de comptages ou d'enquêtes. Les comptages prennent différentes formes : comptage automatique ou manuel des véhicules passant en un point d'une route ; comptage des billets vendus pour les chemins de fer ou l'avion ; comptage des montées et descentes à chaque arrêt pour les lignes d'autobus. Leur coût est modeste, mais ils ne fournissent pas les origines et destinations. On peut toutefois les reconstituer, comme le fait la Direction des Routes dans les cas, de plus en plus rares, où elle ne dispose pas d'enquêtes spécifiques : La distribution des

1. Dans les études relatives à une infrastructure nouvelle, il est frappant de constater la disproportion entre les sommes consacrées à la prévision des coûts de construction et celles consacrées à la prévision des trafics et des recettes.

trafics est supposée suivre une loi gravitaire et la répartition entre itinéraire s'effectue selon la formule d'Abraham : la connaissance des comptages sur chaque arc du réseau permet de caler les coefficients des lois et d'en déduire les matrices origine-destination, et la manière dont elles se déforment lorsque les coûts généralisés de circulation changent.

Plus coûteuses en général sont les enquêtes auprès des usagers, qui peuvent s'effectuer soit durant le transport, soit à domicile. Pour le trafic routier et pour les transports urbains, les enquêtes durant le transport sont difficiles à administrer, et de toute façon légères ; elles prennent préférentiellement la forme d'une distribution de questionnaires qui sont ensuite renvoyés par la poste, et sont utilisées dans le cadre d'enquêtes-cordon, où l'on cherche à cerner les échanges entre une zone (une ville, une région) et le reste du monde. En revanche, les enquêtes auprès des usagers du chemin de fer ou de l'avion sont plus aisées à réaliser.

Les enquêtes *O.D* les plus complètes sont les enquêtes à domicile, par interview face à face ou téléphonique. On peut alors poser un bien plus grand nombre de questions portant à la fois sur les caractéristiques de la personne interviewée, du ménage et des déplacements auxquels on s'intéresse. Une forme particulière de ces enquêtes est constituée par les carnets de voyage : en début de semaine, on distribue aux personnes enquêtées un carnet sur lequel elle reportera tous ses déplacements de la semaine : horaire, coût, moyen de transport, itinéraire, motif du déplacement. Puis on relève ces carnets en fin de semaine.

Une autre forme d'enquête est destinée à s'insérer dans des études de préférences déclarées. On propose à la personne interviewée, à laquelle on aura demandé les renseignements nécessaires sur ses caractéristiques, le choix qu'elle ferait entre différentes alternatives fictives qui lui sont préalablement décrites, et ses réponses sont ensuite utilisées pour calibrer un modèle de choix. Cette méthode est spécialement bien adaptée lorsqu'il s'agit d'apprécier la clientèle d'un mode nouveau de transport pour lequel on n'a pas d'expérience. Elle nécessite néanmoins de très grande précautions de mise en œuvre : les alternatives proposées à l'enquêté doivent être très soigneusement décrites, être le plus proche possible de situations auxquelles il est habitué. On propose à chaque individu plusieurs choix, entre lesquelles les caractéristiques (prix, temps, fréquence...) du mode nouveau et du mode traditionnel alternatif changent de façon à constituer un plan d'expérience. La méthode présente moins de sûreté que l'observation des choix réels.

Une dernière forme d'enquête mérite d'être mentionnée, c'est celle qui est constituée par les panels longitudinaux, à travers lesquels on pose à intervalles réguliers les mêmes questions à un échantillon de personnes. Ces enquêtes permettent d'analyser les liaisons inter-temporelles affectant les comportements et leurs changements. Elles sont spécialement utiles pour les opérateurs qui cherchent en permanence à surveiller leur marché et à déceler le plus rapidement possible les inflexions dans les comportements et désirs de la clientèle.

Le critère de jugement de ces modèles le plus immédiat est la convivialité : commodité d'emploi, rapidité d'exécution, lisibilité des produits, éventail des options ouvertes. Un autre critère moins évident est celui des hypothèses scientifiques de base : qualité du modèle de base, convergence de l'algorithme. Enfin, un troisième critère, qui ne peut être apprécié que par une étude approfondie, est celui du choix des paramètres, fonctions, et hypothèses intermédiaires nombreuses qui parsèment ces énormes modèles, choix qui résulte du calage progressif de l'outil au fur et à mesure de son utilisation. Un quatrième critère est celui de l'adéquation entre prévisions et réalisations.

Le deuxième et le troisième critère sont d'autant plus difficiles à déceler que ces logiciels deviennent de plus en plus des objets commerciaux, vendus par leurs promoteurs à travers des centres de recherche et des bureaux d'étude. Aussi, à défaut d'une analyse approfondie, on se contentera de citer quelques uns d'entre eux :

- DAVIS et ses variantes. Mise au point par l'INRETS, il réalise l'équilibre de choix d'itinéraires d'après le premier principe de Wardrop, selon plusieurs modalités : avec ou sans contrainte de capacité, avec ou sans différenciation de la valeur du temps selon les usagers.

- OPERA : modèle urbain mis au point par le CETE ; réalise les 4 étapes : distribution selon un modèle gravitaire, choix modal agrégé selon une courbe logistique, choix d'itinéraires.

- THERESE : mis au point par le CERTU, centré sur les transports collectifs.

- EMME 2 : mis au point par le CRT Montréal, réalise les 4 étapes, et offre pour chaque étape l'option entre plusieurs modèles (agrégé, désagrégé, logit, probit, emboîté ou non...).

- MINUTP : mis au point par Cambridge Systematics. Offre les mêmes possibilités que le précédent.

- TRIPS : mis en point par MVA consultancy, réalise les mêmes types de fonctions que les deux précédents.

- MATISSE : modèle mis au point par l'INRETS pour les déplacements interurbains, définit de façon intégrée la génération, la distribution, le choix modal et le choix d'itinéraires, avec rétroaction entre les étapes, à travers une segmentation fine de la clientèle.

- IMPACT : modèle mis au point par la RATP, consistant en sous-modèles logit emboîtés, et modélisant les déplacements en région Ile-de-France.

- Modèle de la DREIF : modèle mis au point par la DREIF pour la Région Ile-de-France, utilisant une matrice *O . D* fondée sur l'enquête à domicile périodique, un partage modal agrégé selon une courbe logistique.

- ARIANE : mis au point par la Direction des Routes. Uniquement routier, fondé sur la loi gravitaire et sur la loi d'ABRAHAM.

- PIANO : ce modèle de la SNCF est fondé sur la loi gravitaire pour les phases de génération-distribution, sur le modèle prix-temps probit pour la répartition fer-avion, et sur une loi logistique agrégée pour la répartition train de jour-train de nuit.

Une autre catégorie de données doit être rassemblée pour les études de trafic, celles qui concernent le réseau. L'aire géographique étudiée doit être découpée en zones ; il faut pouvoir affecter à chaque zone les caractéristiques démographiques et socio-économiques qui la concernent et qui seront utilisées pour la modélisation : population, revenus, nombre d'emplois, taux de génération et d'attraction... Le zonage est à adapter au problème étudié : par exemple, dans l'étude de la création d'une section d'autoroute nouvelle, il sera très étroit au voisinage de l'autoroute et de plus en plus large au fur et à mesure qu'on s'en éloigne. Ensuite le réseau de transport reliant les zones est défini par ses arcs et par ses nœuds. À chaque arcs sont affectés son coût de transport, éventuellement variable avec le trafic, sa capacité, la fréquence ou les horaires pour les transports publics. Les nœuds, croisement d'arcs, peuvent comporter des pénalités pour prendre en compte le coût et le temps de parcours à l'intérieur d'un échangeur routier, ou le temps de changement dans une gare ou une station de métro ou le temps de traversée d'un carrefour.

La masse des informations statistiques à rassembler est donc considérable. Si l'on veut étudier un projet assez important et prendre en compte le trafic des arcs du réseau qui lui sont soit complémentaires, soit concurrents, on arrive très vite à un nombre de centroïdes de l'ordre de 100 à 200, c'est-à-dire que la matrice *O . D* comportera de 10 000 à 40 000 cases. Pour remplir ces cases de, mettons 10 réponses en moyenne, il faut de 100 000 à 400 000 déplacements, soit pour des trajets urbains quotidiens, environ 50 000 à 100 000 enquêtes, ce qui représenterait un coût considérable ; ainsi, sauf à dépenser des sommes en général hors de portée dans les recueils d'information, de nombreuses cases seront très peu remplies et les trafic correspondants connus avec une grande incertitude. On comble cette lacune en utilisant une formule gravitaire calée sur les données disponibles.

On pourrait songer à réduire le nombre de zones pour saisir avec plus de précision de trafic entre zones ; mais on perdrait alors sur la connaissance du trafic intrazone, et sur les coûts de circulation à l'intérieur de chaque zone, qui sont d'autant plus mal connus et pèsent d'un poids d'autant plus fort dans les coûts interzones que chaque zone est grande. Il y a un compromis à faire entre la précision sur les trafics émis et reçus par zone, et la précision sur les coûts de circulation entre zones.

3.2. Les logiciels de modèles de trafic

Les études de trafic conduisent à manipuler des quantités énormes de données. En outre, le traitement de ces données nécessite le recours à des chaînes de calculs complexes. On comprend dans ces conditions le développement des modèles informatiques qui assurent à la fois la gestion des données de base, la présentation des résultats, et le traitement scientifique intermédiaire.

3.3. Comparaisons prévisions-réalisations

La multiplicité des modèles pourrait donner à penser qu'on dispose d'outils bien adaptés à leur fonction, et que la précision des prévisions est bonne. Ce n'est apparemment pas le cas si on compare les prévisions et les réalisations, comme le montre par exemple le tableau joint relatif aux prévisions des trafics autoroutiers en France (Setra, 1993)

TABLEAU 4-4

Comparaison trafic prévu - trafic réel sur autoroutes concédées

Sections	Coefficient du redressement	Trafic prévu redressé	Trafic réel/ Trafic prévu redressé en pourcentage
A 8 - Aix-Fréjus	1,15	31 000	- 3,2
A 61 - Toulouse-Narbonne	1,15	18 700	+ 5,3
A 10 - Poitiers-Bordeaux	1,17	16 000	+ 6,3
A 81 - Le Mans-Rennes	1,16	20 300	- 21,6
A 40 - Macon-Genève	1	9 000	+ 54,4
A 42 - Lyon-Pont d'Ain	1,03	9 300	+ 48,4
A 62 - Bordeaux-Toulouse	1,19	13 800	- 5,8
A 11 - Angers-Nantes	1,16	15 800	- 25,3
A 31 - Toul-Langres	1,22	11 200	+ 0,9
A 11 - Le Mans-Angers	1	6 500	+ 69,2
A 72 - Clermont-Ferrand-Saint-Étienne	1,17	13 300	- 21,2
A 71 - Bourges-Clermont-Ferrand	1	8 400	- 12,5
A 51 - Aix-Manosque	1,11	11 700	- 11,1
A 26 - Reims-Arras	1	9 800	- 15,3
Moyenne	1,12	14 180	- 2,0

Les études de trafic effectuées par la RATP ont fait apparaître des résultats assez imprécis au début. Mais, au fur et à mesure que le modèle se calait au fil des études, la précision s'améliorait et pour la ligne A dont la configuration s'est modifiée entre 1982 et 1987, les trafics successifs ont été prévus avec une erreur inférieure à 10 % (cf. RATP, note 9006 « Prévision de trafic sur la ligne A issues du modèle global »).

Une recension effectuée par Flyvbjerg et Skamris (1996) aboutit à la même impression. Ces auteurs analysent les écarts entre prévisions et réalisations pour différents grands projets à travers le monde. Ils font apparaître des erreurs importantes, et en outre biaisées dans le sens de la sur-estimation, montrant des biais stratégiques de la part des analystes.

TABLEAU 4-5

Différences entre prévisions et réalisations pour plusieurs sortes d'études

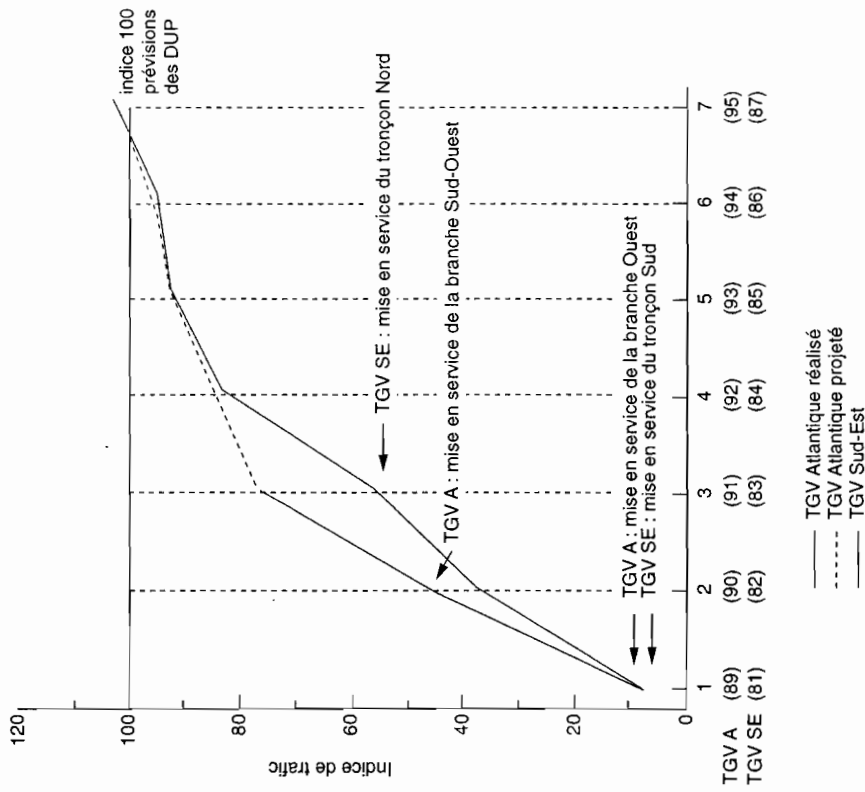
Nature des projets	Sous-estimation 0 à 20 %		Suresstimation -20 % -50 % au-delà	
	1	1	2	5
Métros en pays en développement (source TRRL)	1	1	2	5
Routes au Royaume-Uni (source TRRL)	5 %	10 %	25 %	60 %
Projet de ponts et tunnels au Danemark	1		2	
10 projets ferroviaires suburbains aux USA				Suresstimation moyenne de 61 %

Le sentiment d'insatisfaction que peut laisser la lecture de ces chiffres doit toutefois être tempéré. D'abord, en matière de trafic, il est difficile de comparer des choses comparables. Il est rare qu'un projet se réalise de la manière exacte définie dans l'étude de trafic. Or la comparaison ne peut valablement être faite que si le projet lui-même et ses conditions d'exploitation sont conformes à la réalité ultérieure. Le métro automatique Orly-Val a connu à sa mise en service un trafic à peu près 3 fois plus faible que les prévisions. Mais les auteurs de l'étude ex-ante ont fait valoir les différences considérables entre le projet de réalisation : fréquences inférieures, temps d'attente accru, prix plus élevé, mauvaise signalisation...

Un autre facteur tient à ce que le trafic met du temps à prendre sa valeur d'équilibre. La SNCF a établi le graphique suivant qui fait apparaître pour les TGV un délai de mise en charge de l'ordre de 3 à 5 ans.

Quand les comparaisons sont faites sur des années éloignées de la mise en service, à l'erreur sur la modélisation des trafics proprement dite s'ajoute l'erreur sur la prévision d'évolution des trafics ; en général les évolutions temporelles passent par l'intervention de variables exogènes comme le PIB, le prix du pétrole, dont la prévision est sujette à de nombreux aléas. La figure 4-9, relative aux prévisions de trafic routier total, montre en outre que la prévision est de plus en plus difficile au fil des ans, en raison de l'imprévisibilité croissante des évolutions. Il montre aussi que parfois la prévision à moyen terme est aussi difficile que la prévision à long terme.

FIGURE 4-8
TGV Atlantique et Sud-Est : comparaison des trafics prévus et réalisés

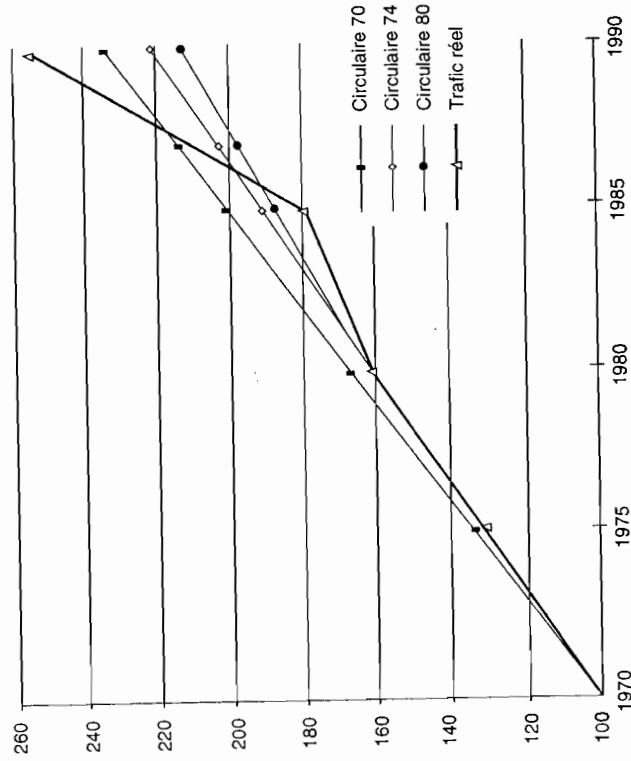


On voit ainsi se dégager les conditions d'une bonne adéquation entre les modèles et la réalité :

- une analyse impartiale,
- de bonnes données de base,
- des investissements d'un type déjà analysé,
- une concordance entre le projet analysé et la réalité ultérieure,
- une certaine patience dans la comparaison avec la réalité (3 ans au moins),
- une correction éventuelle des dérives temporelles des données exogènes.

FIGURE 4-9
Prévision et évolution réelles du trafic en rase campagne

	1970	1975	1980	1985	1987	1990	Écart réel/ prévu en %
Circulaire 70	100	133	167	200	213	233	9,4
Circulaire 74	100	130	160	190	202	220	15,9
Circulaire 80			162	186	196	211	20,9
Circulaire 86				180	187	198	28,8
Circulaire 89					208	233	9,4
Trafic réel en rase campagne	100	130	162	180	208	255	-



Source : SETRA, 1993.

3.4. Des voies de progrès

Il y a des raisons profondes pour que les modèles reproduisent mal la réalité ; elles tiennent à leur horizon temporel mal défini et à ce que le caractère dérivé de la demande de transport est mal prise en compte.

3.4.1. Les échelles de temps courts et les modèles dynamiques

À un bout de l'échelle de temps, celui des durées inférieures à l'heure, les modèles qu'on vient de voir sont mal adaptés car ils raisonnent sur des

situations stationnaires et équilibrées, alors qu'on sait bien que les conditions de circulation changent rapidement entre le moment du départ de chez soi et celui de l'arrivée, surtout dans les situations de congestion, les plus intéressantes à modéliser : les usagers s'adaptent par des processus d'apprentissage, utilisent leur expérience passée et l'information dont ils disposent pour faire face à des situations chaque jour différentes. La prise en compte de ces éléments passe par le développement des modèles dynamiques qui intègrent explicitement les changements de la demande.

Les modèles dynamiques ne s'appliquent pour le moment qu'au choix d'itinéraires et s'affranchissent de l'hypothèse de stationnarité. Ils tiennent compte explicitement des variations temporelles de la demande. Pour cela, le temps est séparé en intervalles de l'ordre de la minute ou fraction de minute, les flux entrant et sortant sur chaque arc sont datés, ainsi que les lois de conservation des flux. Selon les modèles les usagers voient leur date de départ imposée ou au contraire la choisissent compte tenu des conditions de transport et de leur désir d'une date d'arrivée souhaitée, avec pénalités en cas de retard ou d'avance. Le principe d'affectation est une généralisation du premier principe de Wardrop : pour chaque paire origine-destination, les coûts totaux pour chaque couple O, D (itinéraire \times date de départ) possible sont égaux et minimaux pour les couples O, D dont les trafics ne sont pas nuls. Les coûts pour tous les autres couples O, D non choisis (trafic nul) sont supérieurs. La solution passe par des algorithmes de solutions des inéquations correspondant à cette généralisation du premier principe de Wardrop (Wie, Tobin, Bernstein et Friesz, 1994).

Jusqu'à maintenant, ces modèles restent à l'état de recherche, et n'ont été utilisés qu'à titre expérimental sur des réseaux fictifs simplifiés. De Palma et Marchal (1996) présentent le modèle Métropolis qui offre l'avantage de pouvoir être utilisé sur un réseau de taille raisonnable (il a été mis en œuvre sur la ville de Genève et son extension à d'autres agglomérations est en cours). Les comportements des usagers ne sont pas modélisés en cherchant la réalisation du principe de Wardrop, mais par un processus de choix en rationalité limitée avec processus d'apprentissage d'un « jour » sur l'autre. Le modèle comporte une étape de modélisation de la recherche du stationnement. Une autre caractéristique du modèle est qu'il n'implique pas au départ la connaissance désagrégée des caractéristiques des usagers (caractéristiques socio-économiques, heures souhaitées d'arrivée...), mais qu'il les détermine à partir d'hypothèses sur la forme des lois de répartition qu'elles suivent, le modèle déterminant par calibrage les paramètres dont dépendent ces lois. Les exigences en matière de données sont donc faibles. Les sorties du modèle sont les répartitions horaires des instants de départ par noeud, ainsi que la répartition horaire des trafics sur chaque arc.

Delons (1997) a modélisé le choix horaire du départ sous la forme d'une cinquième étape, annexe de la quatrième relative au choix d'itinéraire. Il considère le choix combiné de l'itinéraire et de l'heure de départ

qu'il fait obéir à un principe de Wardrop généralisé : les itinéraires et horaires choisis ont le même coût, les autres ont un coût supérieur. Les valeurs du temps sont réparties selon une loi log-normale. Il ne s'agit donc pas à proprement parler d'un modèle dynamique, mais il rend des services voisins. Ce modèle a testé ce modèle à partir d'une expérience de modulation tarifaire qui s'était déroulée sur le réseau d'autoroutes à péage Cofiroute à l'arrivée à Paris. Le modèle reproduit très correctement la distribution horaire des trafics.

3.4.2. Les échelles de temps longues et les modèles transport-urbanisme

À l'autre bout de l'échelle des temps, celle des durées de l'ordre de l'année, les modifications de l'offre de transport entraînent des transformations multiples ; elles ne portent pas sur les choix d'itinéraires ou de modes, ni sur les heures de départ et le chaînage des déplacements ; elles concernent le nombre de déplacements et des choix plus fondamentaux tels que la décision d'achat d'un véhicule, ou encore les localisations d'emplois et d'activités ou de logements.

La difficulté est qu'on ne peut pas séparer clairement dans les modélisations examinées plus haut quel niveau de conséquences elles prennent en compte et quel niveau elles excluent. *A priori* toutes les conséquences à moyen et à long terme s'expriment à travers le trafic induit ; mais ce trafic induit est la partie la plus mal modélisée de toute la chaîne. Ceci plaide en faveur d'un développement des modèles intégrant explicitement les effets de location et relocalisation, c'est-à-dire intégrant le transport et l'utilisation des sols. Le principe de ces modèles a déjà été évoqué au chapitre II, dans le cadre de l'explication des localisations interrégionales. On présentera ici les principes de modèles spécifiquement urbains, en s'appuyant principalement sur Webster, Bly et Paulley (1988), Anas (1987) et Clément (1996). Ces modèles comportent une partie « transport », une partie « utilisation des sols », et une partie « économique ».

– La partie « transport » définit, à système de transports donné, et à localisation des activités et des résidences donnée, les flux de transport et les coûts de transports. C'est l'équilibre du marché des transports, obtenu à travers des sous-modèles du type de ceux qui ont été vu précédemment.

– La partie « économie » est en général fondée sur les tableaux entrées-sorties. La procédure la plus fréquemment mise en œuvre part de l'existence d'une activité dont le niveau est exogène et qui est appelée activité de base ; les autres activités voient leur niveau déterminé par les tableaux entrées-sorties qui fixent les niveaux de production des autres secteurs nécessaires pour assurer le niveau de production, fixé de façon exogène, du secteur de base, ainsi que les niveaux d'emploi nécessaires.

– Jusque-là aucune relation n'explique les locations des activités ou des individus. C'est à la partie « utilisation des sols » de le faire, et c'est là que les modèles se diversifient.

Citons quelques caractéristiques de certains d'entre eux sur ce dernier plan. Ainsi, dans le modèle Calutas, la localisation résidentielle s'effectue

en fonction de l'attractivité de chaque zone résidentielle possible et du coût du transport entre la zone résidentielle et le lieu de travail ; le remplissage des zones s'effectue sous la contrainte de la surface totale constructible dans la zone. La localisation des activités productives est fondée sur un indicateur d'attractivité des zones différent selon les seconds indicateurs, en tenant compte de la limitation de l'espace disponible. Le modèle Dortmund modélise les déménagements d'une période sur l'autre, à travers des modèles logit qui comparent l'utilité de la localisation actuelle et celle des nouveaux logements ; par ailleurs, dans chaque zone la quantité de nouveaux logements est déterminée en fonction des capacités de la zone, et de l'utilité présentée par les logements de la zone au cours de la période précédente. Dans le modèle Meplan les activités de base, et leur variation d'une période à l'autre sont réparties entre les zones selon des formules fondées sur les accessibilités des zones et les loyers de la période antérieure. Les autres activités sont déterminées par les matrices input-output quant à leur niveau global ; leur répartition spatiale, ainsi que celle des logements, résultent de l'application du modèle de rente foncière vu au chapitre III. Le modèle Tranus présente des caractéristiques similaires (Clément, 1996 ; Clément et Peyrton, 1997).

Le temps est en général découpé en périodes de plusieurs années ; les caractéristiques d'utilisation du sol d'une période sont déterminées par les conditions de transport de la période antérieure ; elles définissent de nouvelles conditions de transport, qui elles-mêmes détermineront à la période suivante les localisations des industries, des services, des logements et les prix des sols, et ainsi de suite.

La mise en œuvre de ces modèles est exigeante en matière de données. Pour les plus satisfaisants sur le plan théorique, leur calibrage, c'est-à-dire l'estimation des paramètres qu'ils comportent, nécessite bien sûr des données concernant les déplacements, et celles-ci doivent être différenciées selon les catégories de ménages et de motifs de déplacements, d'autant que la plupart des modèles mis en œuvre sont désagrégés. Mais, il nécessite des données de même degré de détail sur les activités économiques : niveau d'activité par branches, tableaux d'échanges interindustriels... Enfin, l'information géographique doit être abondante : occupation des sols par types d'activité, prix des sols... Leur adaptation à une situation particulière nécessite donc des études approfondies (Clément et Peyrton, 1997). Il faut néanmoins considérer que, compte tenu de l'intégration de plus en plus forte des politiques de transport et d'urbanisme, l'usage de ces modèles sera d'une utilité croissante pour les prises de décision (Wilson, 1998).

3.4.3. Les transports, demande dérivée et les programmes d'activité

La piste de recherche qu'ils constituent (voir Fischer, 1993) vise à combler une lacune des modèles classiques, en prenant explicitement en compte le fait que la demande de transport est une demande dérivée. Ils

partent de l'hypothèse que les transports ont pour but de réaliser un programme d'activité destiné à satisfaire un complexe comprenant des besoins obligés et des désirs personnels. Des contraintes plus ou moins rigides pèsent sur la réalisation de ces programmes : leur durée, leur localisation dans le temps et l'espace.

L'observation des comportements fait apparaître leur diversité : ils ne résultent pas d'une optimisation formelle, mais d'un processus d'apprentissage progressif, routinier, et aléatoire dans son aboutissement. Les essais de modélisation de ces méthodes se heurtent à la difficulté de définir des catégories claires de comportements. Celles-ci sont d'autant plus nombreuses que la décision n'est pas seulement liée à la personne ; elle dépend des caractéristiques du ménage auquel elle appartient, et qui est la véritable unité de décision.

Les méthodes de programmes d'activités ont eu deux sortes de débouchés :

- D'abord, l'établissement de logiciels modélisant l'adaptation des décisions de programmes d'activités et de transport à de nouvelles conditions de transport. Les méthodes correspondantes restent expérimentales, ne serait-ce qu'en raison de la masse de données nécessaires à leur mise en œuvre (Ben Akiva, Bowman, de Palma, 1996).

- Ensuite, de façon plus modeste, pour susciter une amélioration des modèles traditionnels, par exemple par l'incorporation de nouvelles variables explicatives socio-économiques ou une plus fine classification des catégories d'utilisateurs, dans la voie suivie par MATISSE (Morellet, 1997).

4. Les modèles de trafic de marchandises

La modélisation des marchandises a suscité beaucoup moins de réflexions que celle des voyageurs. La plupart des modèles partent d'un potentiel à transporter défini en origine et en destination, puis ils le répartissent entre les modes en fonction des performances de ces modes ; la phase suivante, celle du choix d'itinéraires, est précédée d'une étape dans laquelle on passe des flux en tonnes aux flux en véhicules, en général par l'intermédiaire d'une clé de passage. La difficulté de cette modélisation tient à ce que les décisions en matière de fret sont étroitement liées à la politique des chargeurs. Ainsi le transport est une des composantes de la politique logistique de l'entreprise qui se combine avec les politiques de stockage et d'approvisionnement pour définir les flux de transport demandés en termes de volume, de fréquences, d'exigences de délais, et d'origine ou destination. Face à cette demande le transporteur ajuste une offre en termes de combinaison de modes.

Si de tels modèles intégrant tous ces aspects peuvent être conçus, il est beaucoup plus difficile de les utiliser de façon opérationnelle, car on ne dispose en général pas des données propres à chaque entreprise qui seraient nécessaires à leur mise en œuvre.

Une modélisation moins ambitieuse et reliant les transports de marchandises à l'activité économique générale est fournie par les modèles régionaux, vus au chapitre II, qui analysent les phases de génération et distribution à travers des matrices inputs-outputs régionalisées. Dans ces procédures les besoins d'une région en un produit particulier résultent du jeu de la matrice input-output ; ces besoins sont satisfaits par la production locale et par des importations en provenance des autres régions. Les flux de transports ainsi engendrés sont liés, par des formules du type logit, aux prix de production de chaque région et aux coûts de transport entre la région origine et la région destinataire.

Les modèles de partage modal utilisent la notion de coût généralisé, incluant le coût de transport et le temps de transport. Leur mise en œuvre implique la connaissance de la valeur du temps¹. Ce paramètre est déterminé selon des méthodes analysées par Calzada et Jiang (1996) ainsi que par Jiang et Calzada (1997), tout à fait similaires à celles utilisées pour les voyageurs. Les auteurs distinguent la méthode du coût d'opportunité ou du coût des facteurs et les méthodes des préférences révélées ou des préférences déclarées. Ils ont effectué des évaluations de valeur du temps à travers un modèle logit utilisant des choix réels de chargeurs résultant d'une enquête effectuée en France en 1988. Ils font apparaître un ensemble d'une dizaine de valeurs qui s'étagent entre 150 et 400 F par tonne pour la route et entre 0,5 et 10 F pour le rail. Les résultats sont donc très dispersés, ce qui ne doit pas étonner en raison de l'extrême dispersion des besoins des chargeurs en termes de qualité de service.

De Jong et Gommers (1992), auteurs de plusieurs études citées dans le recensement précédent, ont également mis en évidence l'importance du facteur d'exactitude, qui ressort avec autant de force sinon plus que le temps total lui-même. Leurs résultats sont issus d'une analyse conjointe, effectuée sur environ 150 chargeurs répartis en France, Allemagne et Pays-Bas, qui visait à évaluer le poids respectif de différents paramètres dans le choix moyen de transport : durée du transport, exactitude, flexibilité, risque de détérioration.

5. Conclusion

Le savoir accumulé en termes d'analyse de la demande est impressionnant par la masse d'informations statistiques collectées, et par le nombre

1. Il s'agit ici de la valeur du temps pour le chargeur, et non de la valeur du temps pour l'opérateur de transport. Cette dernière est l'économie de coût de transport permise par une amélioration des temps de transports, résultant par exemple d'une nouvelle infrastructure ; elle consiste en économie de salaire du chauffeur, amortissement plus rapide du véhicule qui roule davantage dans la même durée... La valeur du temps pour le chargeur représente les gains de frais financiers résultant d'une réduction des stocks, l'amélioration des performances logistiques...

et le raffinement des modèles bâtis autour d'eux. Il est satisfaisant aussi par la connaissance et la capacité de prédiction, notamment en termes de partage modal et de choix d'itinéraires, c'est-à-dire en termes d'ajustement des comportements à court terme même si le stationnement est absent de la plupart des modèles. Certes, la précision des prédictions et le pouvoir explicatif n'atteignent pas la précision qu'on rencontre dans les domaines des sciences exactes, et on se trouve encore à mi-chemin entre l'art et la science. Néanmoins, mis en œuvre par des mains expertes et sans parti-pris, et appliqués à des situations pas trop nouvelles, ils permettent de prévoir avec une précision raisonnable, les partages modaux et les choix d'itinéraires.

Il n'en va pas de même pour les phases de génération et de distribution. Pour elles, les modèles sont plus grossiers ; ainsi, les modèles usuels de génération ne tiennent pas compte de la plus ou moins grande facilité des transports offerts à l'utilisateur, qui constituent à l'évidence un facteur de génération de trafic, à la fois pour le transport de marchandises et encore plus pour le transport de voyageurs. Les constatations empiriques que l'on fait à leur égard, comme la loi de Zahavi (constance du temps total consacré aux transports), sont approximatives et non causales : elles traduisent en fait le résultat d'un processus d'ajustement qui fait intervenir non seulement les décisions de l'utilisateur en termes de transport, mais aussi ses décisions en termes de programmes d'activités et les décisions des autres agents économiques tels que les offreurs d'emploi ou les promoteurs immobiliers. En outre, ces phases de génération et de distribution correspondent à des ajustements lents, à peine perceptibles à court terme ; et à long terme leur effet est tellement enchevêtré avec celui de l'ensemble des autres causes de changement, que l'analyse économétrique ne peut guère les mettre en évidence, et cela d'autant moins que nombre de ces causes sont hors des champs de l'économie et ressortissent à la psychologie ou à la sociologie.

Et pourtant, ces phénomènes lents sont à l'évidence majeurs sur le long terme, quand leurs effets se cumulent. On déplore souvent l'étalement des villes, l'augmentation des distances de déplacement. Même s'ils résultent pour une part d'un attrait plus grand pour des logements individuels spacieux, on ne peut pas ne pas en attribuer une part de responsabilité à des ajustements progressifs des localisations en réponse à des changements dans l'offre de transport. Leur contrôle est d'autant plus difficile que la main invisible fonctionne mal. Il n'y a pas coïncidence entre équilibre et optimum. On l'a vu de façon exemplaire avec le paradoxe de Braess.

Mais les mécanismes qui commandent la mobilité vont être profondément bouleversés par l'intrusion de l'informatique et des télécommunications. Ces innovations technologiques vont modifier à la fois les besoins de transport et les moyens de les satisfaire et peuvent constituer une chance pour la solution des difficultés du transport.

Les besoins vont être modifiés. On s'est beaucoup interrogé pour savoir si les télécommunications allaient ou non faire baisser l'activité des transports. Les analyses actuelles (Crowley, 1997 ; Madre, 1997) donnent à penser

que globalement les besoins seront accrus en quantité. À l'instar de ce qui s'est produit lors de l'arrivée du téléphone. Les besoins de contacts iront croissant dans une civilisation de plus en plus marquée par les échanges d'information, et ces besoins s'exprimeront aussi bien par les transports que par les télécommunications. Le télétravail et le téléachat viendront peut-être un peu réduire, mais surtout transformer les besoins de transports : aux déplacements pour achats des personnes, se substitueront des déplacements pour la livraison des marchandises ; le télétravail et les cassures qu'il permet dans la rigidité des horaires pourrait contribuer à un étallement des pointes. Et surtout les possibilités des télécommunications mobiles vont avoir l'effet majeur de masquer le temps de déplacement et de réduire la valeur du temps, avec les résultats qu'on peut imaginer sur la mobilité.

Le développement des télécommunications permettra aussi une meilleure gestion des trafics. En matière de marchandise, on peut s'attendre à une utilisation plus efficace des parcs de véhicules, se traduisant par un meilleur taux de remplissage et une réduction des retours à vide. En matière de circulation générale, la « route intelligente » et plus généralement les « transports intelligents », en cours de développement, offriront des possibilités d'optimisation des flux de trafic, et ouvriront des espaces de choix actuellement peu étudiés et peu explorés, notamment les reports horaires, les transferts modaux en temps réels, et le remplacement du déplacement physique des personnes par la télécommunication. Il devrait en résulter une élasticité accrue des demandes particulières qui s'expriment à un instant donné sur un axe donné.

On doit donc s'attendre, dans l'avenir, à un remodelage de la demande, qui sur le plan de l'analyse économique se traduira par la nécessaire mise en œuvre de nouveaux modèles, et la révision profonde des paramètres de valeurs du temps et d'élasticités qu'ils comportent. Comme cela se produit fréquemment en économie, c'est la transformation des caractéristiques des biens offerts qui va entraîner une mutation de la demande. L'analyse des modifications qu'elle va connaître constitue à l'heure actuelle un champ de recherche particulièrement important et urgent en économie des transports.

BIBLIOGRAPHIE

- M. Ben-Akiva et S. Lerman (1991), *Discrete choice Analysis*, The MIT Press, Cambridge, Mass.
- M. Ben Akiva, J. Bowman et A. de Palma (1996), « Activity based travel demand model systems », Communication au 25^e anniversaire du CRT, Montréal, octobre 1996.
- M. Boiteux (1994), *Transport : pour un meilleur choix des investissements*, Rapport d'un groupe de travail présidé par M. Boiteux. La Documentation Française, Paris.
- C. Calzada et F. Jiang (1996), « Comment mesurer la valeur du temps en transports de marchandises », *Synthèses SES*, juillet-août.
- L. Clément (1996), *Review of existing land use-transport models*, Rapport CERTU.
- L. Clément et D. Peyrton (1997), *Test of a land use and transport modelling system*, Mimeo, CERTU, Lyon.
- R.A. Cochrane (1975), « A possible economic basis for the gravity model », *Journal of Transport Economics and Policy*, January.
- G. de Jong et M. Gommers (1992), « Time valuation in freight transport ; method and results », PTRC 20th Summer Annual Meeting.
- J. Delons (1997), *Trafic analysis of a congestion pricing operation on the French Highway Network*, note ronéotée.
- A. de Palma et D. Rochat (1996), « Congestion urbaine et comportement des usagers : analyse de la composante horaire », mimeo, Université de Cergy-Pontoise.
- A. de Palma et F. Marchal (1996), « Métropolis : un outil de simulation du trafic urbain », *Revue Transport*.
- P.H. Derycke (1997), *Le péage urbain*, Economica, Paris.
- M. Florian (1991), « The network Equilibrium Model », Communication au séminaire « Optimierung in Verkehrsplanung » ETH Zurich.
- M. Gaudry et M. Wills (1978), « Estimating the functional form of travel demand models », *Transportation Research*, n° 12-4.
- R.M. Gonzales (1997), « The value of time : a theoretical review », *Transport Reviews*, vol. 17.
- B. Goodwin (1992), « A review of new demand elasticities with special reference to short and long run effects of price changes », *Journal of Transport Economics and Policy*, May.
- M. Hainault et J. Karam (1998), *Les modèles de choix modal. Recherche d'élasticités*, Mémoire, ENPC.
- D. Henscher (1995), « Value of travel time savings in personal and commercial automobile travel », in *The full transportation costs and benefits of transportation*, Greene, Jones and Delucchi (1997).
- L. Hivert, J.P. Orfeuil et P. Troulay (1988), « Modèles désagrégés de choix modal », Rapport INRETS n° 67, Paris.
- F. Jiang et C. Calzada (1997), « Shipper's demand characteristics and their value of time : an analysis on the freight mode choice », note, Ministère des Transports et Inrets.
- O. Johansson et L. Schipper (1997), « Measuring the long run fuel demand of cars », *Journal of Transport Economics and Policy*, Sept.

C. Abraham (1961), « La répartition du trafic entre itinéraires concurrents », *Revue Générale des Routes et des Aérodromes*, octobre.

C. Abraham et J.D. Blanchet (1973), « Le modèle prix-temps », *Revue de l'aviation civile*, juin.

A. Anas (1987), *Modelling in urban and regional economics*, Harwood Academic Publishers.

J.P. Arduin, J. Ni et O. Picq (1994), « Valeur du temps, log-normalité des revenus et choix modal », Communication à WC RR 1994 Paris.

S. Kawakami, Y. Hirobata et Z. Xu (1989), A general comparison of stochastic and deterministic equilibrium traffic assignment models. Communication, 5^e WCTR Yokohama.

L. Kechi (1996), « Synthèse de documents sur les élasticités tarifaires », CETUR note ronéotée.

J.G. Koenig (1974), *Théorie de l'accessibilité urbaine*, thèse. Université de Paris VI.

C. Lefevre et J.M. Offner (1990), *Les transports urbains en question*. Éditions CELSE.

F. Leurent (1991), « Traitement mathématique d'un modèle de choix d'itinéraire sur un réseau », note ronéotée. INRETS.

F. Leurent (1997), « Modèles de choix binaires : une comparaison sur les observations du tunnel Prado-Carénage », note de travail 97-5 INRETS, Arcueil.

B. Mandel, M. Gaudry et W. Rothengatter (1996), « A disaggregate Box-Cox logit mode choice model of intercity passenger travel in Germany », in E. Quinet and R. Vickerman (1996), *The Econometrics of Major Transport Infrastructures*, Mac Millan.

J. Massiani (1997), *La non-linéarité de la valeur du temps*, Mémoire, DEA Transport ENPC-Paris XII.

P. Merlin (1991), *Géographie, économie et planification des transports*, PUF, Paris.

Modelistica (1995), *Transus System Overview*, note ronéotée.

O. Morellet (1997), « Modèle MATISSE. Validation du partage du trafic entre itinéraires routiers à péage et hors péage », note du 23 juin, INRETS, Arcueil.

MVA (1987), *The value of travel time savings*, Rapport.

J. Ortuzar et L. Willumsen (1990), *Modelling transport*, John Wiley and sons Ltd.

T.H. Oum, W.C. Waters II et J.S. Yong (1992), « Concepts of price elasticities of transport demand and recent related estimates », *Journal of Transport Economics and Policy*, May

E. Quinet (1993), « Transport between monopoly and competition », in J. Polak and A. Heerje (1993), *European Transport Economics*, Blackwell.

E. Quinet (1994), *Rapport Route-Air-Fer*, Rapport ronéoté pour le compte du Ministère des Transports.

J. Rousseau et C. Saut (1991), « Tests de politiques de transport », RATP Études-Projets, 4^e trimestre.

J. Rousseau et C. Saut (1997), « Un outil de simulation de politiques des transports : IMPACT3 », *Revue générale des Chemins de fer*, décembre, Paris.

C. Segonne (1998), *Comportement de choix d'itinéraires*, Thèse Lyon 2.

E. Selvanathan et S. Selvanathan (1994), « The demand for transport and communication in the United Kingdom and Australia », *Transportation Research*, vol. 28 B, n° 1.

Setra (1993), *Les prévisions de trafic 20 ans après*, note ronéotée.

M. Skamris et B. Flyvbjerg (1996), « Accuracy of traffic forecasts and Cost estimates on large transportation projects », *Transportation Research Record* 1518. National Academy Press, Washington D.C.

K. Small (1992), *Urban transportation economics*, Harward Academic Publishers.

SNCF (1993), *Évaluation des grands projets ferroviaires*, note ronéotée.

H. Varian (1992), *Microeconomic Analysis*, Third edition, W.W. Norton and Company.

J. Villé (1970), *Le paradoxe de la congestion*, note ronéotée. Ministère de l'Équipement.

F.V. Webster, P.H. Bly et N.J. Paulley (1998), *Urban land use and transport interaction*, Aldershot, Gower.

B. Wie, R. Tobin, D. Bernstein et T. Fries (1995), « A comparison of system optimum and user equilibrium dynamic traffic assignments with schedule delays », *Transportation Research C*, vol. 3, n° 6.

A.G. Wilson (1998), « Land use/transport interaction models : past and future », *Journal of Transport Economics and Policy*, January.

Le coût des transports

Le chapitre précédent faisait intervenir le coût payé par l'utilisateur. Mais la notion de coût de transport, entendue au sens large, va bien au-delà de ce que l'utilisateur paie, à savoir les dépenses de l'automobiliste ou le prix du billet de transport collectif. On rencontre aussi des coûts non monétaires supportés par les usagers, sous forme par exemple de temps passé, ainsi que des coûts occasionnés au reste de la collectivité par les utilisateurs des transports ; sous cette catégorie se rencontrent les dépenses d'infrastructures, en général effectuées par la puissance publique, les coûts d'environnement ou d'insécurité ; on rencontre enfin dans les transports publics les dépenses des exploitants, plus ou moins complètement compensés par le prix payé par l'utilisateur à ces exploitants.

Le tableau 5-1, tiré de Greene, Jones et Delucchi (1997), fournit un exemple de décomposition du coût de transport en différents postes, selon ceux qui les subissent et ceux qui les occasionnent, et selon leur caractère marchand ou non, et externe ou non.

La connaissance des coûts de transports est indispensable au décideur. À un niveau micro-économique, une information fine sur les coûts, sur les paramètres dont ils dépendent et sur leurs lois de variation sont des bases incontournables pour la prise des décisions de la part des agents publics ou privés : choix d'investissements, tarification, décisions de gestion de la part des entreprises. À un niveau stratégique et macro économique, la connaissance globale des coûts à l'échelle du pays, d'une région, d'une ville, ou d'un mode éclairent les débats publics sur les choix entre modes de transports, ou sur les niveaux globaux d'investissements et de tarification, ou sur ce que devraient payer les usagers des transports compte tenu des coûts qu'ils font supporter au reste de la collectivité. Un bon éclairage des décisions nécessite de connaître non seulement chacun de ces coûts, mais aussi la manière dont ils varient et les paramètres dont ils dépendent.

TABLEAU 5-1

Les coûts de transport

Coût externe	Coût d'environnement	Faune et flore Énergie Bruit Pollution de l'air, de l'eau, des sols Paysages Vibration
Coût total	Congestion Accidents Utilisation de l'espace	
	Coûts d'infrastructure	
Coût interne	Coûts privés	Carburant Entretien Réparations Assurances, taxes Amortissement du véhicule

Les moyens pour atteindre cette connaissance sont multiples. On peut d'abord recourir à la comptabilité des entreprises ; mais la comptabilité, même analytique, se fonde sur des répartitions des coûts fixes en général arbitraires ; en outre elle ne fournit que des évaluations du coût actuel ; elle ne donne guère d'indication pour connaître la manière dont ces coûts changent quand le volume de production change, ce qu'on appelle les fonctions de coût.

Un autre moyen est de simuler les dépenses d'une entreprise qui aurait à assurer des trafics donnés, et de répéter l'opération pour une batterie suffisamment complète de trafics possibles ; cette méthode d'ingénieurs est coûteuse et longue à mettre en œuvre.

La méthode la plus fréquemment utilisée par l'économiste est de rassembler les caractéristiques de trafics et de coût total d'une série d'entreprise, et d'en déduire par ajustement statistique (par exemple par la méthode des moindres carrés) les paramètres dont dépendent les coûts. Sa mise en œuvre suppose donc une connaissance des formes possibles des fonctions de coût, de leurs propriétés, qui seront examinées dans une première section consacrée aux aspects théoriques de la fonction de coût. Puis chacune des catégories de coût précédemment recensées sera examinée : coût monétaire, dépenses de temps, dépenses d'infrastructure, environnement, insécurité, coût des opérateurs. Enfin on présentera les résultats numériques de quelques études statistiques d'ensemble de ces coûts, avant de conclure sur les enseignements à en tirer pour la prise de décision.

1. La fonction de coût, rappels

Rappelons les principaux résultats et concepts relatifs à la fonction de coût (une présentation plus détaillée est fournie dans l'annexe 4).

La fonction de coût d'un producteur est le coût minimum pour produire une quantité q (scalaire ou vecteur à composantes positives) quand ce producteur peut disposer de facteurs de production x, y, z , aux prix px, py, pz :

$$C = C(q, px, py, pz)$$

On définit à partir de là quelques notions qui interviennent dans l'analyse économique :

- Le coût marginal :

$$\frac{\partial C}{\partial q}$$
- La notion d'économie d'échelle en un point, qui présente deux définitions, l'une globale :

$$C(\lambda q) \leq \lambda C(q) \quad \lambda \geq 1$$

l'autre locale :

$$\frac{C(q)}{q} / \frac{\partial C}{\partial q} \geq 1$$

- La condition d'existence d'un monopole naturel :

$$C(q) < C(q_1) + C(q_2)$$

quels que soient q, q_1 et q_2 tels que :

$$q = q_1 + q_2$$

Les conditions suffisantes pour qu'il en soit ainsi font intervenir :

- une condition d'économies d'échelle, par exemple :

$$C(\lambda q) \leq \lambda C(q) \quad \lambda \geq 1$$

- et aussi, en cas de multiproduction, une condition d'économie d'échelle, par exemple :

$$C(q_1, q_2) < C(q_1, 0) + C(0, q_2)$$

Pour finir ce rappel, il convient de remarquer que les fonctions de production et fonctions de coût ne sont pas de purs résultats des conditions techniques de production. Elles dépendent aussi des structures de marché. Ainsi un marché peu compétitif induit une réduction des efforts de productivité de l'entreprise et accroît son coût. Le même phénomène se produit lorsqu'une entreprise travaille sous contrat pour un commanditaire, et que ce dernier, par exemple par asymétrie d'information, ne peut pas contrôler l'effort de l'entreprise.

2. Le coût monétaire pour l'usager

Dans les situations de transport public, le coût monétaire pour l'usager est en fait un prix fixé par l'exploitant, et le prix payé par l'usager ne représente qu'un transfert entre lui et l'exploitant, dont le montant résulte de la stratégie commerciale de ce dernier.

Dans les situations de transport privé, qui nous occupent seules ici, une analyse plus approfondie est nécessaire, et on la mènera en prenant pour exemple le coût pour l'automobiliste, qui comporte :

- une partie fixe, payée annuellement sous forme de vignette, assurance... ;
- une partie qu'on pourrait qualifier de demi fixe, qui est essentiellement composée de l'amortissement du véhicule, qui dépend de l'âge du véhicule et de son kilométrage.

ENCADRE 5-1

L'amortissement économique

Soit un bien durable dont la valeur décroît au fil des années pour des raisons d'usure, d'obsolescence ou autres ; appelons V_t sa valeur au début de l'année t . Quel est le coût économique de l'utilisation du bien au cours de l'année ? La réponse est donnée par la formule :

$$A_t = V_t - V_{t+1} + iV_t$$

On peut la retrouver en considérant que l'entrepreneur a financé sa possibilité d'user du bien durable pendant l'année t en empruntant pour un an la somme V_t correspondant à l'achat, ce qui lui coûte :

- des intérêts (iV_t)

- le remboursement de V_t en fin d'année, remboursement obtenu pour V_{t+1} par la revente, et pour le reste de sa poche :

$$(V_t - V_{t+1})$$

- une partie variable proportionnelle au trafic, correspondant essentiellement aux dépenses de carburant, d'entretien. Cette partie dépend un peu des conditions de circulation.

Les proportions respectives de chacune de ces parties et leur montant total dépendent du type de voiture, de son parcours annuel et des conditions de circulation qu'elle rencontre. Les prix de revient kilométriques déterminés par l'administration fiscale pour 1996 sont les suivants pour deux catégories de véhicules :

Kilométrage annuel	5 000	20 000
Puissance fiscale		
4 CV	2,52	3,57
10 CV	1,68	2,42

Ces chiffres correspondent à un prix de revient complet, incluant toutes les dépenses. La Direction des Routes, dans son instruction sur les calculs de rentabilité routiers (édition 1996), retient des chiffres moyens plus bas qui ne correspondent qu'à des dépenses marginales :

- entretien et dépréciation : $0,57 F \quad V_{\text{véh}} \times \text{km}$
- carburant : $0,50 F \quad V_{\text{véh}} \times \text{km}$

Ces chiffres sont le fruit de calculs économiques précis. Sont-ils ceux dont les automobilistes ont conscience ? La question se pose car en matière automobile, le paiement des coûts est déconnecté de l'exécution physique de la dépense et effectué de manière globale ; elle est importante dans l'étude des comportements, qui dépendent des coûts ressentis et non des coûts réels. Les enquêtes menées montrent d'abord qu'il y a un écart sensible entre le coût réel et le coût ressenti (Small, 1992 et chapitre IV, § 2.3.4.), ce dernier étant sous-estimé par rapport au coût réel. En pratique, les automobilistes prennent en compte les dépenses de carburants, mais ils sous-estiment les autres dépenses liées à l'usage de l'automobile, en particulier l'amortissement du véhicule.

3. Les dépenses de temps

En dehors des dépenses monétaires qu'on vient d'examiner, les usagers du transport y passent - y perdent - du temps. Ce temps peut être assimilé à un coût, par l'intermédiaire de la notion de valeur du temps, qui a été présentée et développée au chapitre IV. On s'attachera ici à approfondir l'aspect physique du temps passé, dont la complexité vient de ce que le temps perdu dans un transport donné n'est pas une donnée fixe, mais dépend du trafic. On exprime souvent ce fait en disant que le transport est un bien à qualité variable avec le trafic (Lévy-Lambert, 1968) ou soumis à encombrement (Kolm, 1968). Cette dépendance à l'égard du trafic s'exprime différemment selon les modes.

3.1. La congestion routière

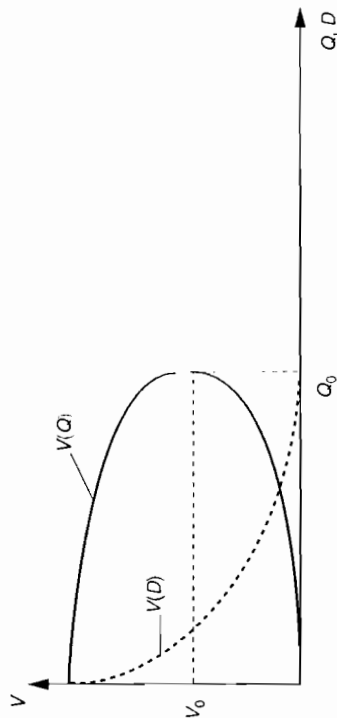
La congestion routière résulte des lois d'écoulement du trafic. En se limitant aux analyses macroscopiques (par opposition aux analyses microscopiques, dans lesquelles on analyse les comportements des véhicules pris un à un et on explique leurs interactions), et en prenant le cas de l'écoulement stationnaire d'un flot de véhicules le long d'une voie indéfinie, on repère les paramètres suivantes :

- vitesse (longueur parcourue par unité de temps) : V
 - débit (nombre de véhicules passant en un point donné par unité de temps) : Q
 - densité (nombre de véhicules sur une longueur unité de voie) : D
- On a entre ces trois quantités la relation de nature comptable, résultant de la cohérence des définitions précédentes :

$$D = \frac{Q}{V}$$

L'expérience montre en outre que la vitesse d'écoulement V décroît avec la densité D . Il en résulte que le débit Q est en relation avec la vitesse selon la courbe $V(Q)$ de la forme figurée sur la figure 5-1. La partie supérieure de cette courbe correspond à un écoulement régulier (la partie inférieure correspondant à un écoulement instable, celui qui résulte d'une faible allure moyenne composé d'arrêts et de redémarrage, que l'on constate en situation de forte densité. L'abscisse Q_0 correspond au débit maximal de la voie, qu'on appelle sa capacité.

FIGURE 5-1
Relation entre débit Q , vitesse V , densité D



Les formes algébriques les plus utilisées en France pour relier débit et vitesse sont les suivantes :

$$T = t_0 (1,1 - A \cdot Q/Q_0) / (1,1 - Q/Q_0)$$

$T(Q)$ est le temps de trajet, inverse de la vitesse.

A est un paramètre caractéristique de la voie, égal à :

- 0 : pour les voies urbaines de centre ville,
- 0,3 à 0,5 : pour les voies primaires,
- 0,666 : pour les voies rapides,
- 0,92 : sur les autoroutes.

t_0 est le temps de trajet à vide

Q est le débit compté en UVP¹

Q_0 est le débit de saturation ou capacité

Ce type de relations n'est valable que tant que $S < 1$, c'est-à-dire que le débit est inférieur au débit de saturation de la route, en général évalué à 1 800 UVP par voie. Au-delà de cette limite tantôt on utilise la formule :

$$T = 10t_0 (1,1 - A)/S^2$$

1. Les UVP sont les Unités Voitures Particulières. On compte en général :

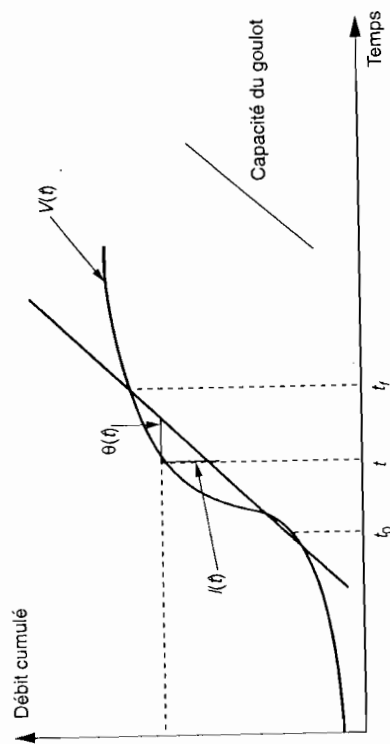
1 deux roues = 0,5 UVP.

1 poids lourd = 2 à 3 UVP.

tantôt on prolonge la courbe par sa tangente à partir d'un point voisin de la saturation, dont le choix présente l'inconvénient de l'arbitraire.

Mais ceci n'est alors qu'une approximation statistique moyenne ; lorsque le trafic rencontre un goulot dont la capacité est inférieure à ce trafic, il faut raisonner en termes de file d'attente, de la manière indiquée sur la figure 5-2, tirée de Arnott, De Palma et Lindsey (1994) ou Vickrey (1963).

FIGURE 5-2
Écoulement du trafic à travers le goulot



Signification des symboles :

$V(t)$: débit cumulé en amont du goulot en fonction du temps t

t_0 : temps de début de l'engorgement

t_f : temps de fin d'engorgement

$l(t)$: longueur de la file d'attente à l'instant t

$\theta(t)$: durée de séjour dans la file d'attente pour le véhicule qui y entre en t

3.2. La congestion aérienne

La congestion aérienne se présente de façon différente de la congestion routière, en ce sens qu'elle ne se traduit que très partiellement par un allongement du temps de trajet. Sa manifestation essentielle est un décalage par rapport à l'horaire souhaité. En effet, la plupart des passagers ont des souhaits d'arrivée précis, par exemple le début de la journée pour les voyageurs d'affaire. Lorsque le volume des souhaits d'arrivée à une certaine heure dépasse les capacités de la piste ou de l'espace aérien, on décale les vols, en avance ou en retard, par rapport à cet horaire souhaité. La ressource rare que constituent les heures de décollage ou d'atterrissage, les « slots », peut être attribuée soit arbitrairement par la puissance publique, soit par des systèmes d'enchères, soit le plus souvent par des

mécanismes qui prolongent les attributions passées et qui confèrent des sortes de droits de propriété (c'est le système des droits du grand-père). La congestion est ainsi planifiée, elle ne se manifeste à l'usager que par des horaires incommodes, les durées de vol ne sont pas augmentées, tout au moins tant qu'il n'y a pas d'incident. Lorsqu'un incident se produit et entraîne un retard pour un vol, les vols suivants, qui le jouxtent avec une marge de sécurité minimale, se trouvent retardés.

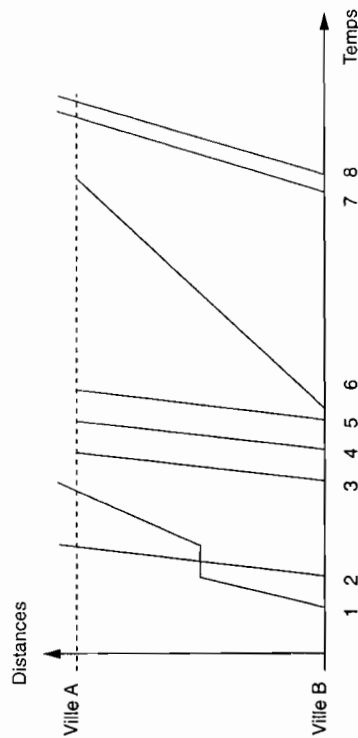
3.3. La congestion ferroviaire

La congestion ferroviaire participe du même principe que la congestion aérienne, elle se traduit par des décalages d'horaires beaucoup plus que par des suppléments de temps de trajets, mais sa représentation est rendue plus complexe car les vitesses des trains ne sont pas les mêmes et leurs arrêts peuvent différer. Ceci apparaît clairement sur un graphique des sillons horaires comme celui représenté sur la figure 5-3, relative à une ligne allant d'une ville A à une ville B.

Sur ce graphique, la marche de chaque train est représentée par une succession de lignes brisées, dont la pente est proportionnelle à la vitesse du train. Les arrêts sont représentés par des paliers. On voit sur ce graphique quelques traits caractéristiques de la congestion ferroviaire :

FIGURE 5-3

Les sillons ferroviaires



Lorsque tous les trains ont la même vitesse (cas des trains 3, 4 et 5), la capacité procurée par leur marche en batterie est très élevée. Un train lent occupe une capacité considérable (cas du train n° 6). Pour permettre aux trains rapides de ne pas être pénalisés, il faut pouvoir l'arrêter et faire passer le train rapide (cas des trains 1 et 2).

Ces caractéristiques font comprendre les traits généraux de la congestion ferroviaire en France : elle ne se produit pas en général sur les trajets en rase campagne, dans la mesure où les trains à grande vitesse ont des lignes spécialisées, et où les trains de marchandises circulent surtout la nuit, et les trains de voyageurs intercity circulent essentiellement le jour. En revanche, elle se produit au voisinage des villes, surtout aux heures de pointe, lorsque se mélangent des circulations dont les vitesses sont très différentes : des trains de banlieues aux arrêts fréquents, des trains de longues distances et des trains de marchandises qui, notamment pour les transports combinés, ont des exigences d'horaires.

3.4. La congestion dans les transports publics urbains

La congestion dans les transports publics urbains présente un tour fort différent de ce qu'on rencontre dans les autres modes. L'élément essentiel, par exemple dans le cas d'une ligne de métro, provient de la réaction de l'exploitant de la ligne, qui va adapter son service au volume de la clientèle en augmentant les fréquences lorsque la clientèle augmente. Il en résulte que le temps total de transport, correspondant à la somme du temps moyen d'attente et du temps de trajet diminue. Il y a donc une congestion à rebours, une externalité positive, dont l'encadré 5-2 présente une modélisation.

ENCADRÉ 5-2

Les coûts du transport public urbain (tiré de Small (1992))

Soit une ligne de transport public qui assure un débit q par l'intermédiaire d'autobus dont le coût d'exploitation est C_v , la capacité N , et l'espacement temporel e . Le temps de trajet est θ et la valeur du temps est prise égale à l'unité.

Le programme que doit résoudre l'exploitant s'écrit :

$$\min_e C = q\theta + q \frac{e}{2} + \frac{C_v}{e} = C^* \quad \text{avec} : \quad qe \leq N$$

Deux cas sont à considérer :

1) $qe < N$

L'espacement optimal est :

$$e = \sqrt{\frac{2C_v}{q}}$$

Le coût de l'opérateur est :

$$\frac{C_v}{e} = \sqrt{\frac{qC_v}{2}}$$

C'est une fonction de coût à rendement croissant, le coût collectif total est :

$$C^* = q\theta + \sqrt{2qC_v}$$

2) $q_e = N$

Le coût collectif total est :

$$C^* = q_e \left(\frac{N}{q} - \frac{1}{q} \right) = \frac{N^2}{q} - \frac{1}{q}$$

Le coût pour l'opérateur est :

$$q \frac{N}{q}$$

Il correspond à une fonction à rendements constants.

Le passage d'un régime à l'autre est réalisé pour un débit quel que :

$$\frac{N}{q} = \frac{2C}{q + \gamma}$$

ou

$$q = \frac{N^2}{2C}$$

Pour ce débit, le coût d'attente est égal au coût de l'opérateur.

4. Les dépenses d'infrastructures

Les dépenses d'infrastructures sont assurées soit par les pouvoirs publics (cas des routes et des voies navigables), soit par des opérateurs plus ou moins liés à l'État : la S.N.C.F. les aéroports et ports sont des établissements publics, le contrôle aérien est assuré par un service public dont le budget est autonome, les sociétés d'autoroute, publiques ou privées, sont liées à l'État par des contrats de concessions.

Tous ces organismes, qui assurent la construction, l'exploitation et la maintenance des infrastructures, présentent quelques caractéristiques essentielles. La première est que la qualité qu'ils offrent dépend de la demande, à travers les mécanismes de formation de la congestion décrits à la section précédente. La seconde est qu'ils fournissent non pas un seul, mais plusieurs produits qui se différencient par la nature du transport (voyageurs, marchandises), ou par l'heure, qui influe sur la qualité offerte à l'utilisateur. La troisième est la présence de discontinuités : on ne peut pas construire des fractions de piste aérienne, on ne peut pas faire des autoroutes avec des fractions de voies, même si pour les routes ordinaires des aménagements tels que giratoires ou carrefours dénivelés, peuvent rétablir une certaine continuité dans la variation des qualités de service et des débits (Small, 1992).

La fonction de coût type et les résultats présentés à la section 1 doivent être adaptés pour tenir compte de ces caractéristiques, qu'on retrouve le plus fortement dans le cas des routes. On analysera d'abord ce mode, puis les infrastructures aériennes. Quand au cas du chemin de fer, comme la gestion des infrastructures est liée à celle de l'exploitation des véhicules, on l'examinera dans la section 7, consacrée au coût des opérateurs.

4.1. Infrastructures routières

Les infrastructures routières produisent à la fois une quantité et une qualité de service, cette qualité étant variable avec la quantité produite c'est-à-dire le trafic. La fonction de coût les représentant doit donc intégrer les effets de ces deux paramètres, et on peut l'écrire (Small, Winston et Evans) :

$$C(q, Q) = qht(q, Q) + \rho K(Q)$$

formule dans laquelle :

- h est la valeur du temps, supposée identique pour tous les usagers,
- q est le trafic écoulé par unité de temps. S'il y a plusieurs catégories de trafic (automobiles, PL...) et que le débit varie par période, q sera un vecteur indicé par les périodes et les trafics,
- Q est la capacité maximale de la route¹,
- $t(q, Q)$ est le temps de trajet par usager, qui dépend du trafic et de la capacité de la route,
- $K(Q)$ est le coût de construction de la route de capacité Q ,
- $\rho K(Q)$ est la charge annuelle du capital comprenant l'amortissement économique et l'entretien, supposé ici pour simplifier indépendant du trafic, et proportionnel au coût de construction.

On peut, à partir de là, définir le coût marginal à court terme

$$\frac{\partial C}{\partial q} = ht(q, Q) + qh \frac{\partial t}{\partial q}$$

Si Q ne peut varier que de façon discrète, on peut tracer une courbe $C(q, Q)$ pour chaque valeur de Q , et la courbe de coût adapté sera constituée de l'ensemble des fragments de courbes $C(q, Q)$ à donné situés le plus bas, sur la figure 5-4.

Si maintenant on suppose que Q peut varier de façon continue, la fonction de coût adapté sera définie par :

$$\frac{\partial C}{\partial q} = 0$$

soit :

$$qh \frac{\partial t}{\partial q} + \rho \frac{\partial K}{\partial Q} = 0$$

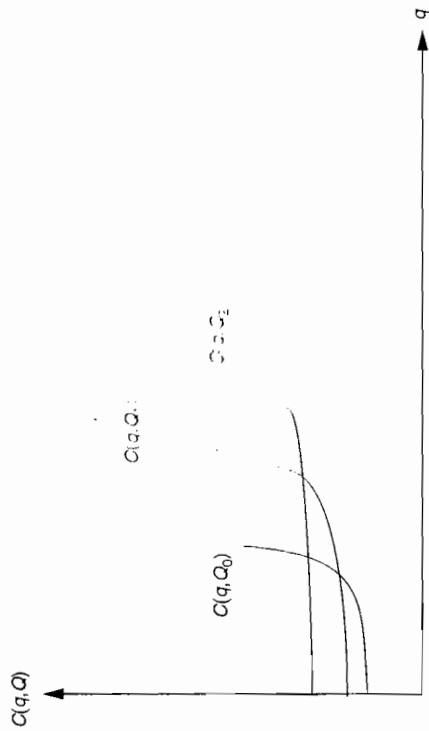
Pour éclairer la question centrale de l'échelle des rendements, il est intéressant de regarder à quelle condition ceux-ci sont constants. Cette constance sera réalisée, il est facile de le voir, si :

$$K = aQ$$

1. Si le trafic comporte plusieurs catégories de véhicules, on admet en général des équivalences d'encorements e_j , entre ces catégories q_j et la capacité maximale s'exprimera par la condition :

$$\sum_j e_j q_j \leq Q$$

FIGURE 5-1
Fonction de coût adapté en situation de capacités discontinues



et si $t(q, Q)$ est de la forme

$$t = t \left(\frac{q}{Q} \right)$$

Alors, la relation précédente s'écrit :

$$q^2 h \frac{t}{Q^2} = \frac{\rho K(Q)}{Q}$$

On en déduit que le coût marginal adapté et le coût moyen sont alors égaux. Ils valent effet, l'un :

$$\frac{dC}{dq} = ht \left(\frac{q}{Q} \right) + \frac{q}{Q} ht' \left(\frac{q}{Q} \right)$$

et l'autre :

$$\frac{C}{q} = \frac{1}{q} \left(qht \left(\frac{q}{Q} \right) + \rho K(Q) \right) = ht \left(\frac{q}{Q} \right) + \frac{q}{Q} ht' \left(\frac{q}{Q} \right)$$

Comment se situe la réalité par rapport à cette situation remarquable ? Examinons d'abord le cas de la rase campagne. Les coûts de construction croissent moins vite que la capacité, comme le montre le tableau suivant, valable pour des situations moyennes en rase campagne.

Par ailleurs, la fonction $T(q, Q)$ n'est pas de la forme :

$$t = t \left(\frac{q}{Q} \right)$$

mais elle manifeste aussi une forme de croissance des rendements, puisque, à valeur de q/Q donné, t est d'autant plus faible que Q est élevé.

Les deux facteurs jouent dans le même sens pour aboutir à l'existence de rendements croissants dans la gestion des infrastructures routières interurbaines. Il n'en va pas forcément de même en milieu urbain, car alors, si les courbes temps-débit et les conclusions qu'on tire de leur forme sont identiques, les coûts de construction, qui dépendent plus fortement du coût d'acquisition des terrains et des contraintes environnementales, peuvent faire monter considérablement les dépenses en fonction de la capacité Q .

TABLEAU 5-2
Coûts et capacités des routes interurbaines

Type de voie	Capacité horaire en Uvp	Coût moyen en MF/km
Route	2 000	16
Route	2 700	20
Route	4 000 à 5 000 (1)	29
Autoroutes	6 000	40
Autoroutes	9 000	50

(1) Dépend du nombre de carrefours plans

Sources : Mémento de la Route 1996 et Guide des Études de Trafic Interurbain 1992.

On a, dans ce qui précède, supposé que le coût d'entretien, inclus dans le terme ρK , était indépendant du trafic. C'est une simplification, et les dépenses d'entretien varient avec le trafic. Elles dépendent du poids des véhicules, selon des lois qui diffèrent selon la catégorie de dépense : ainsi, approximativement proportionnelles à la puissance quatrième du poids des essieux pour l'entretien du corps de chaussée, elles varient à peu près linéairement avec le trafic pour les dépenses de signalisation. Small, Winston et Evans (1989), Brossier (1991) et Newbery (1988) ont étudié ces dépenses et leurs lois de variation avec le trafic.

4.2. Infrastructures aériennes

Il n'y a pas beaucoup de renseignements sur les fonctions de coût des infrastructures aériennes, contrairement au cas du transport aérien proprement dit, abondamment étudié. On présentera successivement l'état de la question pour les aéroports et pour la navigation aérienne.

4.2.1. Les aéroports

Keeler (1973) a trouvé des rendements d'échelle constants à partir d'une étude économétrique portant sur des données anciennes. Doganis et Thompson (1975) arrivent au même résultat, avec une fonction Cobb-Douglas. Tolofari, Ashford et Caves (1990) ont effectué une étude de la fonction de coût des aéroports britanniques, fondée sur l'estimation d'une fonction translog. Ils étudient séparément les dépenses d'exploitation, conduisant aux coûts marginaux et moyens à court terme (hors dépenses

en capital), et la fonction de coût total, correspondant à des infrastructures adaptées (optimisées), et les coûts marginaux et moyens à long terme correspondants. Il en ressort que :

- le coût marginal à court terme est nettement inférieur au coût moyen (dans la proportion de 1 à 2) ;
- cela résulte, au moins pour partie, de ce que les infrastructures aéroporaires sont en général surdimensionnées, c'est-à-dire que la capacité réelle est supérieure à celle qui serait optimale ; ce résultat est peut-être dû aux indivisibilités techniques.
- si l'optimisation était réalisée, il en résulterait, quand même des rendements d'échelle importants, de 1,4 environ.

Enfin, la pratique aéroporuaire, notamment celle développée à Aéroports de Paris, conduit à certaines considérations allant également dans le sens d'économies d'échelle :

- la 40^e heure de pointe, qui sert normalement au dimensionnement des aérogares, est reliée au trafic annuel selon une loi linéaire comportant une ordonnée à l'origine :

$$40^{\text{e}} \text{ heure de pointe} = 400 - 3,5 \text{ trafic annuel en millions}$$

De ce fait, l'importance du trafic est un facteur de productivité des aéroports, conduisant à des rendements d'échelle globaux.

- Il existe des effets de productivité concernant les aérogares : jusqu'à environ 20 millions de passagers, à taux d'occupation équivalents, les coûts totaux n'augmentent pas aussi vite que le trafic, et, à taille d'aérogare donnée, les coûts d'exploitation augmentent moins vite que le trafic. Il existe aussi de forts effets de productivité concernant les pistes et voies de circulation avion, qui présentent des coûts fixes importants.

Ces résultats qualitatifs corroborent l'étude précédente pour aboutir à l'idée que les infrastructures aéroporaires présentent des effets d'échelle. Des études statistiques menées à partir des données françaises confirment ce point (Quinet, 1992) et font apparaître des économies d'échelle qui diminuent avec la taille de l'aéroport et qui sont de l'ordre de 1,1.

4.2.2. Infrastructures de navigation aérienne

Ce domaine complexe a donné lieu à des analyses comparatives de productivité qui montrent que le coût du contrôle aérien dépend de nombreux facteurs tels que le volume de trafic, mais aussi la direction du trafic (flux parallèles ou croisés), la vitesse des avions (des vitesses hétérogènes rendent plus coûteux le contrôle), la répartition horaire du trafic ; enfin, les répercussions géographiques sont importantes : une congestion aux Baléares peut occasionner un encombrement à Amsterdam. En l'absence d'études économétriques, on émet en général les considérations qualitatives les suivantes :

- À équipement donné, le coût marginal est d'abord très faible, puis croissant (on ne peut pas dépasser un nombre d'appareils par contrôleur, on ne peut pas réduire la taille des secteurs, de toute façon bornée par la nécessité d'un espace minimum de manœuvre, sans augmenter les coûts de transmission de l'information d'un secteur à l'autre).

- Quand on fait varier les équipements (radar monopulse, informatisation accrue), les coûts marginaux à court terme et à long terme baissent, et le seuil de saturation est repoussé. Il existe une incertitude technologique, liée à la difficulté d'informatisation complète qui n'est pas actuellement possible.

Ces considérations incitent à favoriser l'hypothèse de rendements croissants amenant à un coût marginal inférieur au coût moyen. Quelle est l'importance des rendements d'échelle ? Selon les avis, les économies d'échelle se situent entre 1 et 2, c'est-à-dire que le coût marginal serait compris entre la moitié du coût moyen et le coût moyen.

5. Les coûts d'environnement

Les transports ont de nombreuses conséquences sur l'environnement.

Citons :

- le bruit,
- la pollution de l'air, locale et globale (effet de serre),
- l'effet de coupure,
- la pollution des eaux et des sols,
- les conséquences esthétiques.

Seuls les trois premiers points, et un peu le 4^e, ont fait l'objet d'évaluations monétaires. Celles-ci sont en effet rendues difficiles pour la raison que ces effets sur l'environnement sont des effets externes, qui ne s'exercent pas à travers le jeu du marché ; on ne dispose donc pas de prix pour les évaluer, il faut reconstituer ces prix.

5.1. Les méthodes d'évaluation

Différentes méthodes générales existent pour évaluer les biens d'environnement :

- La méthode des marchés de substitution

Dans cette catégorie, on peut distinguer :

- Le coût des voyages ; ainsi du coût (en argent et en temps) que les usagers exposent pour se rendre dans un parc de loisirs, on peut en déduire la valeur qu'ils attribuent à ce parc ;

- La méthode dite des « prix hédoniques » : le prix de certains biens marchands (par exemple les logements) est influencé par les caractéristiques de qualité comme l'exposition au bruit et à la pollution. En reliant les variations de prix aux variations de qualité de l'environnement, on peut en déduire la valeur que les individus attachent à la qualité de l'environnement ;

1. Elles sont présentées dans Pearce et Markandya (1989) ou dans Quinet (1994), d'où est tirée la présentation qui suit.

— Lorsqu'il existe un marché de biens à polluer le producteur du marché fournit, si certaines hypothèses de son comportement de ce marché sont remplies, le consentement marginal à payer pour le niveau de pollution permis par le nombre de droits émis.

— L'évaluation des dépenses de prévention. Par une analyse statistique du montant des dépenses que les agents consentent à se protéger d'un certain niveau de nuisance, on peut en déduire une disposition à payer pour éviter la nuisance.

Ces méthodes présentent plusieurs difficultés de mise en œuvre. Ainsi les relations hédoniques trouvées ne sont pas de véritables lois de demande : les occupants des logements les plus exposés les ont choisis en partie parce qu'ils sont moins sensibles au bruit. La valeur des logements est moins diminuée qu'elle ne le serait si la sensibilité au bruit était la même quelle que soit l'exposition. Par ailleurs, comment séparer l'effet des différentes variables, souvent corrélées entre elles.

De même les agents économiques ne sont souvent pas conscients des dommages occasionnés par la nuisance en cause.

• Les évaluations contingentes

Elles consistent à demander aux agents ce qu'ils seraient prêts à payer pour éviter la nuisance, ou prêts à recevoir pour la supporter. Les difficultés de mise en œuvre sont nombreuses. Ainsi :

— Pour obtenir des réponses fiables, il faut apporter beaucoup de soin à l'élaboration des questions, et suivre un code de procédure très complexe.

— La « dissonance cognitive » explique que la disponibilité à payer pour une amélioration soit plus faible que le consentement à recevoir pour une détérioration.

— Enfin, comme pour les méthodes précédentes, les résultats peuvent être faussés par une mauvaise connaissance des dommages.

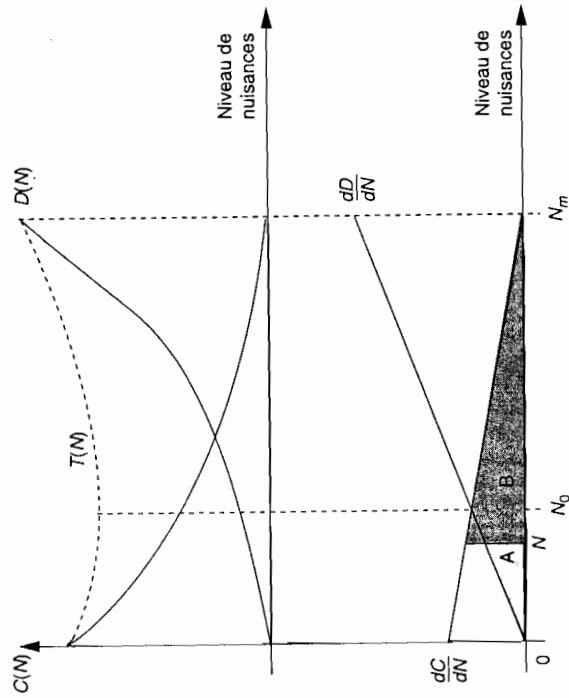
• Les méthodes indirectes : coût d'évitement, coût des dommages

Elles comportent deux phases. La première, de nature technique, vise à apprécier les conséquences de la pollution en termes physiques ; par exemple, pour la pollution de l'air, la fréquence et la gravité des maladies occasionnées, les détériorations d'immeubles ou destructions de végétaux. La seconde consiste à évaluer les dommages correspondants, à travers les évaluations du marché pour les biens détruits lorsqu'il y en a, à travers le coût de remise en état (soins aux malades, ravalement des bâtiments), ou par des évaluations de nature plus subjective (valeur de la vie humaine). Cette méthode évite les défauts d'information des agents. Mais l'évaluation monétaire des dommages est sujette à discussion. En particulier, l'évaluation par le coût de remise en état surévalue le consentement à payer lorsque cette remise en état n'est pas effectuée. Enfin, la remise en état suppose la définition d'un niveau 0 de pollution, définition souvent arbitraire (cas du bruit).

Comme on le voit à leurs définitions, les méthodes de marchés de substitution et les évaluations contingentes déterminent la disponibilité à payer de ceux qui supportent la nuisance ; elles concernent la demande. Les coûts d'évitement ou de dommages concernent l'offre. Si les décisions étaient optimales, toutes ces méthodes aboutiraient au même résultat et on aurait la double égalité que le marché réalise automatiquement dans le cas des biens marchands (voir figure 5-5) :

$$\text{prix} = \text{coût marginal} = \text{disponibilité marginale à payer}$$

FIGURE 5-5
L'optimum en matière d'environnement



Comme on n'est pas, sauf hasard, à l'optimum, ces méthodes ne fournissent des valeurs similaires. Laquelle retenir ? La réponse dépend de la nature du problème à résoudre (Quinet, 1994 ; Mauch et Rottengather, 1996).

Par ailleurs, les coûts comme les disponibilités à payer concernent en général plusieurs acteurs : ainsi pour le bruit, la courbe des disponibilités à payer doit être constituée de la somme des disponibilités à payer de tous les agents touchés par la source de bruit ; de même, la courbe des coûts de protection pourra comporter à la fois le coût d'un écran antibruit, payé par les pouvoirs publics, et le coût de double-vitrages, à charge des occupants des logements.

Lorsque l'effet d'environnement considéré est durable (cas de l'effet de serre), son évaluation conduit à considérer les effets sur les générations futures, ce qui pose des problèmes : comment estimer la valeur que lui attribueront ces générations futures ; et comment ramener cette évaluation à l'instant présent, c'est-à-dire quel taux d'actualisation choisir.

On définit souvent aussi trois catégories de valeurs pour les biens d'environnement, dont la somme constitue la valeur totale du bien :

- La valeur d'usage, qui correspond à la consommation actuelle du bien ; j'apprécie ce parc de loisir parce que je m'y rend.
- La valeur d'option, qui correspond à une possibilité de consommation future du bien ; j'apprécie ce parc parce que peut-être dans 10 ans j'aurai besoin de je.
- La valeur d'existence ; je ne me rendrai pas dans ce parc mais j'apprécie son existence pour les autres ou pour les générations futures.

Ces trois valeurs existent aussi pour les biens matériels, mais on considère que la valeur d'option et la valeur d'existence prennent une importance plus grande pour les biens d'environnement, notamment pour les biens naturels non renouvelables. On n'est pas tout à fait dans ce cas pour les biens de pollution locale. Néanmoins les méthodes hédoniques ou de bien-être nous conduisent que les valeurs d'usage ; pour les méthodes contingentes, on aborde de la manière dont sont posées les questions.

5.2. Les résultats d'évaluation

Les considérations qui précèdent expliquent partiellement les résultats des études de valorisation soient assez dispersées. Elles nécessitent successivement pour chaque type d'atteinte à l'environnement :

5.2.1. Bruit

Les effets du bruit des transports sont difficiles à mesurer. L'unité la plus fréquemment utilisée, le *Leq* est une intégration sur une durée déterminée, du logarithme de la puissance sonore dans laquelle les fréquences sont pondérées en fonction de leur gêne pour l'homme. Il est mesuré en decibel A (dba). Selon certains auteurs, il y a une différence de 5 dbA entre le bruit des chemins de fer et celui de la circulation routière procurant la même gêne. Les méthodes usuelles pour évaluer le bruit sont du type hédonique, fondées sur les pertes de valeur immobilière dues à l'exposition au bruit ; les études convergent pour conclure qu'une augmentation de un dbA entraîne une diminution de 1 % du prix du logement. Les méthodes utilisant le coût de protection souffrent de ce que celui-ci dépend énormément du niveau de protection jugé souhaitable. Ainsi, l'étude Planco (1990) indique pour l'Allemagne les coûts de protection suivants :

Protection à 55 dbA :	1,7 GDM
à 45 dbA :	10,0 GDM

Notons que l'évaluation du bruit faite dans la même étude en prenant en compte la disposition à payer (résultant d'une enquête de préférences déclarées) aboutit à un coût du bruit de 18,0 GDM, ce qui illustre les différences entre les résultats des diverses méthodes.

Quinet (1994) a effectué un recensement de diverses études européennes portant sur l'évaluation du bruit des transports dans différents pays, exprimé en pourcentage du PNB du pays. Les 23 valeurs recensées s'évaluent entre 0,06 % et 1,5 % avec une moyenne à 0,3 %.

Le Livre Vert de 1992 de l'Union Européenne indique la répartition suivante des responsabilités des modes de transports en matière de bruit pour l'ensemble de l'Europe :

Route :	64 %
Rail :	10 %
Air :	26 %

5.2.2. Pollution locale

La pollution locale due aux transports s'exerce à travers différents composés chimiques émis par les véhicules : les oxydes nitriques, facteur d'apparition de l'ozone dans des conditions complexes, les oxydes sulfuriques, les hydrocarbures imbrûlés, les particules organiques volatiles. Les effets sont triples : ils concernent la santé humaine, les effets sur les plantes et les animaux et les effets sur les constructions bâties. De grandes incertitudes règnent sur certains d'entre eux, notamment les effets cancérigènes à long terme des particules. Les méthodes d'évaluation sont de types « coût des dommages », « coûts de protection » et « disponibilité à payer » (résultant d'enquêtes de préférences déclarées).

Les résultats, recensés par Quinet (1994), exprimés en pourcentage des PNB, s'évaluent entre 0,03 % et 1,2 % avec une moyenne à 0,4 % et 10 évaluations sur les 13 se concentrant entre 0,2 % et 0,6 %.

5.2.3. Pollution globale (effet de serre)

Le plus important des gaz à effet de serre est le CO₂, abondamment émis par la combustion des hydrocarbures. L'augmentation des teneurs en CO₂ dans l'atmosphère, constatée depuis le début de l'ère industrielle, s'est accompagnée d'un réchauffement de l'atmosphère. Les modèles élaborés pour prévoir l'évolution du phénomène fournissent des fourchettes assez larges de 0,1 à 0,5 degrés par période de 10 ans si aucune action pour enrayer le phénomène n'est entreprise. Les conséquences de ce réchauffement sont de natures très diverses : disparition d'espèces, élévation du niveau des mers entraînant à la limite la submersion de certains États comme le Bengla-Desh ; adoucissement de certains climats froids. Les études pour évaluer les dommages qui en résulteraient sont peu nombreuses, et la majorité se concentrent sur les conséquences pour l'Amérique du Nord. Les ordres de grandeur obtenus se situent aux alentours d'une réduction de 1 % à 2 % du PIB américain. D'autres études montrent que les pays d'Asie du Sud ou d'Afrique pourraient subir des diminutions de PIB de l'ordre de 10 %. Le tableau ci-dessous fournit les résultats de quelques études récentes (Tinch, 1996).

1. Le niveau de 45 dbA correspond à une situation de silence diurne.

Quelques résultats d'études d'évaluation du réchauffement planétaire

Étude	Hypothèses fondamentales	Fourchette d'estimation (en % du PIB)	
Nordhaus (1991)	2*CO ₂ - 3 °C	2,5 % - 20 % (les omissions)	
		1,5 % - 10 % (estimation la plus plausible)	
		2,5 % - 20 % (estimation la plus plausible)	
Cline (1992)	2*CO ₂ - 2,5 °C	1,5 % - 20 % (les omissions)	
		1,5 % - 10 % (estimation la plus plausible)	
		2,5 % - 20 % (estimation la plus plausible)	
		2,5 % - 20 % (estimation la plus plausible)	
Tol (1994)	2,5 °C - OCDE	1,6 %	
		2,5 °C - non OCDE	1,7 %
		2,5 °C - monde entier	1,9 %
Fankhauser (1994)	2,5 °C - OCDE	1,5 %	
		2,5 °C - non OCDE	1,6 %
		2,5 °C - monde entier	1,7 %

Source : CEMT, 1998.

On peut, à partir de là, calculer le coût d'émission de la tonne de carbone, évalué par l'intermédiaire du coût marginal des dommages. Les chiffres obtenus s'étagent entre 60 et 40 ÉCU par tonne de carbone. Des études plus nombreuses ont porté sur les taxes qu'il serait nécessaire de mettre en place pour réduire les émissions de CO₂ à un niveau acceptable (le terme « acceptable » n'est pas exempt d'un certain arbitraire...). Ces taxes sont sensiblement plus élevées que les valeurs précédentes ; ainsi, la Commission de Bruxelles a proposé en 1994 une taxe de 70 ÉCU par tonne de carbone. Par ailleurs, différentes études, citées dans Quinet (1994) font apparaître des taxes sur le carbone de 120 à 150 ÉCU pour arriver à des réductions des émissions de 20 à 50 %.

6. L'insécurité

L'insécurité est une caractéristique des transports et plus spécialement des transports routiers : il y avait en 1996 environ 8 600 morts dans les transports (8 400 sur les routes, 129 pour le chemin de fer et 75 par avion). L'évaluation du coût de l'insécurité s'obtient en multipliant le nombre de morts, de blessés et d'accidents matériels par le coût unitaire de ces morts, blessés et accidents matériels.

6.1. L'évaluation de la vie humaine

En ce qui concerne la valeur de la vie humaine, notons d'abord qu'il s'agit d'une évaluation statistique, celle d'une vie humaine non identifiée, et que cette valeur est, pourrait-on dire, une valeur marginale, correspondant à la valorisation d'un risque faible.

Deux méthodes sont couramment utilisées pour évaluer la vie humaine. La méthode dite du capital humain se fonde sur la perte de production actualisée qui résulte du décès ; c'est la méthode la plus ancienne. Elle connaît plusieurs modalités selon qu'on apprécie la production brute ou nette, c'est-à-dire diminuée de la consommation. Sa mise en œuvre se heurte à plusieurs difficultés. La première est technique, et résulte de ce qu'il faut d'abord prévoir les productions futures, portant souvent sur un avenir éloigné, et ensuite qu'il faut les actualiser, et le résultat est sensible au taux d'actualisation choisi. Une autre est d'ordre éthique : par cette méthode les pertes de certaines vies, par exemple celles des retraités, seraient nulles, voir négatives dans la variante « production nette ». Enfin, la méthode ignore les éléments subjectifs de douleur et de peine.

L'autre méthode, celle des disponibilités à payer, peut être mise en œuvre sous deux formes différentes. La première passe par la révélation des préférences telle qu'elle apparaît à travers des choix, comme les décisions de protection des individus (coût de soins réduisant les probabilités de décès), ou les salaires de métiers à risque. La seconde correspond aux préférences déclarées : par des techniques d'interviews, on cherche à connaître ce que chacun est prêt à payer pour réduire la probabilité de décès d'un montant donné. Les interviews peuvent cerner deux valeurs qui n'ont pas de raison d'être égales : ce que je serais prêt à payer pour réduire mon risque de décès, ou ce que je serais prêt à payer pour réduire le risque d'un décès indifférencié dans la collectivité.

Les résultats sont assez dispersés, ceux de la méthode du capital humain étant en général plus faibles. Quinet (1994) cite des valeurs en usage en 1989 dans les pays de l'OCDE allant de 100 000 à 1 600 000 ÉCU. Tinch (1996) fournit le tableau de valeurs suivant, tiré de Mauch et Rotengather (1996).

TABLEAU 5-4b

Résultats de certaines études sur l'évaluation de la vie humaine Méthode du capital humain

ÉCU/1991	Coûts			Valeur de la vie	Total des coûts externes
	administratifs (police, justice, etc.)	Coûts médicaux ¹	Pertes de production nettes		
Allemagne	20 114	4 374	1 113 333	1 233 190	
Autriche	19 417	4 771	1 064 788	1 180 106	
Belgique	18 277	4 484	1 002 236	1 109 640	
Danemark	22 884	5 881	1 254 919	1 394 218	
Espagne	16 945	4 111	1 223 228	1 009 288	
Finlande	24 831	6 124	1 361 681	1 465 268	
France	18 425	4 111	1 103 385	1 103 977	
Grèce	14 118	3 423	774 179	803 182	
Irlande	17 119	4 111	1 000 885	1 000 885	
Italie	18 819	4 883	1 124 145	1 124 145	
Luxembourg	18 463	4 111	1 346 475	1 346 475	
Norvège	23 458	5 671	1 433 036	1 433 036	
Pays-Bas	18 618	4 883	1 135 592	1 135 592	
Portugal	12 144	2 946	711 187	711 187	
Royaume-Uni	17 878	4 111	1 048 381	1 048 381	
Suède	26 277	6 362	1 586 434	1 586 434	
Suisse	24 831	6 124	1 661 209	1 661 209	
Moyenne	18 390	4 461	1 106 283	1 106 283	
Pourcentage	1,66 %	0,40 %	91,16 %		

1. Aux coûts médicaux ont été ajoutés les coûts de remplacement de la personne décédée sur son lieu de travail, qui sont à la charge des entreprises.

Source : Tinch (1996) et Mauch et Rotengartner (1996).

On peut se demander si l'existence de sentiments altruistes ne vient pas vicier l'évaluation fournie par la méthode de disponibilité à payer. Un résultat fourni par Jones-Lee (1990) est que l'évaluation de la vie humaine obtenue de cette manière n'a pas à être modifiée pour tenir compte d'un effet d'altruisme, correspondant à une disposition à payer pour sauver la vie d'autrui, sous réserve que l'altruisme porte autant sur la sécurité que sur les autres éléments du bien-être d'autrui (absence de paternalisme).

TABLEAU 5-4b

Résultats de certaines études sur l'évaluation de la vie humaine Méthodes des préférences déclarées ou révélées

Étude	Estimation	Observations
<i>Cité dans Jones-Lee (1990)</i>		
Jones <i>et al.</i> (1985)	3,2 millions d'écus	Enquête réalisée au RU auprès de 1 000 personnes
Persson (1989)	2,6 millions d'écus	Enquête suédoise auprès de 500 personnes
Maier <i>et al.</i> (1989)	3,0 millions d'écus	Échantillon autrichien de 100 personnes
Jones-Lee (1989)	2,9 millions d'écus ¹	Préférence révélée moyenne : 7 pour le RU, 13 pour les EU et 1 pour l'Australie
<i>Cité dans Viscusi (1993) : \$1990</i>		
Dardis (1980)	0,6 million de dollars	Achats de détecteur de fumée
Garbacz (1989)	2,0 millions de dollars	Achats de détecteur de fumée
Marin & Psacharopoulos (1982)	2,8 millions de dollars ¹	Étude salaire-risque (RU)
Marché du travail (données antérieures à 1980)	4,6 millions de dollars (moyennes) ¹	Fourchette de 0,6 à 10,3 millions de dollars
Marché du travail (données postérieures à 1980)	7,5 millions de dollars (moyenne) ¹	Fourchette de 1,6 à 16,2 millions de dollars
Calthrop (1996)	2,4 millions d'écus	Écart-typé : 0,5 million

1. Certaines études sur les salaires utilisent les traitements bruts plutôt que les traitements nets et doivent par conséquent être ajustées à la baisse pour tenir compte des taux d'imposition marginaux.

On peut se demander aussi si la valeur de la vie humaine (VVH) doit être la même dans les différents modes de transport.

C'est ce dernier aspect qui va être maintenant développé¹, en raisonnant par rapport à ce repère qu'est la VVH dans le domaine routier et en recherchant les motifs de différence entre les deux valeurs.

La première et la plus simple tient aux différences de caractéristiques socio-économiques des populations d'usagers des différents modes de transport. Parmi ces caractéristiques, le revenu est probablement la plus importante ; dans la mesure généralement admise où la valeur de la vie est positivement corrélée avec le revenu, il apparaît une différence entre la VVH routière et la VVH des transports collectifs. Plus précisément, il ressortirait de ce facteur que la VVH moyenne

1. À partir de Quinet et Galland (1995).

des transports collectifs urbains devrait être inférieure à celle de l'automobiliste, et celle des transports collectifs interurbains supérieure à celle de l'usager du train 2^e classe, d'avantage pour ceux du train 1^{er} classe et de l'avion.

On peut penser que cet effet est dans l'ensemble adossé par le filtrage du choix du mode en fonction de la VVH. Il est en effet évident que les voyageurs dont la disponibilité à payer pour la sécurité est élevée choisissent le mode le plus sûr, de même que les usagers dont la valeur du temps est la plus forte choisissent le mode le plus rapide. On peut toutefois se demander si cet effet est significatif. En tout cas, aucun modèle de trafic visant à reproduire les comportements ne l'envisage. Cela est peut-être dû à ce que les usagers ont une plus forte disponibilité à payer pour la sécurité sont aussi en général ceux qui ont la plus forte valeur du temps, les deux valeurs étant très corrélées. Et ce fait est aussi la plus forte valeur de relation entre le niveau de sécurité des modes et leur vitesse. Cela peut être du aussi à ce que cet effet d'orientation vers les modes les plus sûrs des usagers ayant la VVH la plus élevée peut se trouver amoindri ou même annulé par l'imparfaite connaissance des conditions réelles de sécurité comparées. L'expérience montre (Viscusi 1993) que les individus sous-estiment les dangers dans les conditions de sécurité des différents modes. Les risques perçus sont supérieurs aux risques réels dans les transports collectifs, et inférieurs dans le transport automobile.

Une autre raison d'écart tient à la contribution entre les efforts de réduction du risque consenti par les automobilistes dans leur façon de conduire et leur disponibilité à payer. Selon un argument présenté par S.C. Kolm (1970), une amélioration de l'équipement routier se traduit par une réduction du risque (par exemple par une conduite plus sage) plus forte de la part de ceux qui ont une VVH élevée que de la part des autres qui convertissent en gain de temps le gain de sécurité qui est proposé. Il en résulte que les décisions publiques en matière de sécurité qui est route doivent être prises avec une valeur moyenne de la vie humaine inférieure à la valeur moyenne des usagers.

Jones-Lee (1994) a introduit comme raison d'écart entre la VVH routière et la VVH des autres modes ce qu'il appelle l'effet de contexte : les conditions d'un accident influent sur le prix qu'on est prêt à payer pour l'éviter. Il est sûr que certaines formes de mort ou de blessure inspirent davantage de répulsions que d'autres, à la fois par les souffrances qui les accompagnent et par l'impression de cruauté ou de violence qui s'en dégage. Une bonne partie des progrès de la médecine vise à atténuer les souffrances et les peines de justice ont été humanisées au cours des âges. Jones-Lee a ainsi mesuré par des enquêtes de préférences déclarées le surcroît de valeur de la vie humaine résultant d'un accident en souterrain par rapport à un accident à l'air libre, et trouve une différence de 50 %.

Le même auteur cite là une autre cause possible d'écart entre la VVH dans les transports collectifs et VVH dans les transports individuels, l'effet de masse : dix morts d'un seul coup sont ressentis plus vivement que dix fois un mort isolé. Les différences d'attention que les médias accordent aux accidents traduisent bien cet effet, ainsi d'ailleurs que certaines pratiques usuelles de l'analyse de risque dans les installations industrielles, mais ceci ne suffit pas à en justifier la légitimité dans les études de surplus collectif. Une justification de cet effet de masse peut être trouvée dans un comportement d'aversion au risque dans l'altruisme (voir Galand et Quinet 1995). Quand à son ordre de grandeur, la seule expérience permettant de l'apprécier est l'enquête menée par Jones-Lee (1994) en termes de préférences déclarées dans le cadre de son étude relative au métro londonien, et dont les résultats tendent à montrer que cet effet de masse est faible, de l'ordre de 10 %.

Les effets qui viennent d'être cités vont tous dans le sens d'une VVH dans les transports collectifs supérieurs à celle des transports individuels. Jones Lee et Loomes (1994) proposent un rapport de l'ordre de 1,5 entre les deux. Quinet et Galland (1995) constatent que le rapport implicite réel qui résulte des décisions de sécurité ferroviaires est considérablement plus élevé. Ils proposent deux raisons pour expliquer cette différence.

D'abord, comme dans les transports collectifs les usagers n'ont pas de responsabilité dans un éventuel accident, il conviendrait de les en protéger davantage. Cet argument peut paraître raisonnable, mais encore faut-il lui trouver une justification dans le cadre de décisions rationnelles sur un plan collectif ; cette justification peut être trouvée dans l'imparfaite information que, dans un transport collectif, le transporteur dispose sur la VVH du transport : il n'en connaît pas la valeur exacte mais simplement la distribution *a priori* sur la population. On peut voir à partir de là que, si l'obtention de la sécurité obéit à une loi des rendements décroissants, le transporteur sera amené à fournir une sécurité moyenne supérieure à celle qui résulterait des décisions prises en situation de connaissance parfaite ou, ce qui revient au même, à calculer la rentabilité de ses actions à partir d'une valeur de la vie humaine supérieure à la moyenne des VVH de la population.

Ensuite, les mécanismes de décision portent à amplifier l'écart : les décideurs en matière de sécurité dans les transports publics supportent davantage les conséquences d'une mauvaise sécurité (responsabilité médiatique, judiciaire, culture d'entreprise) que les coûts de son amélioration (augmentation des subventions).

6.2. Les coûts d'insécurité

Pour calculer les coûts d'insécurité, il faut, en plus du prix de la vie humaine, tenir compte des coûts des blessés et des dégâts matériels.

L'évaluation des accidents matériels est d'habitude fondée sur le coût monétaire des dommages. En ce qui concerne les blessés, l'estimation inclut les coûts directs (soins médicaux, transports des blessés etc.), les coûts indirects (pertes de production) et une estimation relativement subjective, parfois fondée sur les indemnisations judiciaires, du « *pretium doloris* ».

Le Net, dont les estimations sont retenues dans le rapport Boiteux (1994), évalue ainsi les coûts des blessés :

Blessé grave : 370 000 F ; Blessé léger : 79 000 F

Ces chiffres à mettre en comparaison avec la valeur de la vie trouvée dans la même étude, soit : 3 600 000 F.

En appliquant ces valeurs unitaires aux nombres de morts et blessés, et en ajoutant les dégâts matériels, on trouve un coût total d'insécurité, à l'échelle de la France, de 51,4 milliards de F en 1990 (Commission des Comptes de Transport de la Nation, juin 1997).

7. Le coût des opérateurs

7.1. Les outils d'analyse

Les opérateurs de transport constituent une catégorie très large qui couvre à la fois les compagnies aériennes, les entreprises de transport rou-

tier, les exploitants des transports publics urbains, les compagnies de chemin de fer. Ces firmes sont multiproduits. Dans la mesure où cela ne présente pas le même coût de transport du fait de des passagers : ni de transporter vers une destination ou une autre à une période de pointe ou à une période creuse. Les fonctions de coût devraient donc, en toute rigueur, dépendre d'un très grand nombre d'arguments. À tel point qu'elles seraient statistiquement impossibles à déterminer. On remplace donc cette multiplicité d'arguments par quelques résurgences définissant de façon plus maniable les paramètres influant sur les coûts. Dans ces fonctions de coût, qu'on appelle « *hédoniques* » intervennent par exemple :

- le trafic total exprimé en voyageurs, kilomètres Q_V
- en tonnes x kilomètres Q_M
- le taux de chargement moyen T
- la distance moyenne de transport L
- la longueur totale du réseau N

et la fonction s'écrira ainsi :

$$C = C(Q_V, Q_M, T, N, L)$$

On définira à partir de là un certain nombre d'indicateurs dont l'intérêt tient à ce qu'ils permettent de juger de la présence ou non de monopole naturel :

- Les économies d'échelle, calculées à partir d'une variation homothétique de tous les arguments (Jara-Diaz et Cortes (1996) présentent des arguments pour ne pas inclure toutes les caractéristiques de réseau dans l'expression des économies d'échelle) :

$$EE = \frac{C}{Q_V \frac{\partial C}{\partial Q_V} + Q_M \frac{\partial C}{\partial Q_M} + N \frac{\partial C}{\partial N} + T \frac{\partial C}{\partial T} + L \frac{\partial C}{\partial L}}$$

$$= \frac{1}{\varepsilon_{Q_V}^C + \varepsilon_{Q_M}^C + \varepsilon_T^C + \varepsilon_N^C + \varepsilon_L^C}$$

- Les économies de densité, calculées à réseau constant :

$$ED = \frac{C}{Q_V \frac{\partial C}{\partial Q_V} + Q_M \frac{\partial C}{\partial Q_M}} = \frac{1}{\varepsilon_{Q_M}^C + \varepsilon_{Q_V}^C}$$

Dans ces formules, les ε sont les élasticités du coût par rapport aux variables indicées.

1. À noter que la variable la mieux liée aux coûts est en général le siège x kilomètre offert (sko). On pourrait évaluer d'abord une fonction de coût dépendant des sko, puis la relation que les opérateurs font entre sko et demande. Cette procédure est rarement utilisée.

- Les économies d'envergure peuvent être appréciées à travers l'indice d'économie d'envergure :

$$\frac{C(Q_V, Q_M) - C(0, Q_M) - C(Q_V, 0)}{C(Q_V, Q_M)}$$

On peut encore les apprécier par la concavité de la courbe isocoût, qui vérifie :

$$C(Q_V, Q_M) = \text{Constante}$$

Les formes de fonction les plus utilisées sont les fonctions translog, développement limité d'ordre 2 du logarithme de la fonction de coût. Un cas particulier de la fonction translog est la fonction de Cobb-Douglas¹, développement limité d'ordre 1 :

$$C = Q_V^a Q_M^b N^c L^d$$

La supériorité de la fonction translog générale par rapport à la fonction de Cobb Douglas est que l'élasticité de substitution entre facteurs² peut prendre n'importe quelle valeur, alors qu'elle est égale à l'unité pour la fonction de Cobb-Douglas, et quelconque mais constante dans les fonctions CES, du type :

$$y = \left(\sum_j a_j x_j^\rho \right)^{\frac{1}{\rho}}$$

7.2. Les résultats des évaluations

Une recension des évaluations des fonctions de production et de coût des opérateurs de transport a été faite par Winston en 1985. À l'époque, l'analyse portait presque exclusivement sur les économies d'échelle, et les résultats rassemblés font apparaître des rendements d'échelles, élevés dans le chemin de fer (élasticités des coûts aux produits inférieurs à l'unité), et à peu près constants dans les deux autres modes.

Les études effectuées depuis une dizaine d'années ont généralisé l'utilisation de la fonction Translog ; elles ont aussi plus systématiquement pris en compte la dimension multiproduits de l'activité des entreprises de transport. Les résultats auxquels elles aboutissent nuancent les conclusions qui pouvaient être tirées en 1985.

1. Comme indiqué à la section 4.1, la fonction de coût doit aussi contenir comme arguments les prix des entrants qui, pour simplifier l'écriture, n'ont pas été repris dans la formule précédente.

2. L'élasticité de substitution entre deux facteurs x et y d'une fonction de production $q = f(x, y)$ est définie par

$$e = (dTMS/TMS) / (d(y/x)/(y/x))$$

formule dans laquelle le taux marginal de substitution est défini par :

$$TMS = (\partial f / \partial x) / (\partial f / \partial y)$$

TABLEAU 5-5
Élasticité du coût des transports à la production

Étude	Forme fonctionnelle	Spécification de la production	Élasticité du coût au point moyen de l'échantillon
Keeler (1974)	Non linéaire	Multiple	0,57
Harris (1977)	Linéaire	Simple	0,64
Friedlaender et Spady (1981)	Translog	Multiple	0,895
Caves, Christensen et Swanson (1981b)	Translog	Multiple	0,605-0,716
Jara-Diaz et Winston (1981)	Quadratique	Multiple	0,352-0,787
B. Transport routier de marchandises			
Koenker (1977)	Log-linéaire	Simple	> 1
Spady et Friedlaender (1978)	Translog	Simple	> 1
Spady et Friedlaender (1978)	Translog	Simple	< 1
Harmatuck (1981)	Translog	Multiple	= 1
Wang et Friedlaender (1981)	Translog	Multiple	= 1
C. Air			
Eads, Nerlove, Raduchel (1969)	Non linéaire	Simple	≥ 1
Keeler (1972)	Linéaire	Simple	= 1
Douglas et Miller (1974)	Semi-log-linéaire	Simple	= 1
Caves, Christensen et Trettheway (1983)	Translog	Simple	= 1

Source : Winston, 1985.

D'autres études développent les effets de réseaux. T. Walker (1992) étudie, par une méthode de simulation, les effets de la taille et du nombre d'entreprises pour servir un marché de structure géographique donnée. Il le fait par une méthode de simulation fondée sur un algorithme de minimisation des coûts et aboutit à l'existence d'économies d'échelle. Callan et Thomas (1992) étudient, par la détermination d'une fonction de coût translog, les effets de la composition du chargement de ses dimensions et des dimensions du réseau qui influencent les coûts. Ils trouvent des économies de densité, mais pas d'économies d'échelle.

Les études précédemment citées supposent implicitement, par l'usage de la régression linéaire, que les variables explicatives sont exogènes, non déterminées par les entreprises, une condition nécessaire pour que les coefficients obtenus par la méthode des moindres carrés ne soient pas biaisés. Cette hypothèse pouvait être admise dans un environnement réglementé tel qu'il existait avant les années 80. Il n'en est plus ainsi maintenant ; la structure de la production (distance moyenne de transport, taille du réseau, chargement moyen...) résulte d'une décision de l'entreprise, ce qui conduit à considérer, non pas une seule équation, mais un modèle à équations simultanées. Les résultats sont sensiblement différents et font parfois apparaître des économies d'échelles là où l'analyse classique aboutissait à des rendements décroissants.

Xu, Windle, Grimm et Coasi (1994) ont effectué en ce sens une analyse de la fonction de coût du transport routier par chargement incomplet (LTL) dans les années post-dérégulation (1988 à 90) aux USA. Leur modèle fait intervenir, outre les prix des facteurs :

- C le coût total
- TM le produit en Tonnes & Miles
- D la distance moyenne de transport
- T le chargement moyen

Les équations du modèle sont, sous la spécification Translog :

$$\begin{aligned} C &= f(TM, D, T) \\ D &= g(TM) \\ T &= h(TM) \end{aligned} \quad (1)$$

L'analyse économétrique montre d'abord qu'il y a une corrélation significative et positive entre D et TM ainsi qu'entre T et TM . Ceci montre que les grandes entreprises se spécialisent dans les longues distances et les taux de chargement forts.

Il en résulte d'abord que les élasticités de la fonction de coût sont estimées avec biais si l'on ne tient pas compte de ces liaisons. On le voit à la valeur des élasticités qui diffèrent sensiblement selon que les coefficients de l'équation (1) sont estimés isolément ou à travers le modèle constitué par cette équation et les deux autres, comme le montre le tableau ci-dessous où apparaissent les coefficients du premier degré de la fonction translog, ceux qui permettent de déterminer les élasticités au point moyen de l'échantillon :

Coefficients Variables	Équation simple	Modèle
TM	1,08 (0,038)	1,13 (0,035)
D	- 0,07 (0,074)	- 0,55 (0,074)
T	- 1,40 (0,089)	- 0,93 (0,087)

Le secteur le plus étudié est celui du transport routier de marchandises, probablement en raison de l'abondance des données qui facilite l'exploration économique.

Harmatuck (1991), travaillant sur la base d'une fonction bédonique Translog séparant les chargements complets (Truck Load) et les autres (Less than Truck Load) et tenant compte des caractéristiques de réseau, aboutit au résultat que, dans l'ensemble, les économies d'échelle sont voisines de 1 et plutôt supérieures à l'unité, mais que des écarts importants existent par rapport à cette valeur. L'auteur teste aussi l'existence d'économies d'envergure : il également les situations existent selon le transporteur, mais il semble qu'en général les économies d'envergure existent notamment pour les gros transporteurs, résultat qui rejoint celui trouvé par Harmatuck (1981) et par Chian, Judy et Friedlaender (1984). Beuthe et Sayez (1994) analysent la fonction de coût des transports routiers à longue distance en France, en distinguant 3 natures de transports : les camions citernes, les transports réfrigérés et les autres. Les auteurs trouvent des résultats similaires à ceux de Harmatuck (1991) : une diversité de situation, à l'intérieur desquelles prévalent de légères économies d'échelle et d'envergures.

L'analyse par régression simple montre une faible influence de la distance de transport D . Le modèle fait apparaître un effet beaucoup plus grand et significatif de cette distance de transport. Or la distance de transport et le chargement moyen T sont bien corrélés avec le trafic total en tonne-kilomètre $T.M$: les firmes de grande taille s'implantent sur les marchés de transport à longue distance, sur lequel elles peuvent compenser les légères déséconomies d'échelle par un gain sur la distance de transport, et par un gain de chargement et avoir ainsi un avantage de coût important. Cette capacité des grandes firmes à accroître leur distance de transport peut provenir de plusieurs facteurs tels qu'un meilleur système d'information ou une possibilité de services plus diversifiés dus à une meilleure couverture géographique, ou une plus grande capacité financière.

Cette différence dans l'évaluation des coefficients se traduit dans l'évaluation des économies d'échelle. Celles-ci peuvent alors être appréciées :

- soit, de manière classique, à travers l'équation simple faisant intervenir à la fois le trafic total et les caractéristiques du réseau :

$$EE = \frac{1}{\epsilon_{T.M}^C}$$

- soit, pour tenir compte des corrélations entre variables du réseau, et en utilisant les élasticités fournies par le modèle à 3 équations, par :

$$EE = \frac{1}{\epsilon_{T.M}^C + \epsilon_{T.M}^D + \epsilon_{T.M}^F}$$

La première formulation indique des déséconomies d'échelle ($EE = 0,92$) alors que la deuxième conduit à des économies d'échelle plus élevées ($EE = 1,25$), le tout au point moyen de l'échantillon.

Au total, dans le transport routier de marchandises, il y a quelques économies d'échelle, et surtout des rendements croissants sur d'autres paramètres tels que la distance de transport ou le chargement moyen, rendements croissants qui sont principalement exploités par les firmes de grande taille.

Ces considérations s'appliquent aussi aux autres modes de Transports. Ainsi, les analyses usuelles du transport aérien font apparaître des rendements de densité élevés et des rendements d'échelle voisins de l'unité. C'est ce que trouvent Kirby (1986) pour les compagnies intérieures Australiennes, Oum et Zhang (1991), Gilen, Oum et Tretheway (1990) pour les compagnies canadiennes, Cave, Christensen et Tretheway (1984) pour les compagnies américaines. Les mêmes auteurs mettent en évidence l'importance de l'aspect multiproduit de l'activité aérienne, et notamment de variables telles que le nombre de points desservis, la longueur des étapes, la densité du trafic.

Pavaux (1994) s'appuie d'une part sur des analyses qui montrent que les étapes longues, ou une densité de trafic élevée, facteurs d'abaissement des coûts, se rencontrent surtout dans les grandes compagnies, leur conférant un avantage que l'analyse élémentaire de la fonction de coût ne fait pas apparaître, puisque dans cette fonction, l'élasticité du coût au trafic total est voisine de l'unité.

Oum et Zhang (1997), reprennent les méthodologies de Xu *et al.* (1994), et les appliquent aux données de Caves *et al.* (1984). En tenant compte de ce que les distantes de transport sont corrélées avec la taille de l'entreprise, ils trouvent des rendements d'échelle supérieurs à ceux de Caves *et al.* Cet effet est renforcé par les résultats d'exercices de simulation qui, comme pour Walker (1992) en ce qui concerne la route, montrent que, toutes choses égales par ailleurs, de légères économies d'échelle apparaissent.

En ce qui concerne les autobus, Gagnepain (1998) fait apparaître des économies d'échelle en France.

Berechman (1983) met en évidence de fortes économies d'échelle dans l'industrie des autobus israéliens, en contraste avec la plupart des études antérieures. Tauchen, Fravel et Gilbert (1983) trouvent aussi pour les autocars des économies d'échelle élevées pour les petites compagnies, alors que pour les grandes compagnies, elles sont pratiquement égales à l'unité. Ils trouvent aussi des économies d'envergure entre services réguliers, scolaires et charters.

L'analyse des fonctions de coût du chemin de fer présente deux caractéristiques : d'abord cette fonction de coût englobe à la fois les infrastructures et l'exploitation, contrairement à celles qui ont été analysées jusqu'ici, et qui ne portaient que sur l'exploitation. Ensuite, dans les chemins de fer plus que pour les autres modes, l'adéquation du niveau de capital au niveau de production n'est pas assurée et constitue en soi un sujet d'analyse.

Lorsque l'on passe sur ce deuxième aspect, on trouve des rendements de densité élevés, comme c'est le cas de Geehan (1993) pour l'Irlande, ou de Preston (1994) pour les chemins de fer de 14 pays européens. Quant aux rendements d'échelle, ils dépendent de la taille du réseau d'après Preston (1994) ; supérieurs à l'unité pour les « petits » réseaux européens (Suisse, Portugal, Belgique, Hollande...), ils sont inférieurs à l'unité pour les « grands » réseaux (Grande Bretagne, France, Allemagne...).

Friedlander *et al.* (1993) étudient une fonction de coût variable prise sous forme translog, et dépendant (outre les prix des facteurs de production) des caractéristiques du produit (tonne x routes, % de produits agricoles et de produits miniers transportés), du réseau, de la longueur de transport moyenne, et du niveau d'équipement. À partir de la connaissance du coût variable, on peut déterminer une fonction de coût à long terme optimisée, et comparer le capital optimal au capital réel. L'analyse porte sur une période 12 ans et concerne 27 compagnies ferroviaires américaines. Elle montre que la quasi-totalité de ces compagnies présente un excès de capital. Par ailleurs, les économies de densité (réseau constant) sont élevées et les économies d'échelle (à réseau variable) sont plus faibles, mais en général supérieures à l'unité.

Au total les analyses économétriques les plus récentes nuancent et relativisent les résultats d'ensemble établis en 1985 par Winston : les économies d'échelle et d'envergure dans les chemins de fer sont moindres qu'on ne le pensait il y a quelques décennies ; si dans le transport routier les rendements constants sont plus fréquents, on rencontre aussi des situations d'économies d'échelle, le transport aérien étant dans une situation intermédiaire. Ces résultats dépendent de la structures géographique des réseaux desservis, et du caractère plus ou moins adapté des équipements. Par ailleurs les entreprises choisissent, en général en fonction de leur taille, des créneaux de marché permettant de compenser les effets négatifs de certaines déséconomies auxquelles elles sont confrontées ; les grandes entreprises compensent les diminutions d'économies d'échelle dues à leur taille par des économies de densité et la concentration de leurs activités sur de longues dessertes.

8. Le coût total des transports : résultats numériques

On a, au cours des sections précédentes, indiqué la forme des fonctions de coût correspondant à chaque poste de dépense, et les paramètres intervenant dans ces fonctions de coût. On voudrait ici à titre de synthèse, présenter un certain nombre d'évaluations du coût total des transports, en rassemblant les résultats de quelques études récemment publiées.

Ces études sont hétérogènes dans leur présentation et dans la définition des postes de coûts qu'elles retiennent. On peut toutefois faire à leur sujet quelques remarques générales :

- Quant aux postes de coûts retenus. L'attention s'est focalisée sur les coûts collectifs (coûts externes et coûts de la puissance publique). Les coûts monétaires pour l'utilisateur ne sont souvent pas repris, et les coûts du temps passé sont traités de façon très disparate (coût marginal, moyen, supplément par rapport à des conditions de circulation usagées normales).
- Quant à l'arbitraire des décompositions. Plusieurs postes de coûts, notamment les coûts d'infrastructures et les coûts d'opérateurs, présentent des coûts fixes pour lesquels il existe de nombreuses répartitions entre les catégories de trafic, et aucune n'échappe à l'arbitraire (cf encadré 5-3 joint). Certains auteurs substituent aux coûts d'infrastructures les coûts de congestion, définis eux-mêmes d'une façon différente d'une étude à l'autre ; d'autres les ajoutent partiellement.
- Quant à la finalité des calculs, qui sont souvent souvent un bilan entre ce que les transports coûtent au reste de l'économie et ce qu'ils apportent financièrement sous forme de taxes diverses.
- Quant aux différences entre résultats américains et européens. En Amérique du Nord les taxes sur les carburants sont sensiblement plus faibles, ce qui influe sur le sens des bilans entre ce que les transports paient et ce qu'ils rapportent au reste de la collectivité. En terme de méthode, les études américaines prennent en compte beaucoup plus systématiquement les coûts de parking, qui sont moins d'être négligeables.

ENCADRÉ 5-3

Quelques aspects de la répartition des coûts fixes

La répartition des coûts fixes d'une activité multiproduits n'a pas de solution incontestable. On peut envisager ce problème avec un objectif d'efficacité, et on aboutit alors à des formules du type Ramsey-Boiteux faisant intervenir la demande. On peut aussi l'envisager avec un objectif d'équité qui sera développé ici.

Prenons pour simplifier la situation où les différentes catégories de trafic - qu'on suppose pour simplifier au nombre de trois - ne supportent que des coûts fixes. On est alors en présence d'un coût total :

$$C(q_1, q_2, q_3)$$

$$C(q_1, 0, 0), C(q_1, q_2, 0)$$

et de coûts partiels

et les équivalents par changement d'indice.

L'idée d'imputer à chaque trafic la dépense supplémentaire qu'il occasionne se heurte au fait que cette dépense supplémentaire dépend de l'ordre dans lequel on prend les trafics :

$$C(q_1, q_2, q_3) - C(0, q_2, q_3)$$

sera en général différent de :

$$C(q_1, q_2, 0) - C(0, q_2, 0)$$

On est amené à considérer des répartitions telles qu'aucun sous-ensemble de trafics n'ait intérêt à se grouper pour réaliser et utiliser une autre infrastructure, ce qui revient à rechercher une imputation qui soit dans le cœur du jeu coopératif défini par la fonction caractéristique C . Mais le cœur peut être vide (cette situation se présente si le monopole n'est pas soutenable). Prenons par exemple le cas de 3 trafics dont l'écoulement nécessite 2 voies, alors que 2 trafics seulement n'en justifieraient qu'une ; cela correspond à la situation où :

$$C(q_1, q_2, q_3) = 2$$

et pour toutes les autres combinaisons :

$$C = 1$$

Alors, on voit facilement qu'aucun système de répartitions P_i ne vérifie à la fois :

$$P_1 + P_2 + P_3 = 2 = C(q_1, q_2, q_3)$$

et :

$$P_1 + P_2 \leq 1$$

et les relations similaires obtenues par permutation de la précédente.

Une autre méthode de répartition des dépenses fixes consiste à prendre une moyenne pondérée des différents coûts incréments envisageables, par exemple pour le trafic q_1 :

$$P_1 = \frac{1}{3} \left((C(q_1, q_2, q_3) - C(0, q_2, q_3)) + \frac{1}{2} (C(q_1, q_2, 0) - C(0, q_2, 0)) \right) + \frac{1}{2} (C(q_1, 0, q_3) - C(0, 0, q_3)) + \frac{1}{2} (C(q_1, 0, q_3) - C(0, 0, 0)) + \frac{1}{3} C(0, 0, 0)$$

Plus généralement :

$$P_i = \frac{1}{n!} \left(\sum_{s=0}^{s=n-1} s!(n-s-1)! \sum_{\bar{s}} (C(q_{\bar{s}}, q_i) - C(q_{\bar{s}}, 0)) \right) + \frac{C(0)}{n}$$

dans cette formule

$$\sum_{\bar{s}} C(q_{\bar{s}}, q_i)$$

représente toutes les combinaisons possibles de choix de s produits parmi les $(n-1)$ en dehors de i , plus le produit i . Cette allocation possède les propriétés intéressantes d'être indépendante de l'ordre de prise en compte des produits, de ne dépendre que des coûts incréments, et de fournir un prix nul pour un produit qui ne coûte rien à fournir.

Mais elle n'est pas forcément dans le cœur du jeu coopératif correspondant (Curien et Gensollen, 1992.)

D'autres modèles de répartition, fondés également sur la théorie des jeux coopératifs, sont développés, notamment aux USA (voir Quinet, 1997).

8.1. Études américaines

Les études américaines sont nombreuses. Gomez-Ibanez (1995 et 1997) en a recensé plusieurs. Elles sont relatives au transport routier, et exprimées en cents par miles ; elles ne comportent ni le coût de congestion, ni le coût monétaire pour l'automobiliste (tableau 5-6).

Direction (1992) fournit une estimation des coûts par passager et kilomètre pour les différents modes de transports interurbains de voyageurs au Canada.

TABLEAU 5.-
Coût par passager kilomètre au Canada
en cent canadien par passager kilomètre

Type de coût	Automobile			Autocar		
	Usagers	Autres	Total	Usagers	Autres	Total
Infrastructure	0,0	2,1	2,1	0,0	0,3	0,3
Environnement	0,0	0,6	0,6	0,0	0,2	0,2
Accidents	3,7	0,1	3,8	3,4	0,0	0,4
Taxes/redevances						
spéc. de transp.	1,2	-1,2	0,0	1,5	-0,3	0,0
Véhicules/transporteurs	10,9	0,0	10,9	8,4	0,2	8,6
Total	15,8	1,6	17,4	9,1	0,4	9,5
	Avion			Train		
Type de coût	Usagers	Autres	Total	Usagers	Autres	Total
Infrastructure	2,2	3,4	5,6	2,5	0,0	2,9
Environnement	0,0	1,0	1,0	0,0	0,6	0,6
Accidents	0,1	0,0	0,1	0,0	0,0	0,0
Taxes/redevances						
spéc. de transp.	0,6	-0,6	0,0	0,4	-0,4	0,0
Véhicules/transporteurs	14,4	0,1	14,5	7,2	3,5	40,2
Total	17,3	3,8	21,1	10,0	3,0	43,9
	Bac			Voyages interurbains totaux		
Type de coût	Usagers	Autres	Total	Usagers	Autres	Total
Infrastructure	0,0	4,7	4,7	0,0	0,0	2,4
Environnement	0,0	2,0	2,0	0,0	0,6	0,6
Accidents	0,1	0,0	0,1	0,0	0,2	0,2
Taxes/redevances						
spéc. de transp.	0,9	-0,9	0,0	0,0	-1,1	0,0
Véhicules/transporteurs	24,1	11,6	35,7	11,2	0,2	11,4
Total	25,1	17,4	42,5	11,2	0,0	17,8

TABLEAU 5-6
Coûts Sociaux de Transports de voyageurs aux USA (en US cents par voyageur-mile)

Estimation du WRI (automobile seulement)	Estimation de Apogees Research		Estimation du NRDC		Estimation de Litman	
	Auto	Bus urbain	Méto	Automobile	Bus	Zone urbaine
Coûts gouvernementaux	2,8	20,8	33,7	0,4	7,0	14,0
-Capital	0,3	59,3	44,5	2,4	43,0	31,0
-Fonctionnement	0,9	1,2	-	0,9	0,1	1,4
-Autres (police)	1,8	80,1	78,2	50,1	44,1	7,1
-Sous-total	0,8	4,3	3,7	0,1	7,5	2,2
Coûts sociaux	9,5	1,0	7,0	4,5	5,0	1,6
-Pollution de l'air	(-)	(-)	(4,6)	(2,7)	(1,1)	0,7
-Pollution de l'air (dont changement climatique)	(0,7)	1,0	0,0	4,5	5,0	0,7
- Bruit	-	-	0,1	0,1	1,2	0,1
- Pollution des eaux	1,4	0,6	1,1	0,7	0,6	0,1
- Accidents	0,7	1,9	0,5	2,8	1,3	0,0
- Energy	2,7	5,2	3,2	-	10,9	2,2
- Parking	5,9	14,4	0,01	7,4	8,2	0,0
- Autres	14,4	3,0	1,1	19,2	50,2	12,0
-Sous-total	1,0	14,8	37,0	14,0	14,0	19,1
-Péages	3,0	3,0	0,0	0,0	0,0	0,0
-Taxes	1,0	14,8	37,0	14,0	14,0	0,9
-Sous-total	8,3	15,7	42,3	22,2	43,5	12,1
Subvention nette	19,1	19,1	19,1	19,1	19,1	45,7

Source : Gomez-Ibanez (1997).

Enfin, citons les résultats d'une étude intermédiaire sur le corridor Los Angeles - San Francisco, qui fournit les chiffres suivants, en dollars par passager* kilomètre (Levinson et al., 1994).

TABLEAU 5-5
Coûts à long terme sur le corridor Los-Angeles/San Francisco
(US cents par voyageur-kilomètre)

Catégorie de coût	Avion	Train à grande vitesse	Autoroute
Infrastructure (construction, entretien)	1,82	12,90	1,20
Opérateur : capital	4,56	1,00	0,00
Opérateur : exploitation	2,43	5,00	0,00
Accidents	0,02	0,00	2,00
Congestion	0,00	0,00	0,46
Bruit	0,43	0,20	0,45
Pollution	0,09	0,00	0,31
Usager : coût monétaire	0,00	0,00	8,60
Usager : temps	0,00	4,40	10,00
Total	13,92	23,50	23,02

8.2. Études européennes

On peut d'abord prendre une vue générale des coûts selon leur catégorie économique. Quinet (1994) indique les ordres de grandeur suivants, en pourcentage de PIB :

dépenses monétaires	: 14 %
temps passé	: 8 %
sécurité	: 2 %
pollution	: 0,5 %
bruit	: 0,3 %

On peut aussi en prendre une vue plus détaillée selon le pays ou les conditions de transport. Les chiffres suivants, tirés de Quinet (1997) recensent différentes études récentes européennes qui comparent les coûts de transport pour la collectivité et les taxes qui sont appliquées aux différents modes (tableau 5-9).

TABLEAU 5-9
Quelques coûts sociaux des transports - Poids lourds
En FF par veh. x km

Études Pays Années	Jeanrenaud Suisse 1988		Brossier France 1992		Quinet France 1992
	Côté total 19 t	Côté marginal 44 t	Côté total 44 t	Côté marginal 19 t	
Routes nationales hors péage					
Autres particularités	Côté total 19 t	Côté marginal 44 t	Côté total 44 t	Côté marginal 19 t	Tous réseaux confondus
Congestion	-	-	-	0,382	0,669
Infrastructure	1,13	0,636	1,008	0,186	0,318
Accident	-	-	-	0,156	0,156
Bruit	0,82	-	-	-	-
Pollution locale	-	-	0,944	0,281	0,944
Pollution globale	-	-	-	-	-
Total	1,96	0,918	1,951	1,005	2,086
Coût des taxes actuelles	1,16	0,416	0,680	0,416	0,680
					0,70

Orfeuil (1997) établit pour la France un bilan très complet des coûts et revenus occasionnés par la circulation routière, d'où il ressort que la circulation interurbaine paie globalement ses coûts, mais pas la circulation urbaine (tableau 5-10).

Les valeurs précédemment citées concernent surtout les coûts collectifs. Paradoxalement, il est plus difficile de trouver en Europe des valeurs de coûts monétaires pour les opérateurs.

8.2.1. Coûts des opérateurs voyageurs

Rappelons les coûts de l'automobile cités précédemment ; ils varient en fonction du kilométrage annuel et de la puissance du véhicule entre 1,68 F et 3,57 F. En zone urbaine, avec un taux d'occupation moyen de 1,2 passagers, on arrive à des valeurs par passagers et kilomètres situées entre 1,50 F et 3 F. En parcours interurbain, le taux d'occupation est de l'ordre de 2 à 2,5 et le coût par passager se situe entre 0,80 F et 1,80 F.

Quin (1995) a calculé des coûts kilométriques pour différents types de transports de voyageurs urbains. Il étudie pour chaque mode, différentes situations de coût d'investissement et de volume de trafic.

Il aboutit aux fourchettes figurées dans le tableau 5-11.

TABLEAU 5-10
Bilan de la circulation routière (centimes par km, 1991)

	Deux roues		Voitures		Utilitaires		Poids lourds		Autobus Autocars		Total
	E	D	E	D	E	D	E	D	E	D	
<i>Circulations urbaines</i>											
Revenu des taxes	15	43	20	38	21	105	117	35			
Dépenses routières	13	23	23	30	31	235	83	27			
Coûts externes	75	43	42	38	37	162	142	45			
Bruit	6	6	6	7	6	20	13	6			
Pollution	20	9	8	11	9	72	95	11			
Effet de serre	1	3	3	3	5	20	17	4			
Accidents	47	14	14	2	2	15	17	13			
Congestion	0	11	11	16	16	35	0	12			
Bilan	-73	-23	-46	-30	-47	-292	-108	-37			
<i>Rase campagne</i>											
Revenu des taxes	23	36	18	28	19	94	129	36			
Dépenses routières	13	16	16	18	18	83	77	22			
Coûts externes	40	12	12	7	7	36	34	15			
Bruit	0	1	1	1	1	2	2	1			
Pollution	4	1	1	1	1	13	14	2			
Effet de serre	1	2	2	2	3	10	9	3			
Accidents	34	8	8	3	3	11	9	9			
Congestion	-	-	-	-	-	-	-	-			
Bilan	-30	+9	-9	+3	-6	-25	+18	-1			

Source : Orfeuil, 1997.

TABLEAU 5-11
Coûts de quelques opérateurs de transport urbain

	Coût pour l'exploitant en F. par passager et kilomètre :		Min.	Max.
Autobus urbain (1)	1,0		1,0	2,0
Tramway (2)	0,6		0,6	2,0
Métro léger (2)	2,0		2,0	4,0
Métro classique (2)	1,5		1,5	2,0
Rappel : Automobile (coût pour l'utilisateur)	0,8		0,8	0,8
rase campagne	1,8		1,8	1,8
milieu urbain				

(1) Non compris dépenses d'infrastructures.

(2) Y compris dépenses d'infrastructures.

Source : Quin, 1995.

En matière ferroviaire, divers recoupements conduisent aux ordres de grandeur suivants en F par voyageur et kilomètre :

- banlieue parisienne : 0,7 F
- services régionaux de province : 1,0 F
- Trains rapides et express : 0,5 F

TABLEAU 5-9 (suite)
Quelques coûts sociaux des transports (automobiles) en F par véh. x km

Étude	Pays	Années	Autres particularités	Londres	Inter-urbain	Inter-urbain	Coût marginal	Coût (2) marginal	Urban	Non-urbain	Inter-urbain	Urban	Inter-urbain	Urban	Coût total	Coût actuel
Pierson et al.	UK	1990		1,98	1,02	1,8	0,46	0,82	1,50	1,50	1,50	1,50	1,50	1,50	3,57	4,28
Jeanrenaud	Suisse	1988		2,66 (3)	0,32	0,42	0,13 (1)	1,50	1,50	1,50	1,50	1,50	1,50	1,50	3,79	4,28
Brossier	France	1992 (4) (5)		1,47	0,27	0,82	2,09	1,2	1,50	1,50	1,50	1,50	1,50	1,50	4,72	4,28
Direction	Canada	1991		1,2	0,27	0,82	2,09	1,2	1,50	1,50	1,50	1,50	1,50	1,50	3,5	0,7
Quin	France	1990		2,13	0,12	0,30	2,52	0,31	0,60	0,60	0,60	0,60	0,60	0,60	3,42	3,81
Hansson	Suède	1995		1,8	0,47	0,47	1,05	0,31	0,60	0,60	0,60	0,60	0,60	0,60	2,54	3,35
Quinet	France	1992		1,21	0,66-6,6	0,47	1,38	0,22	0,54	0,54	0,54	0,54	0,54	0,54	3,42	4,76
				18,10	1,98	1,8	1,8	8,54	2,04	2,04	2,04	2,04	2,04	2,04	13,62	13,50
				7,85	6,18	6,18	2,66 (3)	1,47	1,50	1,50	1,50	1,50	1,50	1,50	3,57	4,28
				1,98	1,02	1,8	0,46	0,82	1,50	1,50	1,50	1,50	1,50	1,50	3,79	4,28
				2,66 (3)	0,32	0,42	0,13 (1)	1,50	1,50	1,50	1,50	1,50	1,50	1,50	3,79	4,28
				0,32	0,10	0,47	1,05	0,31	0,60	0,60	0,60	0,60	0,60	0,60	2,54	3,35
				0,47	0,47	0,47	1,05	0,31	0,60	0,60	0,60	0,60	0,60	0,60	2,54	3,35
				2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	13,62	13,50
				2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	13,62	13,50
				13,62	7,92	7,92	7,92	7,92	7,92	7,92	7,92	7,92	7,92	7,92	30,66	30,66
				30,66	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	30,66	30,66
				30,66	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	13,62	30,66	30,66
				13,50	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	13,50	13,50
				13,50	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	12,05	13,50	13,50

(1) Dommages aux bâtiments seulement.
(2) Répartition selon les règles officielles britanniques.
(3) Tiré de Ferrer et Jacard.
(4) Chiffres par véh. x km.
(5) Trafic interurbain seulement.

Source : Quinet, 1997.

Ces derniers chiffres ne comprennent pas la part des dépenses d'infrastructure financée par l'État. Leur réintégration conduirait à augmenter d'environ 30 % à 40 % les ordres de grandeur précédents.

Par ailleurs, Metzler (1995) fournit la figure 5-6 qui donne le coût au voyageur \times km et au train \times km en fonction du nombre de voyageurs.

Enfin, en divisant les charges des entreprises de Transport aérien par leur trafic total, on aboutit aux chiffres suivants :

Air France : $0,8 \text{ F par passager} \times \text{ km}$

Air Inter : $1,2 \text{ F par passager} \times \text{ km}$

Ces valeurs sont des moyennes autour desquelles la dispersion est grande.

8.2.2. Coût des opérateurs Marchandises

La Commission des Comptes de Transports de la Nation (édition de 1997) fournit des prix moyens du transport routier de marchandises, un secteur où les prix sont proches des coûts. Ces prix varient de 1,10 F pour les transports de zone courte (inférieur à 50 km) à 0,45 F/tonnes \times km pour les relations internationales à longue distance. Les chiffres sont très dispersés, et même pour la dernière catégorie, on cite des transports qui s'effectuent aux prix de 0,30 F/t \times km et même moins. Mais en fait la bonne mesure du coût du transport routier est le véhicule-kilomètre. Son prix de revient varie de 6 à 8 F pour un véhicule de 40 tonnes de PTCA (la variation n'est pas proportionnelle au tonnage), et le coût à la tonne-kilomètre s'en déduit par division par le tonnage transporté.

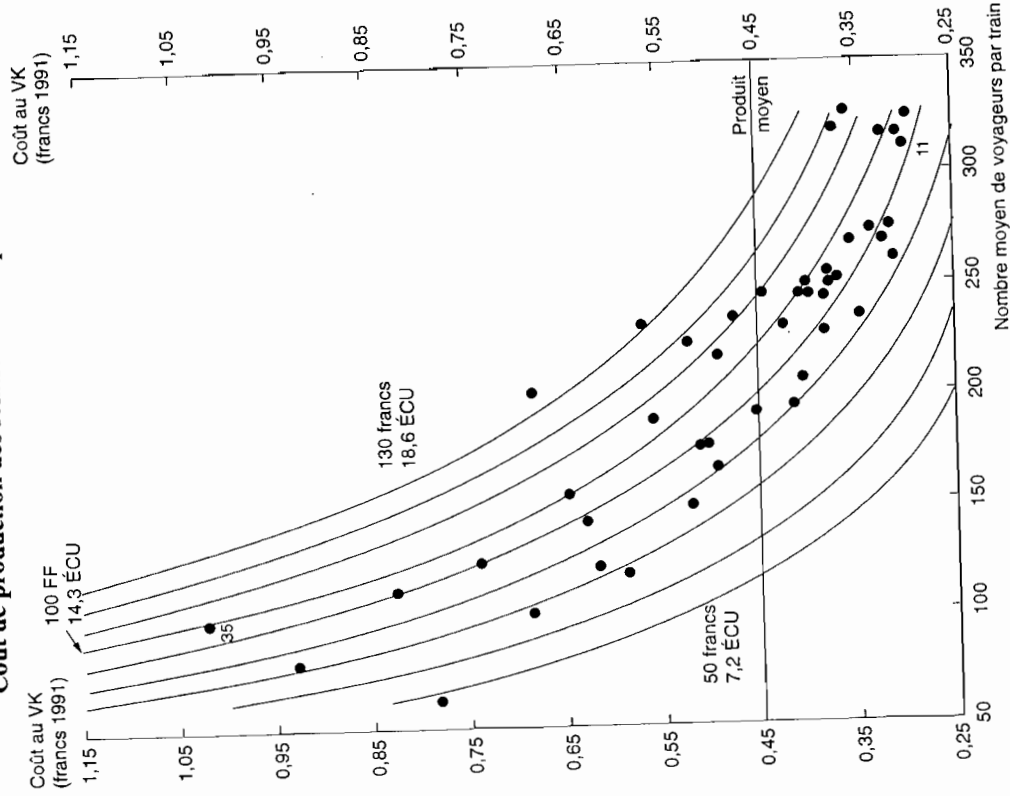
Divers recoupements donnent à penser que le coût moyen ferroviaire, de l'ordre de 0,30 F/t \times km, recouvre des écarts très grands entre les trains complets (environ 0,10 F/t \times km) et les wagons isolés (de l'ordre de 0,8 F/t \times km).

Enfin, le trafic par voie navigable s'effectue à un coût moyen de l'ordre de 0,20 F par tonne et km. Mais cette moyenne recouvre aussi des écarts sensibles entre les transports par les grands convois poussés de plusieurs milliers de tonnes (moins de 0,10 F/t \times km) et les transports par péniches Freycinet de 38,5 m et environ 380 t (0,50 F t \times km).

9. Conclusion

On peut considérer dans le coût des transports des éléments de moins en moins bien connus, en allant des dépenses monétaires aux coûts d'environnement et en passant par les coûts d'infrastructures. Les dépenses monétaires directes de l'automobile comportent une part d'amortissement dont le montant est sujet à appréciation. Une incertitude frappe aussi les coûts d'infrastructure dans la mesure où ces coûts comportent une proportion élevée de dépenses fixes dont la répartition entre les catégories de trafic présente une importante dose d'arbitraire. Enfin, les coûts

FIGURE 5-6
Coût de production des relations classiques en 1991



Notes : Chaque point indique le coût au VK d'une relation.

Chaque courbe représente un même niveau de coût au TK.

Ex. : Paris-Strasbourg (n° 11) coût au VK = 0,3 F ; coût au TK = 93 F.

Paris-Dieppe (n° 35) coût au VK = 1,02 F ; coût au TK = 98 F.

d'environnement, d'insécurité et aussi, dans une moindre mesure, les dépenses de temps, sont entachés d'un plus grand aléa qui est à la fois d'ordre physique (quelles quantités de substances polluantes sont reçues par les riverains ?) et d'ordre économique (quel est le coût des dommages causés par ces substances polluantes ?).

Les coûts de transports dépendent aussi d'un grand nombre de paramètres : le type de véhicule, la composition du trafic, le lieu, l'heure, la nature de l'infrastructure. Le raisonnement en moyenne auquel leur lecture invite est très trompeur, comme l'illustre l'encadré 5-4, d'autant que les unités statistiques de trafic (voy. \times km et tonne \times km) ne sont pas les bons paramètres de coût.

On en tire souvent des comparaisons en équité sur ce que sont diverses catégories de trafic. On trouve par exemple, c'est le cas dans beaucoup de pays européens, que l'automobile paie sous formes de taxes, à peu près son coût. Mais c'est un raisonnement moyen, et une analyse plus fine montrerait que le bilan est fortement négatif pour les déplacements urbains. De même si, en gros les poids lourds paient un peu moins que ce qu'ils coûtent, ceci est la moyenne de situations très contrastées. Les petits poids lourds paient en fait plus que leur coût (ceci n'est d'ailleurs vrai qu'en moyenne là aussi : ils paient nettement plus en rase campagne et plutôt moins en zone urbaine...) et les gros poids lourds nettement moins dans des proportions qui croissent avec leur taille.

Et ce qui est vrai en terme de niveau l'est aussi en terme de variations. Les économies d'échelle se logent un peu partout, et si elle sont plus fréquentes dans les chemins de fer, on en rencontre aussi dans le transport routier de marchandise.

Enfin, les évolutions temporelles sont fortes. Elles se marquent d'abord par une décroissance des coûts monétaires, comme on l'a vu au chapitre 1 : sur plusieurs décennies, les coûts de transports n'ont cessé de décroître par rapport aux prix du PIB, et pour beaucoup de modes, en particulier l'automobile les prix du transport pour le consommateur se sont aussi réduits par rapport aux prix à la consommation. Ce que d'ailleurs les données statistiques sur les prix ne font pas apparaître, c'est que cette baisse du prix du transport s'est accompagnée d'une amélioration de la qualité, en vitesse, confort et commodité : quoi de commun entre les avions actuels et ceux des années 50 ?

En même temps que les coûts monétaires se réduisaient, les autres coûts, ceux qui ne sont pas directement perçus par l'utilisateur se sont accrues. Certes, comme l'ont fait voir les chiffres fournis en section 8, ils restent minoritaires par rapport aux coûts supportés par l'exploitant ou l'automobiliste. Mais leur part augmente. C'est le cas des dépenses d'infrastructure, pour une raison qui n'est pas forcément durable : leur croissance rapide a été occasionnée par le développement des trafics, selon le mécanisme bien connu de l'accélérateur qui relie le niveau de l'investissement à la croissance du trafic : si le trafic est stable, aucun investissement n'est nécessaire, si ce n'est de remplacement ; si le trafic augmente de 10 %, un investissement de I est nécessaire ; si l'augmentation est de 20 %, c'est en gros $2I$ qu'il faut. C'est ce mécanisme, du type de l'accélérateur, qui explique le développement depuis la guerre des investissements routiers, autoroutiers et aéroportuaires, ainsi que l'explosion des infrastructures de transport en ville, liée à l'urbanisation.

ENCADRÉ 5-4

Les erreurs issues de raisonnements en moyenne

Soient deux modes F et R dont les coûts unitaires CF et CR dépendent d'un paramètre L , uniformément réparti entre 0 et L_1 (L peut être la distance de transport, ou le volume de l'envoi) :

$$CF(L) = a_F + Lb_F$$

$$CR(L) = Lb_R$$

Le jeu du marché est tel que les envois pour lesquels : $L < L_0 = (a_F)/(b_R - b_F)$ prendront R et les autres F : les statistiques font donc apparaître un coût moyen (par km ou par tonne, selon la signification de L) observé de :

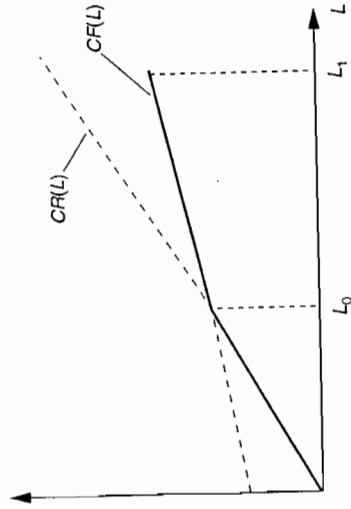
$$CMR = b_R \quad \text{pour le mode } R$$

$$CMF = b_F + a_F/2(L_0 + L_1) \quad \text{pour le mode } F$$

et bien évidemment :

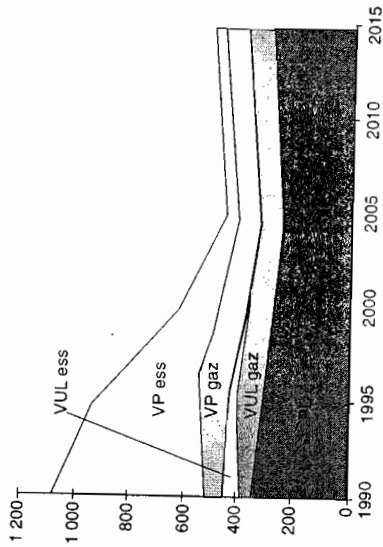
$$CMF < CMR$$

Un raisonnement trop rapide sur les moyennes observées conduira à préconiser un transport de trafic de R vers F . L'erreur provient de ce qu'on a raisonné sur des moyennes en omettant de tenir compte des dispersions autour de ces moyennes.



Et maintenant, ce sont les coûts d'environnement qui apparaissent et prennent une importance qui va croissante. Cette importance s'explique pour diverses raisons : d'abord, les atteintes à l'environnement se développent avec la croissance des trafics. Certes les moyens de lutte contre ces atteintes se développent aussi, mais avec retard. Ainsi, la pollution par les oxydes d'azote est combattue par les pots catalytiques qui vont pratiquement la supprimer ; mais le renouvellement du parc automobile et la généralisation au sein de ce parc de l'essence sans plomb prendront de nombreuses années, et la décroissance des émissions de NO_x sera lente, comme le montre la figure 5-7, tiré de Girault et Kail (1997).

FIGURE 5-7
Évolution prévue des émissions de NOx



En même temps, le coût unitaire de ces atteintes à l'environnement va croissant, et cela pour deux raisons. La première tient aux développements de la connaissance ; ainsi, en matière de pollution de l'air, l'attention s'est d'abord portée sur les dérivés azotés et sulfuriques dont on a assez rapidement analysé les dangers pour les voies respiratoires ; et maintenant l'attention se déplace vers les particules qu'on soupçonne d'être cancérigènes. La seconde est subjective : les préférences pour les valeurs de l'environnement vont croissantes comme le montrent les débats suscités dans l'opinion publique par les épisodes de pollution de l'air.

Le ciseau qui se produit entre les baisses des dépenses monétaires relatives à des transports de qualité sans cesse améliorée donc en volumes croissants, et les hausses de leur coûts externes, en particulier ceux que subissent les non-usagers, est à la fois l'expression et la cause d'une bonne partie des sentiments d'insatisfaction que suscite actuellement l'évolution du secteur des transports.

BIBLIOGRAPHIE

- W. Baumol, J. Panzar et R. Willig (1982), *Contestable markets and the theory of Industrial Structure*, New York ; Harcourt Brace Jovanovitch.
- J. Berechman (1983), « Costs, economics of scale and factor demand in bus transport », *Journal of Transport Economics and Policy*, January.
- M. Beuthe et A. Savez (1994), *The multiproduct cost function of long distance trucking in France*, Faculté catholique de Mons, ronéoté.
- A. Bonnafous (1992), « Transports et environnement : comment valoriser et maîtriser les effets externes ? », *Économie et statistique*, oct.-nov.
- C. Brossier (1996), *Nouvelle étude de l'imputation des coûts d'infrastructure de Transports*, Ministère des transports, Paris.

- S. Callan et J. Thomas (1992), « Cost differential among household goods carriers », *Journal of Transport Economics and Policy*, January.
- D.W. Caves, L.R. Christensen et M.W. Tretheway (1984), « Economies of diversity versus economies of scale. Why trunk and local service Airlines costs differ ? », *Rand Journal of Economics*, 15.
- CEMT (1998), *Les transports efficaces pour l'Europe*.
- W. Chiang, S. Judy et A. Friedlander (1984), « Output aggregation, network effects, and the measurement of trucking technology », *Review of Economics and Statistics* 64.
- Direction (1992), « The final report of the Royal Commission on National Passenger Transportation », Ministry of Supply and Services, Canada.
- R. Doganis and G.F. Thompson (1975), « The economics of regional airports », *International Journal of Transport Economics*.
- D. Encaoua et A. Perrot (1991), *Concurrence et coopération dans le transport aérien en Europe*, Rapport de recherche, Université de Paris I.
- A. Friedlander, E. Berndt, J. Shaw, E.W. Chiang, M. Showalter et C. Velluro (1993), « Rail costs and capital adjustments in a quasi-regulated environment », *Journal of Transport Economics and Policy*, May.
- P. Gagnepain (1998), *Efficacité, incertains et réglementation : économétrie des contrats de transport urbain*, thèse sous la direction de C. Crampes, Toulouse I.
- H. Geehan (1993), « Railway costs and productivity growth », *Journal of Transport Economics and Policy*, January.
- D. Gillen, T.H. Oum et M. Tretheway (1990), « Airline cost structure and policy implications », *Journal of Transport Economics and Policy*, January.
- M. Girault et J.M. Kail (1997), « Perspectives de la demande de transport et des émissions de polluants à l'horizon 2015 », Communication au Congrès ATEC Janvier, Versailles.
- J. Gomez-Ibanez (1995), « Pitfall in estimating whether transport users pay their way » Communication à la Conférence : « Measuring the full social costs and benefits of transportation », Irvine, June 95.
- J. Gomez-Ibanez (1997), « Estimating whether Transport users pay their way. The state of the art », in Greene, Jones and Delucchi (1997).
- D. Greene, D. Jones et M. Delucchi (1997), *The full social costs and benefits of Transportation*, Springer-Verlag.
- D. Harmatuck (1981), « A motor carrier joint cost function », *Journal of Transport Economics and Policy*, May.
- D. Harmatuck (1991), « Economies of scale and scope in the motor carrier industry », *Journal of Transport Economics and Policy*, May.
- S.R. Jara-Diaz et C.E. Cortes (1996), « On the calculation of scale economies from transport cost functions », *Journal of Transport Economics and Policy*.
- C. Jeanrenaud (1993), *Les coûts sociaux des transports en Suisse*, IRER, Neuf-châtel.
- M.W. Jones-Lee (1990), « The value of transport safety », *Oxford Review of Economic Policy*, vol. 6.
- M.W. Jones-Lee and G. Loomes (1994), « Towards a willingness to pay based value of underground safety », *Journal of Transport Economics and Policy*, vol. XXVIII, n° 1.
- M. Keeler (1973), « Airport cost and congestion », *The American Economist*.

- M. Kirby (1986), « Airline economies of scale and Australian domestic air transport policy », *Journal of Transport Economics and Policy*, Sept.
- S.C. Kolm (1968), « Prix publics optimaux », *Cahiers du séminaire d'Économétrie*, n° 9, Paris.
- S.C. Kolm (1970a), « L'inégalité des valeurs des vies humaines », *Cahiers du séminaire d'Économétrie*, n° 18. CNRS.
- S.C. Kolm, (1970b), *Le service des masses*. Dunod. Paris.
- D. Levinson, D. Gillena, Kanafani et J.M. Mathieu (1996), « The full cost of intercity transportation », Research Report. Institute of Transport Studies. UCB ITS RR 96-3, University of California at Berkeley.
- H. Lévy-Lambert (1968), « Sur la tarification des biens à qualité variable avec le trafic », *Econometrica*, 36.
- H. Mauch et W. Rotengather (1996). *External cost of land transport*, Rapport pour l'UIC.
- J.M. Metzler (1995), « Rapport France », in *Les coûts du transport interurbain*, CEMT, Paris.
- C. Nash (1997), « Does monetary valuation of environnement make sense ? » in de Rus and Nash, *Recent developments in Transport Economics*, Ashgate.
- D. Newbery (1988), « Road damage externalities and road users charges », *Econometrica*, March.
- J.P. Orfeuil (1997), *Les coûts externes de la circulation routière*, Rapport INRETS n° 216.
- T.H. Oum et Y. Zhang (1991), « Utilisation of quasi-fixed inputs and estimations of cost functions », *Journal of Transport Economics and Policy*, May.
- T.H. Oum et Y. Zhang (1997), « A note on scale economies in transport », *Journal of Transport Economics and Policy*, September.
- J. Pavaux (1994), *L'économie du transport aérien*, Economica, Paris.
- D. Pearce et A. Markandya (1989), *L'évaluation monétaire des avantages des politiques de l'environnement*, Rapport OCDE, Paris.
- J. Peirson, I. Skinner et R. Vickermann (1994), « The taxation of external effects and the efficient supply of transport », discussion paper n° 94-3. CERTE, Canterbury.
- Planco (1990), *Coûts externes du transport par route, fer et voie navigable*, Rapport.
- J. Preston (1994), « The economics of rail privatization », Conference au séminaire de l'École Polytechnique, octobre 1994.
- E. Quinet (1992), *Route, Air, Fer*, Rapport pour le Ministre des Transports, Paris.
- E. Quinet (1994), « The social costs of Transport », in OCDE. *Internalizing the social costs of transport*.
- E. Quinet (1997), « Full social cost of transportation in Europe », in Greene, Jones and Delucchi (1997).
- E. Quinet et C. Galland (1995), *Les dépenses de sécurité et la valeur de la vie humaine dans les transports collectifs de voyageurs*. Rapport CERAS-ENPC pour le Ministère des Transports.
- C. Quin (1994), *Analyse des coûts de déplacement : élaboration d'une méthodologie dans le cadre d'un compte transport de voyageurs*. Rapport pour le Ministère des Transports, Paris.
- K. Small, C. Winston et C.A. Evans (1989), *Road works*, Washington D.C., the Brookings Institution.
- K. Small (1992), *Urban transportation economics*, Harwood Academic Publishers, Chur (Switzerland).
- H. Tauchen, F. Fravel et G. Gilbert (1983), « Cost structure of the intercity bus industry », *Journal of Transport Economics and Policy*, January.
- R. Tinch (1996), *Estimation des prix fictifs des externalités dans les transports*, Rapport CEMT/CS/SOC (96) 5.
- Tolofari, N. Ashford et Caves (1990), *The cost of air service fragmentation*, Loughborough University.
- H. Varian (1992), *Micro-economic analysis*, W.W. Norton and Company.
- W.K. Viscusi (1993), « The value of risks to life and health », *Journal of Economic Literature*, vol. XXXI.
- T. Walker (1992), « Network economics of scale in short hand truckload operations », *Journal of Transport Economics and Policy*, January.
- C. Winston (1985), « Conceptual developments in the economics of transportation : an interpretative survey », *Journal of Economic Literature*, vol. XIII.
- K. Xu, R. Windle et C. Grimm (1994), « Reevaluating returns to scale in transport », *Journal of Transport Economic and Policy*, September.

TROISIÈME PARTIE

L'organisation de l'offre

Nature des marchés et interventions des pouvoirs publics

L'offre prend des formes très diversifiées dans les transports. Certains opérateurs intégrés établissent un service complet, comme c'est le cas actuellement pour les chemins de fer ; d'autres se bornent à fournir à d'autres opérateurs une partie du service de transport. On y rencontre dans certains sous-secteurs des monopoles et des oligopoles. Dans d'autres les entreprises sont très nombreuses. On y trouve enfin une implication variable, mais toujours forte, des pouvoirs publics, allant de la gestion directe à une réglementation de l'activité privée, en passant par les concessions de service public à des opérateurs des statuts divers. Passons en revue les formes d'organisation de la production, qui se caractérisent par trois dimensions : intégration ou fragmentation, monopole ou concurrence, libéralisme ou intervention.

ENCADRÉ 6-1

Production des transports.

La diversité des situations en France

- Les ~~transports ferroviaires~~ interurbains sont assurés par la SNCF, service public disposant d'un monopole encore quasi-total.
- Les ~~ports autonomes~~ et les ~~aéroports~~ sont des établissements publics, concessionnaires exclusifs d'un service public sur le territoire de leur compétence.
- Il en est de même pour la RATP en ce qui concernent les transports collectifs de la région Ile-de-France dans leur zone centrale. Les transports collectifs de la zone périphérique de l'Ile-de-France sont assurés par l'APTR et l'ADATRIF, regroupements d'opérateurs privés liés par contrat avec la Région.
- Les ~~routes~~ ordinaires sont construites et exploitées directement par la puissance publique (Etat ; départements, communes).
- Les sociétés d'ouvrages à ~~péage~~ (société autoroutières, ponts, tunnels) sont concessionnaires de l'ouvrage pour des périodes de quelques dizaines d'années. Elles sont de statuts publics ou privés.

intelligents en cours de développement ; c'est aussi le niveau de la gestion des sillons dans les chemins de fer ;

- le niveau des services, constitué par l'offre présentée aux consommateurs finaux, et caractérisé par les prix, les fréquences, les qualités de services... Dans l'organisation interne des opérateurs de ce niveau on pourrait d'ailleurs distinguer également trois niveaux d'activité ; ainsi une entreprise de transport routier ou aérien est composée d'un premier niveau, à caractère d'infrastructure, constitué par les dépôts, entrepôts et véhicules ; l'infrastructure est composée du système de gestion des véhicules et de gestion de la demande, par exemple les Systèmes Informatisés de Réserve (SIR) des compagnies aériennes ; enfin le niveau des services correspond à ce qui est présenté à la clientèle.

1.2. Coordination hiérarchique ou par le marché

L'articulation entre ces trois niveaux peut s'effectuer à travers des opérateurs distincts en relation de marché ou de marchandage, ou par un opérateur intégré gérant à la fois deux ou trois des niveaux précédents, ou encore par un système de partenariat de long terme intermédiaire entre les deux cas extrêmes précédents. Les avantages et défauts de l'une ou l'autre de ces organisations sont évalués différemment selon les pays et selon les époques. Ainsi dans les chemins de fer, les trois niveaux sont intégrés en France, séparés en Angleterre.

La théorie économique fournit à ce sujet les grandes orientations suivantes :

- Le marché n'est efficace, n'aboutit à un optimum que si les agents qui y opèrent sont nombreux, chacun petit et parfaitement informés ; il perd ses vertus en situation de monopole ou oligopole.
- Le marchandage aboutit, selon le théorème de Coase, à un optimum, mais sous des réserves importantes (Henry, 1997) :
- Tous les agents concernés doivent participer à la négociation et pouvoir effectuer des transferts
- La négociation est sans coût
- L'information est parfaite.

- Dans plusieurs situations l'intégration est préférable à ces deux formes d'articulations. Ce sera par exemple le cas d'asymétrie d'information (encadré 6-2).

Ce sera aussi le cas lorsque la coordination des actions est essentielle. Ainsi, la coordination des trains appartenant à différents exploitants et circulant sur un même réseau ne peut pas, pour des raisons techniques évidentes, être effectuée selon des procédures aussi simples que celles des automobiles sur une route. On pourrait l'envisager sous la forme de contrats ; mais leur coût - le coût de transaction, voir encadré 6-3 - serait considérable ; on est alors amené à envisager l'existence d'un organisme répartiteur des sillons. Mais sa séparation du gestionnaire d'infrastructure risque d'entraîner de nombreuses inefficacités : interventions en cas d'ac-

- Les transports publics urbains sont assurés dans quelques agglomérations par des régies municipales, et dans la majorité des cas concédés à des sociétés privées pour des durées de 4 à 5 ans. La plupart du temps, le matériel appartient à la commune.

- Les transports publics interurbains par autocars sont réalisés dans leur majorité par des entreprises privées, qui sont sous contrat avec les autorités locales pour les lignes régulières.

- Le transport routier des marchandises est assuré soit par les entreprises bénéficiaires pour leur compte propre, soit par des transporteurs pour compte d'autrui qui ont le statut privé (même si les plus gros sont des filiales de la SNCF...), sont très nombreux (cf. chapitre 1), et soumis à une réglementation complexe à défaut d'être efficace.

- Le transport de marchandise par voie navigable est de même produit soit pour compte propre par les entreprises (cas par exemple des compagnies pétrolières, qui opèrent d'ailleurs souvent à travers des filiales), soit pour compte d'autrui par une profession composée de quelques grosses compagnies qui exploitent des barges et pousseurs de grandes capacités (quelques milliers de tonnes) et d'un grand nombre d'artisans exploitant chacun quelques péniches de faible tonnage (380 tonnes).

- Le transport aérien est assuré selon une structure très concentrée par une très grosse compagnie, Air France, compagnie nationale, et par quelques compagnies de taille réduite (Air Liberté...) et à statut privé.

Ce type de structure se retrouve par ou prou, avec quelques variantes, dans les autres pays européens. Les différences les plus fortes concernent les chemins de fer, les compagnies sont souvent plus morcelées et plus loin de l'État qu'en France. La situation est inverse pour les infrastructures routières, où c'est en France qu'on trouve la plus forte extension du système des concessions à péages. Le système aérien est beaucoup plus libéralisé et privatisé au Royaume-Uni et aussi, à un moindre degré, en Allemagne.

1. Intégration ou fragmentation

1.1. Les trois niveaux d'organisation

En adoptant une terminologie établie par Curien et Dupuy (1997), on peut distinguer trois niveaux dans l'offre de transport :

- le niveau de l'infrastructure, la partie la plus matérielle et la plus physique du système, qui est en gros constituée par ce qui est inscrit dans le sol : les routes, les rails, les aéroports, les ports ;

- le niveau de l'infrastructure, qui a pour but d'assurer la bonne utilisation de l'infrastructure. Ce niveau, pour le moment moins important que dans d'autres secteurs, se développe rapidement. C'est celui qui est assuré par les services de la navigation aérienne pour le transport aérien ; par la police de la route et par les services d'exploitation et d'information routiers : coordination des faux tricolores en ville, opérations de délestage lors des pointes de trafic interurbain, information routière comme le système SIRIUS en Région Ile-de-France, ou a travers les systèmes de véhicules

Marchandage et asymétrie d'information

Supposons qu'un vendeur propose une machine qui est dans un état d'entretien θ :

$\theta = 0$ si elle ne marche pas, $\theta = 1$ si elle est comme neuve.

Compte tenu d'un état d'entretien θ donné la valeur de la machine (somme des bénéfices actualisés qu'on peut en retirer) est de θV_1 . pour le vendeur, θV_2 pour l'acheteur. Si θ est connu du vendeur, mais inconnu pour l'acheteur, l'échange risque de ne pas se faire. En effet si le vendeur propose un prix p , l'acheteur en inférait que θ est inférieur à p/V_1 . Pour schématiser sa croyance, on peut prendre comme hypothèse qu'il pensera que θ est uniformément réparti entre 0 et p/V_1 . Donc la valeur la plus probable pour θ sera $p/2V_1$. Si l'acheteur est neutre au risque, l'utilité qu'il retirera de la machine sera en espérance mathématique :

$$pV_2/2V_1$$

et il ne fera l'achat que si

$$pV_2/2V_1 > p$$

soit :

$$V_2 > 2V_1$$

condition beaucoup plus restrictive que la condition d'optimalité à laquelle aboutirait un marchandage en information parfaite ; l'échange est avantageux dès que :

$$V_2 > V_1$$

cident, entretien des voies... Ne faut-il pas alors fusionner le gestionnaire de l'infrastructure et le réparateur des sillons, correspondant aux niveaux d'infrastructure et d'infrastructure ? C'est la solution adoptée en Grande-Bretagne. Dans la plupart des autres pays, où existe un opérateur principal, c'est lui qui a en charge la répartition des sillons.

Une autre situation de ce type est celle des investissements spécifiques, (encadré 6-4) qui ne peuvent être utilisés que pour le partenaire auquel ils sont destinés ; la défaillance du partenaire les rendrait inutilisables ; un transporteur routier qui équipe ses camions pour le transport combiné devient tributaire des décisions de la SNCF concernant les services ferroviaires de transport combiné qu'elle offre à ses clients : un exploitant de Système Informatisé de Réserveation dépend de la Compagnie aérienne pour laquelle il opère... Des contrats de longue durée sont nécessaires ; mais, à partir d'un certain degré de complexité de ces contrats (longue durée et/ou aléas), la fusion présente des avantages.

Coûts de transaction

La notion de coût de transaction a été introduite par Williamson (1975).

Milgrom et Roberts (1992) en fournissent une typologie distinguant :

- les coûts de coordination : transmission de l'information, puis de la décision dans une organisation hiérarchique ; organisation du marché.
- les coûts de motivation, liés aux asymétries d'information, par exemple l'acheteur d'une voiture d'occasion connaît moins bien que le vendeur l'état du véhicule, ou encore le patron ne peut pas mesurer exactement l'effort au travail de son employé ; liés aussi à l'engagement imparfait des parties, et la difficulté pour chacun d'éviter l'opportunisme de l'autre : les contrats sont rarement complets, c'est-à-dire qu'ils ne prévoient pas toutes les circonstances susceptibles de se produire.

Les caractéristiques suivantes déterminent l'ampleur des coûts de transaction (Dang Nguyen, 1995) :

- l'information limitée des décideurs
- la difficulté pour eux de mettre en œuvre des contrats complets prévoyant toutes les éventualités futures,
- la spécificité des transactions
- la rationalité limitée des décideurs
- le comportement opportuniste des parties
- la fréquence des transactions.
- la difficulté de mesurer les résultats.

Investissement spécifique

La spécificité peut être illustrée par le modèle simple suivant, tiré de Tirole (1990) : supposons 2 firmes. La firme n° 1 est en relation avec le marché final et produit un bien dont le prix sur le marché est v . Elle achète un intrant de la firme n° 2, à un prix $p(I)$, I étant l'investissement spécifique réalisé par cette firme, dont le coût de production est $c(I)$ (on suppose $C'(I) < 0$ et $C''(I) > 0$). Alors l'optimum de I est tel que :

$$\max_I (v - C(I) - I) \quad \text{soit} \quad C'(I) = -1$$

En revanche, la négociation du prix entre les deux firmes si elle s'effectue une fois les investissements faits et sur la base d'un partage des profits à partir de la situation « sans échange » aboutit à un prix $p(I)$ tel que :

$$v - p(I) = p(I) - C(I) \quad \text{soit} \quad p(I) = (v + C(I))/2$$

Alors la firme 2 fixe son investissement de façon que :

$$\max_I (p(I) - C(I) - I) = \max_I \left(v - \frac{C(I)}{2} - I \right) \quad \text{soit} \quad C'(I) = -2$$

La valeur de I est sous-optimale, en raison de la crainte d'opportunisme de la part de la firme 1.

2. Entre monopole et concurrence

2.1. Le champs du monopole

Le monopole se rencontre normalement en présence d'économies d'échelle et d'envergure. C'est ce qu'on appelle la situation de monopole naturel, et alors le monopole est à la fois la structure de production la plus efficace, et celle vers laquelle tend le mouvement naturel : en effet, si au départ il y a deux entreprises, la plus importante a des coûts plus faibles et prend le pas sur l'autre. Comme on l'a vu au chapitre 5, les conditions du monopole naturel, économies d'échelle et économies d'envergure, se rencontrent essentiellement dans les infrastructures ; elles se caractérisent par des coûts fixes élevés et des coûts de fonctionnement faibles. D'autres raisons justifient dans ce cas des infrastructures l'attribution de droits spéciaux à l'opérateur : la nécessité de prérogatives de puissance publique pour l'achat des terrains lors de la construction, les pouvoirs de police nécessaires au maintien de la sécurité et au bon écoulement du trafic¹.

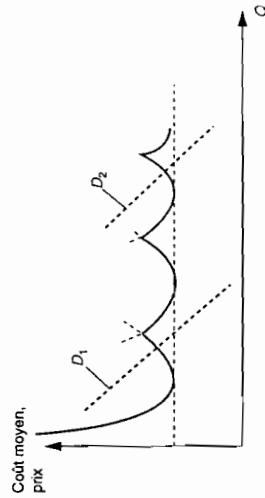
ENCADRÉ 6-5

Structure de marché et volume de la demande

On peut illustrer les liens entre structure de marché et volume de la demande dans le cas d'une production unique. Soit une activité assurée par une ou plusieurs entreprises ayant toutes la même fonction de coût de forme classique en $U : C(q)$, minimale pour q_0^* . Selon le niveau de la demande, il y aura, pour une production efficace, n entreprises sur le marché, n étant défini par :

$$\min_n C(Q/n)$$

On aboutit au graphique joint, dans lequel la fonction de coût de l'industrie en cause a la forme d'une courbe en festons d'égales largeurs et de plus en plus plats au fur et à mesure que la quantité à produire augmente.



1. Dans ces deux cas, la coordination par le marché impliquerait des coûts, des délais, et des risques de non-optimalité considérables par rapport à la délégation des pouvoirs de décision à une autorité.

Si la demande a le niveau D_1 , la situation la plus efficace sur le plan collectif sera le monopole. C'est ce que l'on appelle le monopole naturel. Pour le niveau D_2 , on est en présence d'un oligopole (dans le cas de la figure, 3 entreprises). Enfin, au fur et à mesure que la demande augmente, on se rapproche de la situation de concurrence parfaite, où les coûts sont indépendants du volume de production et où les rendements sont constants (courbe de coût moyen plate). Si la courbe de coût $C(Q)$ présente un méplat au voisinage du point de coût minimum, alors la hauteur des festons ira en diminuant pour finir par disparaître.

Les économies d'échelle et d'envergure se rencontrent aussi dans l'exploitation des transports. Mais elles y ont une beaucoup plus faible extension que dans les infrastructures.

2.2. L'étendue et les modalités de la concurrence

Il convient d'abord de remarquer que la présence de monopole n'empêche pas une certaine forme de concurrence. Les ports d'une même façade maritime sont en compétition, et cela d'autant qu'ils sont proches. Le Havre et Dunkerque subissent la concurrence de Anvers et Rotterdam, il y a de même une compétition entre les aéroports européens pour les trafics intercontinentaux.

Il convient aussi de distinguer entre monopole du produit et du service. Ainsi la SNCF a le monopole du produit ferroviaire ; sur les services qu'elle vend elle est concurrencée par les autres modes, la route et l'avion en particulier.

ENCADRÉ 6-6

La notion de marché pertinent

La notion de marché pertinent est utilisée dans le droit de la concurrence pour évaluer le pouvoir de marché d'une entreprise. C'est la zone de concurrence sur laquelle des entreprises se rencontrent. La façon la plus simple de définir l'étendue d'un marché est de raisonner en terme de substituableité des produits : deux produits feront partie d'un même marché si leurs élasticités croisées dépassent un certain seuil.

Ce critère est bien adapté pour déterminer si deux produits bien identifiés font partie d'un même marché. Il l'est moins quand on l'applique aux entreprises, puisque dans un marché de concurrence parfaite, l'élasticité croisée d'une entreprise aux prix d'une autre est nulle. On peut à la place demander à chaque entreprise de définir les concurrents auxquels elle se heurte, en gardant à l'esprit que les réponses risquent d'être biaisées pour des raisons stratégiques.

On ne peut donc pas juger du degré de concurrence en fonction simplement du nombre d'opérateurs présents sur une aire géographique donnée. Il faut analyser plus finement les conditions de l'offre et celles de la demande pour préciser la notion de marché et pouvoir juger de sa structure, il faut appliquer au cas des transports la notion de marché pertinent utilisée en droit de la concurrence (encadré 6-6).

Dans un sens étroit un produit de l'activité transport et un marché au sens élémentaire seront définis par le transport d'une personne - ou d'un objet - d'une origine à une destination donnée dans des conditions données de coût, de délais, de sécurité et de confort. On voit facilement que changer une de ces caractéristiques entrant dans cette définition change la valeur du bien pour le client et change donc le produit et le marché. Parmi ces caractéristiques certaines sont fondamentales ; ce sont celles qui concernent l'origine et la destination, pour lesquelles les possibilités de substitution sont limitées. Les caractéristiques de qualité de service sont plus substituables. Les substitutions ont été illustrées par le tableau des élasticités croisées entre modes, fourni au chapitre I, qui fait apparaître la concurrence que se livrent le transport aérien et le transport ferroviaire à grande vitesse pour les longues distances, ainsi que la concurrence entre l'automobile et le train normal pour les plus courtes distances.

Une autre forme d'interaction entre marchés élémentaires résulte des conditions de l'offre. Ainsi, les marchés de transport Milan-Amsterdam et Francfort-Madrid sont sans lien au regard de la demande, mais peuvent être reliés au regard des coûts, à travers par exemple les économies d'entrevures : le fait d'exploiter les deux lignes ensemble permet une réduction de la dépense totale, résultant d'un meilleur amortissement des pointes de trafic ou d'une gestion plus facile des incidents et des pannes dont on peut atténuer les effets sur un matériel plus important. De même, pour certains modes, les moyens de transport peuvent être déplacés d'un marché où ils sont excédentaires vers un autre marché où ils sont déficitaires, contribuant à une égalisation des coûts entre ces marchés, aboutissant à un résultat analogue à celui de la concurrence parfaite, même si le nombre d'entreprises opérant sur chaque marché élémentaire est très réduit. Ce phénomène se rencontre dans les transports maritimes, aériens et routiers. Il est plus faible pour le transport ferroviaire, dans la mesure où le déplacement du matériel est coûteux et où le matériel est adapté à l'infrastructure et à l'exploitation ; les déplacements d'un pays à l'autre ne sont possibles que si les normes sont compatibles ; et même en France, les rames Eurostar ne pourraient pas fonctionner entre Paris et Lyon.

En matière de transport routier de marchandises, de nombreux marchés locaux sont dans des situations d'oligopole, dans la mesure où un nombre limité d'exploitants opèrent, souvent sous la forme de relations contractuelles poursuivies avec le même client chargeur qui utilise régulièrement un ou deux, ou deux ou trois transporteurs. Mais cette situation n'entraîne pas les conséquences de l'oligopole classique, parce que, grâce à la mobilité du capital en véhicules, un abus de position dominante des

transporteurs en place entraînerait l'entrée sur le marché élémentaire correspondant de nouveaux transporteurs géographiquement voisins.

Ce phénomène de fluidité de l'offre joue à plein dans le transport maritime de vrac où les possibilités d'entrée sur le marché sont totales. Pour les lignes régulières, il est limité par les Conférences Maritimes, alliances de plusieurs compagnies sur les prix, les fréquences et les partages de profits pour assurer un service régulier sur une ligne donnée. Mais ces accords, qui ne sont protégés par aucune autorité publique, sont exposés, en cas d'excès, à l'entrée d'un outsider.

La fluidité de l'offre et l'unification des marchés à laquelle elle conduit jouent très diversement pour le transport aérien, et dépendent des possibilités légales d'entrées sur le marché. On sait, en effet, que le transport aérien international est régi par des accords bilatéraux qui donnent une extension plus ou moins aux possibilités d'activités commerciales des compagnies des pays en cause. Selon le contenu des accords bilatéraux, les possibilités d'entrées sur chaque marché élémentaire sont plus ou moins fortes.

ENCADRÉ 6-7 Les libertés de transport aérien

La convention de Chicago, qui date de 1944, établit les principes régissant le droit public aérien, fondé sur la souveraineté des États, et définit les « libertés » que ceux-ci peuvent s'accorder bilatéralement :

- 1^{re} liberté : survol sans possibilité d'atterrissage.
- 2^e liberté : possibilité d'atterrissage pour des raisons non commerciales.
- 3^e liberté : débarquement de passagers à partir d'un avion étranger.
- 4^e liberté : embarquement de passagers sur un avion étranger.
- 5^e liberté : même chose que pour les 3^e et 4^e libertés, mais à partir de et vers un pays tiers.
- 6^e liberté : combinant les 3^e et 4^e libertés, réalise la 5^e liberté avec une escale.
- 7^e liberté : cabotage dans un pays tiers.

On voit sur cet exemple apparaître le poids de l'intervention publique qui, par son pouvoir législatif et réglementaire, peut modifier le jeu des forces naturelles, et en particulier créer ou défaire des monopoles, réduire ou augmenter le nombre d'opérateurs et le degré de leur compétition. Ces dernières considérations amènent à se pencher sur les interventions de la puissance publique dans les marchés des transports, leurs raisons d'être et leurs limites.

3. Entre intervention et libéralisme

La politique des transports est une perpétuelle oscillation entre libéralisme et interventionnisme. Pour prendre le cas de la France, sa position a été libérale depuis le milieu du 19^e siècle jusqu'à l'entre-deux guerres, avec

notamment le développement des chemins de fer à travers des sociétés concessionnaires privées, puis une grande liberté laissée à l'initiative privée au début de l'automobile. L'intervention publique apparut lorsque le développement du transport routier mit en péril les chemins de fer en réduisant leur clientèle et en creusant leur déficit. L'aboutissement fut la fusion des compagnies de chemin de fer en 1937, puis la nationalisation de la SNCF, et, après la guerre, un ensemble de contrôles et de règlements visant à freiner le développement du transport routier. La politique générale d'inspiration interventionniste de la Libération a conduit à la création des autres entreprises publiques du secteur : Air France, Air-Inter, RATP, CGM.

Mais, au fil des années, les entreprises publiques rencontraient des difficultés croissantes pour équilibrer leurs budgets ; les mesures pour enrayer le développement du transport routier n'atteignaient pas leur but. Dans les années 80, les idées libérales reprirent le dessus avec la déréglementation américaine, le tachtérisme anglais, et les orientations nouvelles de la Commission Européenne. La France subit avec un certain retard cette vague de libéralisme, qui toucha d'abord le transport routier de marchandises, puis le transport aérien, et qui maintenant ébranle l'édifice de la SNCF.

À ces variations dans le temps font écho des variations selon les pays. En Europe, la France est plus interventionniste que les pays du Nord, et en particulier le Royaume-Uni où les chemins de fer, préalablement fragilisés en une centaine d'organismes ont progressivement été privatisés, suivant ainsi la voie imposée aux compagnies aériennes et aux aéroports, tandis que le marché était presque totalement libéré pour les transports routiers de voyageurs, tant urbains qu'interurbains.

ENCADRÉ 6-8

Les modalités d'intervention publique

Outre la gestion directe en régie ou par l'intermédiaire d'un service public (établissement public, entreprise publique), la puissance publique peut agir à travers une concession concernant la construction et l'exploitation d'un ouvrage (cas des sociétés d'autoroutes), ou simplement l'exploitation d'un réseau (cas des transports collectifs urbains), la concession pouvant être accordée soit à une entreprise publique (société d'économie mixte par exemple) soit à une entreprise privée.

La puissance publique peut aussi agir à l'égard des entreprises privées à travers la réglementation qu'on analyse en trois catégories :

- la réglementation technique, celle qui touche les règles de construction et de circulation des véhicules,
- la réglementation sociale, qui touche les conditions de travail,
- la réglementation économique, qui concerne les prix (fixation des tarifs, établissement de taxes) ou les quantités (autorisation nécessaire pour entrer sur le marché, limitation des capacités de production).

Si des raisons historiques ou géographiques expliquent les fluctuations dans l'intervention publique, des raisons plus fondamentales justifient sa permanence. Ces raisons peuvent se regrouper sous deux têtes de chapitre : les défaillances du marché et les missions de service public.

3.1. Les défaillances du marché

L'État doit bien sûr d'abord veiller à un fonctionnement correct des marchés, et pour cela, par exemple, contrôler le niveau de sécurité des moyens de transport, que les usagers n'ont pas les moyens de vérifier. Il doit aussi pallier les défaillances du marché, particulièrement présentes dans les transports où elles résultent de l'existence de caractéristiques qui éloignent le marché de l'optimum économique, essentiellement les externalités et les rendements croissants¹.

ENCADRÉ 6-9

Externalités, monopole et optimum économique

Les relations entre optimum économique et économie de marché sont réglées par les deux théorèmes fondamentaux :

« Un équilibre concurrentiel, où chaque agent détermine son comportement en prenant le prix comme des données, est un optimum de Pareto. »
 et « Si l'ensemble de production est convexe (pas de rendement croissant) un optimum de Pareto peut être obtenu par un système de marché. »

Ces deux théorèmes établissent l'optimalité du système de marché sous certaines hypothèses (Laffont, 1982) qui, en dehors d'hypothèses techniques, peuvent s'exprimer ainsi : les fonctions d'utilité des consommateurs ne dépendent que des biens qu'ils consomment ; la fonction de production de chaque entreprise ne dépend que des inputs de l'entreprise et est à rendement constant. Ces conditions excluent les effets externes et les comportements de monopoles. On peut analyser les conséquences distorsives par rapport à l'optimum de ces deux dernières caractéristiques, dans le cadre d'une analyse partielle.

Effets externes

Sur le graphique ci-après sont figurées :

- la courbe de demande D ,
- la courbe d'offre qui, dans le cas d'un bien produit sur un marché concurrentiel, définit le coût marginal de production : Cm privé,
- la courbe de coût collectif Cm collectif, qui dans le cas d'un effet externe (par exemple nuisance liée à l'activité) diffère du coût marginal privé.

1. Certaines activités du secteur des transports peuvent aussi s'analyser à travers le concept de bien public, cas limité d'une activité à coût fixe uniquement, le coût marginal étant nul : un bien public est tel que sa consommation par un usager ne réduit pas la consommation des autres usagers ; l'utilisation d'infrastructures non saturées peut s'analyser comme la fourniture d'un bien public, au même titre qu'une émission de télévision ou un parc de loisir, tant qu'il ne reçoit pas un nombre élevé de visiteurs.

Quant aux monopoles, ils sont sources d'une double inefficacité. L'inefficacité statique (voir l'encadré 6-9) résulte de ce que le monopole tarifie au-dessus de l'optimum et réduit les quantités produites. L'inefficacité dynamique résulte de la moindre pression qui s'exerce sur l'entreprise en situation de monopole. De nombreuses analyses statistiques les mettent en évidence, au premier rang desquelles figurent les observations faites à l'occasion de la déréglementation, qui a conduit à supprimer les droits exclusifs dans certaines activités et à y remplacer le monopole par la concurrence.

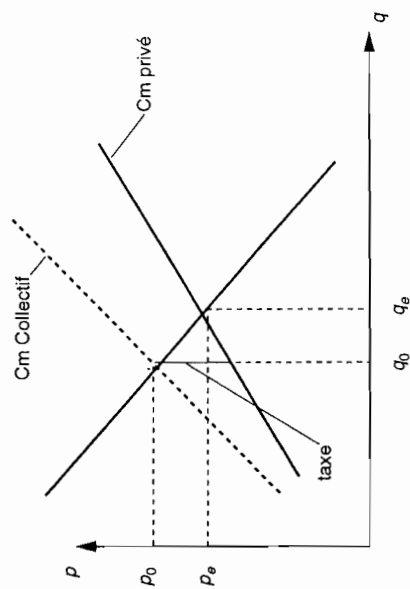
3.2. Les missions de service public

L'autre source d'intervention publique est l'exécution de missions de service public. La notion de service public est ancrée dans le droit français qui lui donne une place extensive, et lui reconnaît traditionnellement plusieurs attributs : la continuité, l'universalité, l'adaptabilité et l'égalité.

Bien que ces principes se retrouvent plus ou moins dans les droits de la plupart des pays européens, le traité de l'Union ne leur accorde guère de place. Ce dernier a pour but essentiel de fixer les règles permettant à la concurrence de s'exercer sagement sur l'ensemble du territoire européen. De nombreux articles traitent des situations de position dominante (article 86), ou des actions publiques, notamment les subventions, susceptibles de fausser le jeu de la concurrence (article 92). Seul l'article 90 mentionne « ... les entreprises publiques et les entreprises auxquelles ils (les États) accordent des droits spéciaux... », et c'est pour restreindre leurs droits, qui ne doivent pas aller à l'encontre des règles de saine concurrence et de la libre circulation des biens et services « ... dans les limites où l'application de ces règles ne fait pas échec à l'accomplissement en droit ou en fait de la mission particulière qui leur a été impartie... ». Une particularité juridique importante : c'est la Commission, c'est-à-dire l'administration bruxelloise et la sorte de Gouvernement européen non élu que sont les Commissaires, qui dit le droit, et a donc un pouvoir sur les États en la matière, ce pouvoir n'étant limité que par la jurisprudence issue des arrêts de la Cour Européenne concernant les affaires dont elle est saisie.

Quels sont les activités susceptibles d'être qualifiées de « services d'intérêt économique général » ? Le traité n'est guère explicite, mais la jurisprudence l'est un peu plus. Dans les arrêts Corbeau et Almelo, la Cour de Justice a reconnu le caractère de service d'intérêt économique général au service des postes belge et à la distribution régionale d'électricité, se fondant sur les caractéristiques des missions qui leur sont confiées : obligation de desserte, continuité du service, uniformité du prix, le tout étant assemblé sous le terme de service universel.

La jurisprudence communautaire a une vue limitative des restrictions de concurrence permises pour l'exécution de la mission de service public (Stoffaers, 1995) : elles doivent s'avérer « ... nécessaires pour permettre à l'entreprise investie d'une telle mission d'intérêt général d'accomplir



L'équilibre naturel du marché se fixe en q_e, P_e , alors que l'optimum est en q_0, P_0 ... On peut restaurer ce dernier en instituant une taxe t à la production.

Comportement de monopole

Si le bien est produit par un monopole qui cherche à maximiser son profit, son comportement sera déterminé par :

$$\max_q \Pi(q) = \max_q (p(q)q - C(q))$$

expression dans laquelle :

Π est le profit du monopole

$p(q)$ est la courbe inverse de demande

$C(q)$ est la fonction de coût

on en tire, en se limitant aux conditions du premier ordre :

$$p = dC/dq - q dp/dq$$

Alors que l'optimum, défini par la maximisation du surplus, est réalisé par :

$$\max_q \left(\int_0^q p(x) dx - C(q) \right) \quad \text{soit} \quad p = dC/dq$$

Le monopole fixe un prix supérieur au coût marginal, au détriment du consommateur, qui paie plus cher des quantités plus faibles. Si le monopole suivait la règle de vente au coût marginal, l'optimum serait assuré.

Pour pallier les effets négatifs qu'ont les externalités sur le bien-être collectif et internaliser leurs conséquences pour ceux qui les provoquent, la puissance publique dispose d'un certain nombre de moyens : la taxation (ou subvention dans le cas d'une externalité positive), la réglementation ou l'instauration de marchés de droits.

celle-ci... » (Arrêt Almelo). Ainsi la protection d'activités rémunératrices permettant de compenser les pertes résultant de l'exécution du service public est admise, sauf si les activités rémunératrices en cause sont «... des services spécifiques dissociables du service d'intérêt général qui répondent à des besoins particuliers d'opérateurs économiques et qui exigent certaines prestations supplémentaires que le service postal traditionnel n'offre pas... » (Arrêt Corbeau).

Le choc entre l'interprétation restrictive communautaire de la notion de service public et la place extensive qu'elle prend en France a conduit à développer les réflexions sur ces sujets en France. Une mission confiée à Stoffaers (1995) et une autre confiée à Denoix de Saint-Marc (1996) ont contribué à dégager les deux principes qui définissent le service public (Henry, 1997a) : son caractère essentiel pour la vie économique et sociale du pays, et le remède aux défaillances du marché. Sur cette base apparaissent trois catégories fondamentales de services publics (Denoix de Saint-Marc, 1996 et Henry, 1997b) :

- ceux qui rendent accessibles des services essentiels aux usagers menacés d'exclusion,
- ceux qui contribuent à la cohésion sociale et au sentiment d'appartenance à une communauté,
- ceux qui favorisent une bonne utilisation, dans l'espace et le temps, des ressources communes.

Le service public trouve de nombreuses raisons de s'exercer dans les transports. Leur présence et leur continuité sont évidemment nécessaires au bon fonctionnement économique et social, comme on le constate lorsque par suite de grèves ou d'accidents, ils sont interrompus. D'ailleurs la loi d'orientation des transports intérieurs du 31 décembre 1982 (LOTI) a inscrit dans son texte le droit au transport dans des conditions acceptables de coût. Il est alors nécessaire que la puissance publique intervienne lorsque ce droit au transport n'est pas assuré par les mécanismes naturels du marché ou lorsque, par la présence d'un monopole naturel, son exercice dépend d'une entreprise unique dont les décisions pourraient le mettre en péril.

Cette nécessité se présente par exemple lorsque le prix assurant le droit au transport est inférieur au coût ; elle se présente aussi lorsque le service universel, ou le service public dans la conception française, conduit au mécanisme des subventions croisées, par lesquels les bénéfices d'une partie du service (par exemple les lignes les plus chargées) financent les déficits des lignes non rentables (par exemple les lignes suburbaines des transports urbains, ou les dessertes rurales en interurbain), permettant de pratiquer un prix égal pour tous sans faire appel à des subventions extérieures, et malgré l'inégalité des coûts. Le monopole est dit alors non soutenable, car un concurrent pourrait éliminer le marché, en offrant un prix un peu inférieur pour les lignes rentables, et détruire ainsi le mécanisme de compensation assurant le financement du service universel (cf. encadré 6-10).

Soutenabilité, subvention croisées et écrémage

Soit une entreprise assurant une production multiple définie par un vecteur q , selon une fonction de coût $C(q)$. Il y a subvention croisée entre deux groupes de produits q_1 et q_2 vérifiant : $q = q_1 + q_2$ et commercialisés avec des vecteurs de prix p_1 et p_2 , si les conditions suivantes ne sont pas satisfaites :

$$C(q) - C(q_2) \leq p_1 q_1 \leq C(q_1)$$

$$C(q) - C(q_1) \leq p_2 q_2 \leq C(q_2)$$

on vérifie que si : $p_1 q_1 + p_2 q_2 = C(q)$ deux quelconques d'entre elles résultent des deux autres.

Alors si par exemple le premier couple d'inégalités n'est pas satisfait, un concurrent peut fabriquer q_1 seulement, et le vendre à un prix un peu inférieur à p_1 , et ainsi écrémer les marchés des biens q_1 en éliminant l'entreprise en place.

La pratique de la subvention croisée fait l'objet de critiques : elle éloigne les prix des coûts et aboutit donc à une perte économique. Elle conduit souvent à jeter le brouillard sur les coûts et peut ainsi entraîner des erreurs supplémentaires d'allocations de moyens et de fournitures de biens et services. Son avantage est qu'elle permet de financer des obligations de service public sans avoir recours à l'impôt.

La subvention croisée est une des sources d'insoutenabilité, une notion qui se définit ainsi :

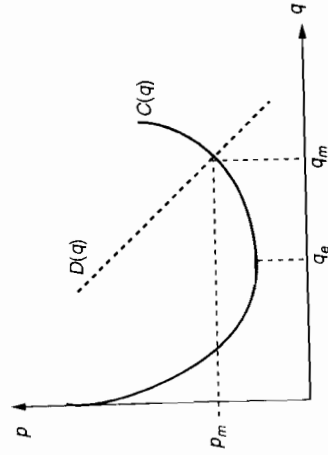
« Si la firme en place produit q (scalaire ou vecteur) et le vend au prix p vérifiant :

$$q = D(p) \quad \text{et} \quad pq \geq C(q)$$

Il existe un prix p_e et une quantité q vérifiant :

$$p \geq p_e \quad \text{et} \quad D(p_e) \geq q_e \quad \text{ainsi que} \quad p_e q_e \geq C(q_e)$$

Un cas d'insoutenabilité structurelle est celui décrit par la figure jointe, où sont représentées les courbes de demande D et de coût moyen, et où le monopole est astreint à l'équilibre budgétaire : la situation (q_m, P_m) est optimale sous la contrainte imposée, mais elle n'est pas soutenable : un concurrent produisant q_e à un prix légèrement inférieur à P_m détruirait l'équilibre du monopole.



d'intérêt général, ce qui se produit essentiellement dans la gestion des infrastructures. Le chapitre VIII analyse les formes de la concurrence et les effets des réglementations qui lui sont usuellement appliquées. Enfin le chapitre IX présentera les situations, en général associées, de monopole et d'exécution de service public.

BIBLIOGRAPHIE

- N. Curien et G. Dupuy (1996), *Réseaux de communication, marché et territoire*, Presses des Ponts et Chaussées.
- G. Dang N'Guyen (1995), *Économie industrielle appliquée*, Vuibert.
- R. Denoix de Saint-Marc (1996), *Rapport à Monsieur le Premier Ministre sur le Service Public*, La Documentation française.
- C. Henry (1997a), *Cours d'économie publique*, vol. 2, École Polytechnique.
- C. Henry (1997b), *Concurrence et service publique dans l'Union Européenne*, PUF, Paris.
- J.J. Laffont (1982), *Cours de théorie micro-économique*, Economica.
- F. Levêque (1997), *L'économie de la réglementation*, Collection Repères, La Découverte.
- P. Milgrom et J. Roberts (1992), *Economics, organisation and management*, Prentice-Hall International Éditions.
- E. Quinet (1993), « Transport between monopoly and competition », in J. Polak and A. Heertje (1993), *European Transport Economics*, Blackwell.
- C. Stoffaes (1995), *L'Europe de l'utilité publique*, ASPE Europe Éditions.
- J. Tirole (1990), *The theory of industrial organization*, The MIT Press.
- O.E. Williamson (1975), *Markets and Hierarchies : Analysis and Antitrust implications*, The Free Press, New-York.

Ainsi, lorsque les autorités anglaises ont aboli les monopoles d'exploitation des bus dans les zones urbaines et entre villes, on a vu disparaître l'exploitation des lignes non rentables, qui n'ont pu être rétablies que par des concessions subventionnées : la suppression des monopoles de droit a fait disparaître les subventions croisées et a conduit à substituer un financement par l'impôt à un financement par les usagers des autres lignes. Dans cet exemple le monopole était une des modalités d'exécution du service public, mais ce n'était pas la seule : une organisation plus complexe composée de la concurrence sur les lignes rentables, et de concessions sur celles qui ne l'étaient pas, atteignait le même but.

3.3. Les limites à l'intervention publique

Ces limites sont de deux ordres.

Elles tiennent d'abord aux vertus de la concurrence, vantée par Schumpeter comme processus de destruction créatrice. La concurrence pousse à réduire les coûts, à adapter les produits aux désirs et besoins de la clientèle, à développer le progrès : les opérateurs, voyant leurs profits érodés par leurs concurrents, cherchent à les rétablir par un processus dynamique d'innovation et d'adaptation.

L'autre limite réside dans une analyse critique de l'action publique, dans une reconnaissance des défaillances du Gouvernement. La puissance publique n'est pas l'entité à la recherche de l'intérêt collectif que suppose la théorie. Elle est en fait composée d'une multiplicité d'acteurs qui sont aussi animés par des objectifs propres : les organisations administratives cherchent à développer leur pouvoir, l'importance de leur service en nombre de personnes et en volume financier à leur disposition. Les hommes politiques sont contraints par leurs soucis de réélection, et soumis à l'agenda électoral. Le contrôle du Parlement sur les actions administratives est très imparfait. De toute façon, la puissance publique est victime d'une asymétrie d'information en sa défaveur, les entreprises à contrôler connaissant bien mieux que leur tutelle les conditions de l'offre et de la demande qui règnent dans leur marché. Enfin, le réglementé fera tout son possible pour retourner à son profit le réglementeur par des actions pouvant aller de la diffusion d'informations adaptées et soigneusement filtrées jusqu'aux pratiques usuelles de lobbying : c'est le mécanisme connu sous le nom de capture du régulateur.

4. Conclusion

Les considérations qui précèdent font comprendre les raisons de la diversité des structures des marchés dans le secteur. Elles vont être développées dans les chapitres qui suivent. Le chapitre VII relatif aux décisions publique optimales traitera du comportement de la puissance publique lorsqu'elle est opérateur direct et qu'elle poursuit pleinement ses objectifs

Décisions publiques optimales

On prend ici le point de vue de pouvoirs publics qui, en présence de consommateurs et de producteurs dont l'objectif est de maximiser leur utilité ou leur profit, cherchent à satisfaire à l'intérêt collectif à partir des moyens à leur disposition : l'information, la réglementation, la fixation de taxes et tarifs, et la production directe de biens, cette dernière action s'exerçant surtout en matière d'infrastructures. On va successivement rap- peler les modalités d'expression de l'intérêt collectif et les critères per- mettant de le mesurer (section 1) puis l'application de ces critères au domaine où se manifeste le plus fortement l'intervention de la puissance publique, les infrastructures ; la section 2 concernera essentiellement leur exploitation (tarification, réglementation) ; la section 3 sera consacrée à leur construction et traitera des problèmes de choix des investissements.

1. Définition et mesure de l'intérêt collectif

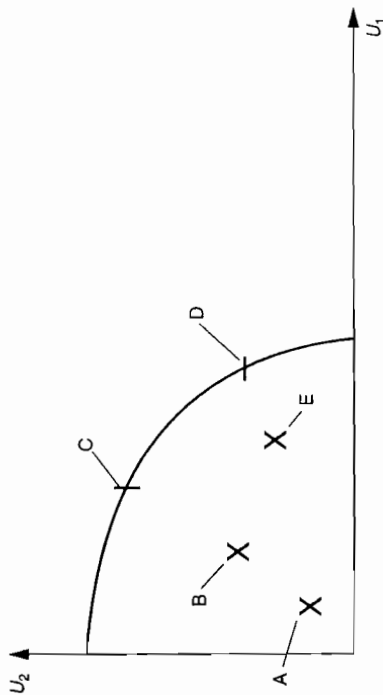
Lorsque la puissance publique met en œuvre une décision, celle-ci se traduit par une modification de l'état d'équilibre du système économique, entraînant des changements dans les prix des biens et dans les quantités consommées de ces biens. Le mécanisme de l'équilibre général walrasien met en évidence l'interdépendance de tous les marchés, et apprend que des modifications initiales, mêmes si elles ne concernent qu'un marché, entraînent des transformations des prix et des quantités échangées sur tous les marchés.

La théorie du comportement du consommateur permet d'apprécier l'effet d'une transformation des prix et quantités sur l'utilité de chacun des individus qui composent la société considérée. Elle s'exprime par la varia- tion d'utilité U_i de l'individu i entre l'état initial (indice 0) et l'état final (indice 1) les quantités consommées étant notées q :

$$\Delta U_i = U_i(q_i^1, \dots, q_{ni}^1) - U_i(q_i^0, \dots, q_{ni}^0) = U_i^1 - U_i^0$$

On peut ainsi, ne serait-ce que par la pensée, déterminer pour une décision envisagée, les variations d'utilité qui en résulteront pour chaque individu de la société. Faisons-le dans le cas où la société comporte deux individus, en traçant un graphique où figurent en abscisse l'utilité de Monsieur 1 et en ordonnées l'utilité de Monsieur 2. Notons d'abord que, en raison de la limitation des ressources, la zone des possibles est limitée, par exemple comme l'indique la figure 7-1, sous la courbe.

FIGURE 7-1
Optimum de Pareto



Le point B est visiblement supérieur au point A, parce que tous les membres (ici les deux membres) de la société y gagnent ; et de même C est supérieur à B. De même D est préférable à E, mais on ne peut pas comparer E et B ou E et C ; on sait simplement que E et B peuvent être améliorés pour tous les membres de la collectivité. Mais il n'y a aucune situation qui soit pour tout le monde préférable à C. On ne peut pas, sur la base du critère précédent, choisir entre C et D. C est un optimum de Pareto comme D, et il y en a ici une infinité, définis par la frontière entre le possible et l'irréalisable.

Le critère de Pareto ne permet donc pas de comparer toutes les situations : seules celles unanimement préférées sont comparables. Mais les actions où tous sont bénéficiaires sont rares dans les sociétés composées d'un grand nombre d'individus ; des comparaisons entre situations comme B et E, où certains perdent et d'autres gagnent, sont souvent nécessaires ; et de toute façon il faut bien se décider entre deux situations d'optimums paretiens comme C ou D.

Cette analyse fait apparaître plusieurs problèmes :

- comment pour un même individu, calculer la variation d'utilité entre deux états ?
- comment intégrer dans un critère global, les variations d'utilité des différents membres de la collectivité ?

1.1. Le calcul des variations d'utilité individuelles

Si la transformation est marginale, la variation d'utilité peut être exprimée sous diverses formes, issues des relations fondamentales de la théorie du consommateur présentées à l'annexe 3, et qu'on résumera :

Si le consommateur *i* est muni d'une fonction d'utilité :

$$U_i(q_{i1}, q_{i2}, \dots, q_{in})$$

et d'un revenu maximum à ne pas dépasser r_i , son comportement sera dicté par le programme suivant :

$$\max_{q_{i1}, q_{i2}, \dots, q_{in}} U_i$$

tel que :

$$\sum_j p_j q_{ij} \leq r_i$$

La dérivation du lagrangien (Annexe I) :

$$L = U_i - \lambda_i \left(\sum_j p_j q_{ij} - r_i \right)$$

donne :

$$\frac{\partial U_i}{\partial q_{ij}} = \lambda_i p_j$$

Donc :

$$dU_i = \sum_{j=1}^n \frac{\partial U_i}{\partial q_{ij}} dq_{ij} = \lambda_i \sum_{j=1}^n p_j dq_{ij}$$

ou, puisque :

$$\sum_j (p_j dq_{ij} + q_{ij} dp_j) = dr_i$$

$$dU_i = \lambda_i \left(dr_i - \sum_j q_{ij} dp_j \right) \quad (7-1)$$

Le calcul des dU_i passe par la connaissance de l'utilité marginale du revenu λ_i . Ce problème peut être tourné, puisque les utilités ne sont définies qu'à une fonction monotone près, et donc que leur variation marginale ne l'est qu'à un facteur positif multiplicatif près. On peut donc choisir pour ce facteur la valeur unité. On a alors l'expression :

$$dU_i = \sum_j p_j dq_{ij} \quad (7-2)$$

Si seulement un prix varie, les autres prix et le revenu ne changeant pas, on aboutit à :

$$dU = -q_{ij} dp_j$$

En normalisant les utilités individuelles selon la convention classique qui fixe à l'unité les utilités marginales de la monnaie des différents individus, on voit que :

$$dW = \sum_i \left[F_i \sum_j (p_j dq_j) \right]$$

Si la répartition initiale des revenus est optimale, alors les poids F_i attribués aux individus sont égaux, ce qui entraîne que les F_i sont indépendants de i , peuvent être pris égaux à 1 et :

$$dW = \sum_i p_i dq_i$$

dq_j étant la variation totale de quantité consommée du bien j .

Si la collectivité n'accorde pas le même poids aux individus qui la composent, si elle veut favoriser certains groupes, alors cela pourra se traduire par des coefficients différents selon les individus.

1.3. Problèmes pratiques de calcul des surplus

Le calcul des surplus au sens de Dupuit ou de variation équivalente est *a priori* très exigeant en informations : à la suite de l'action que l'on envisage, tout le système des prix et des revenus va être modifié, donc toutes les quantités de tous les biens consommés par tous les individus ; il faudrait calculer d'abord toutes ces variations ensuite les surplus de Dupuit de chaque individu de la collectivité en cause.

Cette tâche écrasante peut être simplifiée dans deux situations :

- Si on peut agréger les fonctions de demandes des individus en celle d'un individu représentatif. Cela se produit dans deux situations : la première évidente se présente lorsque tous les individus ont des fonctions d'utilité et des revenus identiques ; la deuxième beaucoup plus réaliste se présente lorsque les fonctions de demande individuelles sont du type (Varian, 1992) :

$$q_j^i = \alpha_j^i (p_1, \dots, p_n) + r^i \beta_j (p_1, \dots, p_n)$$

Alors la fonction de demande agrégée s'écrit, en posant : $R = \sum p^i$

$$Q_j = \sum_i q_j^i = \sum_i \alpha_j^i (p_1, \dots, p_n) + R \beta_j (p_1, \dots, p_n)$$

Pour un bien donné, on n'a à considérer qu'une seule fonction de demande, au lieu d'avoir à prendre en compte autant de fonctions de demande qu'il y a d'individus.

- Dans le cas de transformation marginales, si le système économique est dans une situation d'optimum parétien, dans lequel les entreprises maximisent leur bénéfice à prix constants, alors une mesure (marginale bien sûr) prise sur un bien particulier peut être évaluée en analysant uni-

L'annexe 5 analyse quelques uns des problèmes posés par l'extension de ces résultats au cas de transformations non marginales. Avec des hypothèses simplificatrices (par exemple fonction d'utilité quasi-linéaire), et dans le cas où le prix d'un seul bien change, on aboutit à la formule connue qui exprime la variation de l'aire sous la courbe de demande :

$$\Delta U_i = - \int_{p_j^0}^{p_j^1} q_{ij} dp_j \quad (7-3)$$

C'est l'expression classique du surplus de Dupuit, qui permet de comparer et hiérarchiser différentes transformations pour un même individu, à partir des données mesurables que sont les prix et les quantités.

1.2. Comment agréger les utilités individuelles ?

Comment agréger les variations d'utilités individuelles des individus qui composent la société à considérer ? Si ces variations d'utilité entre deux états sont positives pour tous les individus, la transformation est souhaitable. Une difficulté se présente lorsque la transformation améliore la situation de certains et détériore celles d'autres. Le critère admis est le critère de compensation : les gagnants ont de quoi compenser les perdants. Sa mise en œuvre rigoureuse est très délicate dans le cas de transformations finies (voir l'annexe 5).

La situation est plus simple dans le cas des transformations marginales. Les transferts étant marginaux également, leur effet sur le système de prix n'a pas à être pris en compte. Le critère de compensation se réduit alors à :

$$\sum_j p_j dq_j \geq 0$$

dq_j étant la somme algébrique des variations de quantités du bien j consommées par tous les individus.

Ce critère a été présenté dans le cas où la fonction d'utilité est uniquement composée de biens marchands. Il peut être étendu au cas des effets externes et à celui des biens non marchands comme le temps (Morisugi, 1983).

Une autre présentation du critère du surplus dans le cas de transformations marginales consiste à supposer que la collectivité est munie d'une fonction d'utilité collective dépendant des utilités de chacun des individus qui la composent :

$$W = F(U_1, \dots, U_n)$$

Alors la variation de l'intérêt collectif résultant d'une action donnée de caractère marginal est :

$$dW = \sum_i F_i dU^i$$

quement le marché du bien en cause : l'ensemble des effets sur le système économique et des surplus qui s'y produisent peut être mesuré à travers le surplus de Dupuit sur le marché en cause, l'analyse globale peut être résu- mée à travers une analyse partielle, limitée au secteur concerné. La démonstration, adaptée de Lesourne (1967), et reprise dans l'encadré joint sur un exemple particulier, suppose des transformations marginales, le plein emploi des ressources naturelles, et l'équilibre du commerce exté- rieur. Si la transformation n'est pas marginale, le calcul complet portant sur l'ensemble du système devient nécessaire. Kanemoto et Mera (1985) ont déterminé des formules approchées permettant de l'exécuter. Mori- sugi (1997) a effectué des calculs numériques mesurant l'écart entre le sur- plus de Dupuit et la variation équivalente, et montre que cet écart est faible même dans le cas d'actions importantes comme par exemple le pro- gramme de construction des chemins de fer japonais.

ENCADRÉ 7-1

Équivalence entre analyse globale et analyse partielle dans le cas d'une économie gérée à l'optimum, à partir d'un exemple tiré du secteur des transports

Supposons que le transport utilise simplement un bien a de prix p_a , en quantité X_a par km parcouru. La variation de surplus collectif est alors :

$$dU = p_a X_a dT + \sum_i p_i dq_i \quad (1)$$

Dans cette relation :

dq_i représente la variation de consommation du bien i ,

T représente le nombre de km parcourus.

De l'égalité entre l'offre et la demande de chaque bien i , on tire :

$$dq_i = \sum_h dq_{ih} + dq_{ia}$$

Dans cette relation :

l'indice h représente les entreprises,

q_{ia} représente la quantité de bien i consommée par l'entreprise qui fabrique le bien a .

La convention est faite que les variations de quantité consommées sont négatives et produites positives.

On en tire :

$$\sum_i p_i dq_i = \sum_i p_i \sum_h dq_{ih} + \sum_i p_i dq_{ia} \quad (2)$$

Dans l'hypothèse où les entreprises sont gérées à l'optimum, c'est-à-dire où elle recherchent à maximiser leur bénéfice à prix constant, on a :

$$dB_h = \sum_h p_i dq_{ih} = 0 \quad (3)$$

et :

$$\sum_i p_i dq_{ia} + p_a dq_a = 0 \quad (4)$$

Or dq_a , production de l'entreprise a , est égal à :

$$dq_a = d(X_a T) = X_a dT + T dX_a \quad (5)$$

en combinant les relations précédentes, on obtient :

$$dU = -p_a T dX_a$$

C'est-à-dire que le surplus collectif est égal au surplus de l'utilisateur.

On voit en outre, par les relations (2) et (3), que si une entreprise h n'est pas gérée à l'optimum, il convient d'ajouter au surplus de l'utilisateur sa variation de bénéfice à prix constant :

$$dB_h = \sum_i p_i dq_{ih}$$

On peut prendre la mesure de la simplification ainsi permise : suppo- sons qu'on réduise le coût de transport sur une relation particulière, et, sans nous préoccuper pour le moment de la dépense éventuelle nécessaire pour obtenir cette réduction du coût de transport (supposons par exemple que cette réduction de coût est obtenue par une mesure d'organisation à dépense constante), analysons ce qui va se passer au niveau de l'utilisa- tion. Alors tous les prix de tous les biens de l'économie vont changer, ainsi que les quantités consommées par chaque consommateur. Mais le résultat précédent permet de calculer le surplus collectif sans avoir à déterminer cette myriade de chiffres. Il montre en effet que ce surplus collectif est égal au surplus des usagers qui empruntent la liaison, et qui vaut :

$$dU = -T dp_i \quad (7-3)$$

relation où :

T est le trafic de la liaison,

dp_i est la variation de coût de transport sur la liaison.

En effet, pour chaque usager j utilisant la liaison :

$$dU_j = \sum_i p_i dq_{ij} = dr_j - \sum_i q_{ij} dp_i$$

Or leur revenu ne change pas, ni non plus les prix des biens, sauf celui du transport qui varie de dp_i .

On peut maintenant lever l'hypothèse que la mesure s'est effectuée sans coût. Si sa mise en œuvre a nécessité une dépense supplémentaire dC , le même raisonnement montre qu'il suffit de mettre au passif de l'opéra- tion la variation dC qui résume toutes les conséquences de la dépense sup- plémentaire en termes de variations d'utilités des individus. Le surplus collectif devient :

$$dU = -T dp_i - dC$$

Ceci bien sur, sous réserve des conditions de validité du résultat : transformation marginale, économie à l'optimum, possibilité de compensation ou répartition optimale des revenus.

Si l'économie n'est pas gérée à l'optimum, c'est-à-dire si certaines entreprises ont un comportement de monopole maximisant son profit, les résultats présentés dans l'encadré 7-1 montrent qu'il convient d'ajouter au surplus de l'usager les variations de bénéfice à prix constants de ces entreprises.

Si la répartition des revenus n'est pas optimale, ou si les compensations ne sont pas possibles, alors il faudrait déterminer les gains et pertes de chacun et les pondérer par l'importance que leur donne la collectivité.

Remarquons enfin que cette théorie exclut la présence d'externalités. On vient d'examiner les outils théoriques essentiels de la prise de décision collective.

On va maintenant étudier l'utilisation de ces outils dans la gestion des transports, qui se présente surtout là où la puissance publique intervient le directement, c'est-à-dire dans la gestion des infrastructures, à travers les décisions d'exploitation, objets de la section 2 suivante, et à travers les décisions de choix d'investissement qui seront examinées dans la section 3.

2. L'exploitation des infrastructures : tarification, réglementation, marché de droits

On va d'abord présenter les règles de gestion à mettre en œuvre dans une recherche d'optimum premier ; puis on analysera les modifications entraînées dans l'outil principal de gestion, la tarification, par des situations d'optimum second, par exemple lorsque l'opérateur est astreint à l'équilibre financier ou lorsque le reste de l'économie n'est pas géré à l'optimum ; enfin on analysera la pratique de la tarification.

2.1. Règles de gestion optimale

Le problème qu'on se pose ici est celui du niveau optimal de production. Lorsqu'il n'y a pas d'effets externes, ce niveau est obtenu par l'outil tarifaire à travers le principe de la tarification au coût marginal : lorsque le tarif est ainsi fixé, la quantité vendue, qui s'en déduit par la loi de demande, est optimale.

Le principe de la tarification au coût marginal peut être retrouvé à partir de nombreuses présentations. Les plus élaborées font appel aux théorèmes de l'économie du bien qui expriment que, sous des hypothèses dont les plus importantes concernent la convexité des ensembles de production et des fonctions d'utilité, un équilibre concurrentiel est un optimum (cf encadré 6-7). Or, dans l'équilibre concurrentiel, les entreprises, qui maximisent leur profit à prix constants, poussent leur production jusqu'au niveau où le coût marginal est égal au prix du marché.

Une voie simple pour arriver à ce résultat est de recourir, en analyse partielle, à la notion de surplus de Dupuit pour le marché considéré. Celui-ci est égal à l'aire sous la courbe de demande diminuée du coût de production :

$$S(q) = \int_0^q p(t) dt - c(q)$$

Il est maximum pour :

$$p(q) = \frac{dc}{dq}$$

On remarque que ce résultat tient quels que soient les rendements, décroissants ou non, du producteur.

En revanche, il ne tient pas compte de la présence d'effets externes et suppose que l'économie est gérée de façon optimale. Analysons d'abord comment il peut être adapté au cas des effets externes.

2.1.1. Prise en compte des effets externes – Les principes

Les effets externes apparaissent, comme les biens non marchands, dans les fonctions d'utilité, ils ne dépendent pas des décisions de l'individu qui les subit ou en profite, mais de décision d'autrui. Leur présence est un obstacle au bien-être collectif, dans la mesure où celui qui les engendre, insensible aux bénéfices ou gênes qu'il procure à autrui, les exerce trop ou trop peu. Restaurer l'optimum de Pareto revient alors à corriger l'effet du système de prix naturels.

On peut illustrer ces mécanismes par un exemple simple, celui de deux entreprises 1 et 2, opérant chacune sur un marché concurrentiel. La première fabrique un produit x à un coût $C_1(y)$ ($C_1' > 0$, $C_1'' > 0$). La deuxième fabrique un produit x à un coût qui dépend à la fois de la quantité de x qu'elle produit et de la quantité produite par y (on peut supposer par exemple que l'entreprise 1 émet des effluents en quantité égale au niveau de production y , qui polluent la rivière et rendent l'eau utilisée par 2 moins propre à l'usage) :

$$C_2(x, y) = k(y) + c(x) \quad \text{avec : } dc/dx > 0, \quad dk/dy > 0, \quad d^2c/dx^2 < 0, \quad d^2k/dy^2 < 0$$

Si le prix de marché du bien 1 est p_1 , et celui du bien 2 p_2 alors les productions de 1 et 2, celles qui maximisent les profits de chaque entreprise, vont vérifier, dans l'équilibre naturel correspondant à la recherche du profit maximum :

$$\begin{aligned} \frac{\partial C_1}{\partial y} &= p_1 \\ \frac{\partial C_2}{\partial x} &= \frac{\partial c}{\partial x} = p_2 \end{aligned}$$

Système dont les solutions sont : x_n et y_n .

L'entreprise 1 va produire jusqu'à ce que son coût marginal égale le prix du marché, sans se soucier des coûts qu'elle occasionne à 2. Leur profit seront :

$$\begin{aligned} \Pi_1^n &= p_1 y_n - C_1(y_n) \\ \Pi_2^n &= p_2 x_n - C_2(x_n, y_n) = p_2 - c(x_n) - k(y_n) \end{aligned}$$

Comme l'entreprise 1 ne se préoccupe pas de l'effet de son activité sur le profit de 2, la situation n'est pas optimale pour l'ensemble des deux entreprises. Tout se passe comme si l'entreprise 2 supportait un coût fixe $k(y)$ sur lequel elle n'a pas de maîtrise et qui dépend de l'entreprise 1. Leur profit total pourrait être augmenté. Il faudrait pour cela que y soit fixé au niveau y_0 maximisant :

$$\Pi_1 + \Pi_2$$

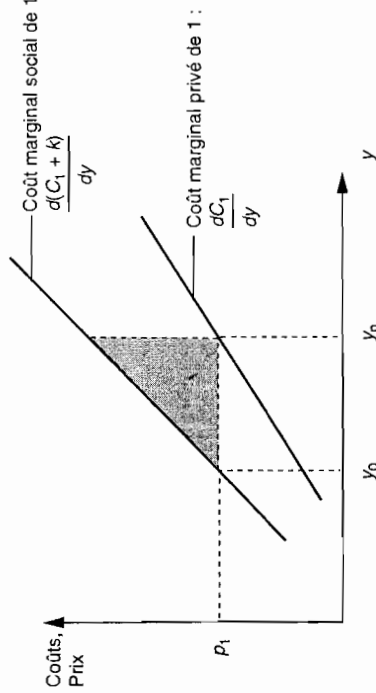
C'est celui pour lequel :

$$p_1 = \frac{\partial C_1}{\partial y} + \frac{\partial k}{\partial y}$$

En somme, l'entreprise 1 doit fixer son prix en tenant compte, non seulement de son coût marginal propre, mais du coût marginal que son niveau de production entraîne pour l'entreprise 2.

Ces résultats sont illustrés sur la figure 7-2 où est figuré en gris la perte de profit résultant de l'externalité.

FIGURE 7-2
Externalités et optimum de Pareto



Il est intéressant de recenser les moyens d'atteindre l'optimum, c'est-à-dire de ramener la production de l'entreprise 1 au niveau y_h au niveau y_0 .

Le premier est de nature réglementaire : la puissance publique fixe une norme d'émission du polluant telle que cette norme limite la production de l'entreprise 1 à y_0 .

Le second est d'ordre tarifaire : la puissance publique établit sur la production de y une taxe de niveau unitaire : dk/dy .

Le troisième moyen est un moyen de marchandage, application du théorème de Coase déjà vu au chapitre précédent : l'entreprise 2 a intérêt à passer un contrat avec 1, par lequel 2 paiera à 1 la somme $(y_h - y_0) dk/dy$ si 1 limite sa production à y_0 : l'optimum est atteint par un marchandage libre où chacun gagne. Dans une forme ultime, l'entreprise 2 achète l'entreprise 1.

Un quatrième moyen serait pour la puissance publique de vendre aux enchères des bons de pollution permettant l'émission d'une quantité de polluant égale au nombre de bons possédés ; le bon de polluant devient un facteur de production ; pour un niveau de prix q donné de ce bon, l'entreprise 1 demandera un nombre y de bons vérifiant :

$$y(p_1 + q) = dC_1/dy$$

Cette relation définit une fonction d'enchère pour l'entreprise 1, et s'il y a plusieurs entreprises du type 1, chacune enchérissant, la somme de leurs enchères définit une courbe d'enchère qui donne une relation entre le volume de bons émis et le prix d'adjudication de ces bons, permettant à la puissance publique d'atteindre le niveau optimum de pollution. Dans le cas présent, pour un prix du bon π , l'entreprise achètera y_π bons tel que $k'(y_\pi) = \pi$.

Tous ces moyens peuvent rétablir l'optimum. En dehors des problèmes de mise en œuvre, en termes notamment d'informations nécessaires, de coûts, de délais et de contrôle, leur différence essentielle réside dans la répartition des bénéfices entre les trois acteurs du jeu : l'entreprise 1, l'entreprise 2 et la puissance publique.

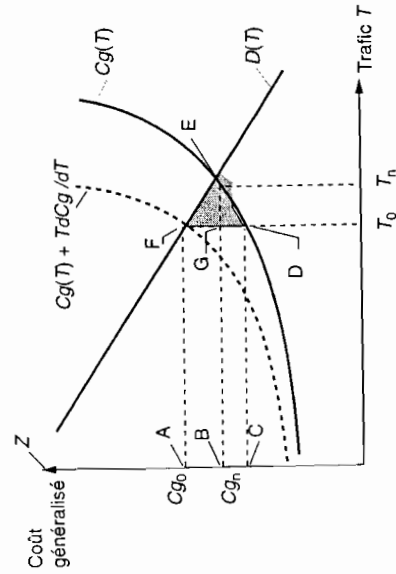
Munis de ces outils, analysons successivement les différents externalités : pertes de temps, atteinte à l'environnement, et insécurité.

2.1.2. Les pertes de temps

• Congestion routière

Voyons le cas de la congestion routière, en analysant d'abord la situation la plus simple, celle d'un flot constant sur une route uniforme indéfinie. Les conditions de circulation sont représentées par la figure 7-3 ci-dessous. On y suppose que les usagers prennent leur décision de voyager en fonction de leur coût généralisé et que ce coût généralisé est le même pour tous : les véhicules sont homogènes et les valeurs du temps sont toutes égales.

FIGURE 7-3
Le coût marginal de congestion



l'amélioration du surplus collectif s'accompagne, paradoxalement, d'une réduction du surplus des usagers, compensée, et au-delà, par des recettes supplémentaires de l'exploitant. Il pourrait en être différemment dans d'autres situations, bien sûr.

Le modèle est schématique, notamment dans la mesure où il suppose que tous les usagers ont la même valeur du temps. Si l'on abandonne cette hypothèse, des conséquences paradoxales peuvent apparaître. Ainsi, on verrait facilement que, dans un réseau très simple composé d'une origine, d'une destination et de deux arcs identiques reliant l'une à l'autre, il est avantageux, au regard du surplus collectif, d'instaurer un péage sur un des arcs : les usagers ayant les valeurs du temps les plus élevées se concentrent sur l'arc à péage, sur lequel ils gagnent du temps, et les autres à valeur du temps plus faible, choisissent l'itinéraire libre de péage où ils mettent plus de temps qu'auparavant. Pour de niveaux de péage adéquats, les gains des usagers dépassent les pertes des autres.

Pour analyser plus complètement les situations de congestion, revenons au modèle de congestion du chapitre V, et fondé sur la présentation par Arnoit, de Palma et Lindsey (1994). Dans ce modèle, les N usagers ont une heure d'arrivée préférée t^* , et trois valeurs du temps :

- α : valeur du temps de trajet
- β : coût d'une arrivée en avance
- γ : coût d'une arrivée en retard.

On suppose que le seul temps de trajet est le temps d'attente éventuelle au goulot d'étranglement de la route, où la capacité est s . Comme on a vu au chapitre V, l'équilibre naturel pouvait être représenté par un schéma du type de celui figuré sur la figure 7-4 : la queue commence à se former en t_g , elle augmente jusqu'en t_n , puis se résorbe progressivement et disparaît à t_f . Le trajet présente le même coût pour tous les usagers. Le premier et le dernier partis ne supportent qu'un temps d'attente. L'utilisateur partant en t_n arrive à l'heure désirée t^* et ne supporte qu'un temps de trajet. On a :

$$t_q = t^* - \frac{\gamma}{1-\gamma} \frac{N}{s}$$

$$t_f = t^* + \frac{\beta}{\beta+\gamma} \frac{N}{s}$$

$$t_n = t^* - \frac{\beta\gamma}{\alpha(\beta+\gamma)} \frac{N}{s}$$

Le coût total du trajet est : $\delta N^2/s$.

Il se partage pour moitié en coût d'attente, et pour moitié en coût de décalage. Cette situation n'est pas optimale. L'attente au goulot pourrait être supprimée sans que les heures d'arrivée soient modifiées. Il suffirait pour cela qu'un régulateur fixe à chaque usager son heure de départ, de façon que le débit soit constant et égal à la capacité du goulot. Alors il n'y aurait plus de coût d'attente, et on démontre que le coût total serait divisé par deux. Mais la répartition entre les usagers serait inégale, certains subiraient un coût élevé, d'autres, ceux que le régulateur ferait partir à t^* , un coût nul ; de toute façon sa mise en œuvre dans le cas routier est utopique : on ne voit pas comment il serait possible d'établir un ordre de priorité entre les nombreux usagers. Un moyen toutefois de réaliser une telle

Sur ce graphique sont figurées :
 - la courbe de demande $D(T)$, qui définit le volume de trafic T désiré de circuler pour un coût généralisé donné Cg ,
 - la courbe de coût moyen $Cg(T)$, qui définit le coût généralisé pour un volume de trafic donné. Cette courbe est croissante en raison de la gêne mutuelle et des pertes de temps, croissantes que s'occasionnent les véhicules entre eux quand leur nombre augmente.

Le trafic d'équilibre s'établit en T_n conduisant à un coût généralisé Cg_n . Le surplus collectif correspondant s'écrit :

$$S(T_n) = \int_0^{T_n} D(T) dT - T_n Cg(T_n)$$

Mais ce n'est pas pour T_n que le surplus est maximum. C'est pour une valeur T_0 du trafic vérifiant :

$$\max_T S(T) = \max_T \left[\int_0^T D(t) dt - TCg(T) \right]$$

Soit :

$$D(T_0) = Cg(T_0) + T_0 \frac{dCg}{dT_0}$$

On voit qu'à l'optimum le coût généralisé diffère du coût moyen de la quantité :

$$T_0 \frac{dCg}{dT_0}$$

Cette quantité représente la gêne que chaque usager occasionne aux autres : par sa présence, chacun des autres usagers voit son coût généralisé augmenter de dCg/dT_0 et ils sont au nombre de T_0 . Comment faire en sorte que le trafic se fixe au niveau T_0 et non au niveau T_n ? Le moyen le plus évident consiste à taxer le trajet, par exemple au moyen d'un péage, d'un montant de $T_0(dCg/dT_0)$. C'est le principe de la tarification au coût marginal social. On démontre que ce résultat, obtenu ici dans la situation simple d'un arc isolé, se généralise au cas d'un réseau (par exemple Yang et Huang, 1998) : l'optimum est obtenu lorsque, sur chaque arc, l'utilisateur qui l'emprunte paie un coût égal au coût que l'utilisateur marginal impose aux autres usagers présents sur cet arc.

Il est intéressant d'analyser sur le graphique précédent ce qui se passe quand on établit une telle taxe. Tout d'abord le surplus collectif est augmenté : il était auparavant de BEZ et il passe à FDCZ. Mais sa répartition change ; les usagers y perdent puisque leur surplus passe de BEZ à AFZ. En revanche les pouvoirs publics, collecteurs de la taxe, y gagnent ; leur revenu, nul auparavant, passe à AFDC. Sur le cas de figure considéré,

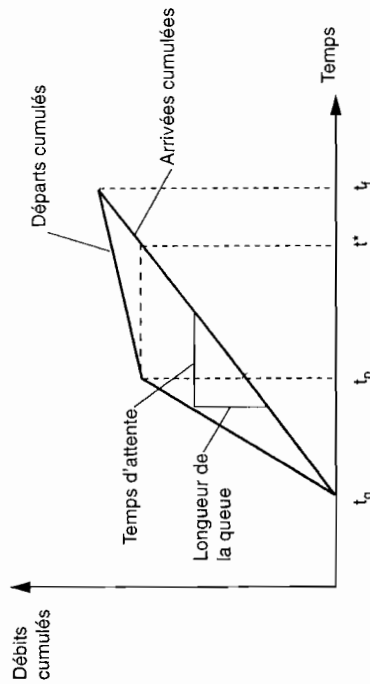
1. Pour un trafic supérieur, certains usagers auraient un coût généralisé supérieur à ce qu'ils sont prêts à payer, et vice-versa. Le seul point d'équilibre (de Nash) est en T_n . Cg_n .

situation est d'instaurer un péage variable dans le temps et qui, à chaque instant, serait égal au coût d'attente de l'équilibre naturel. L'équilibre de Nash qui en résulte pour les usagers aboutit à la suppression des files d'attente. Le débit des départs est égal à la capacité des goulots, et bien sûr à celui des arrivées.

Le coût pour les usagers est le même que dans la situation sans péage, mais les pouvoirs publics bénéficient maintenant d'une recette égale à la moitié du coût pour les usagers. On verrait facilement qu'à chaque instant le péage est égal au coût total que l'usager à qui il s'applique occasionne, par sa présence, aux autres usagers, plus précisément à ceux dont il accroît le décalage par rapport à l'horaire désiré.

En dehors des problèmes pratiques de mise en œuvre d'une taxe qui varie de façon continue, une difficulté réside dans la fixation de son montant, qui nécessite des informations précises, sur l'heure désirée, la valeur du temps et le flux de trafic.

FIGURE 7-4
Analyse de la congestion



Les modèles avec des utilisateurs différents sont un peu plus complexes. Arnott *et al.* (1994) ont exploré les conséquences de l'existence de plusieurs classes d'utilisateurs quant aux valeurs du temps et quant aux heures d'arrivée désirées. Les résultats sont que, dans le cas d'une infrastructure linéaire avec un seul goulot, un péage variable dans le temps permet toujours de supprimer les queues (il peut ne pas être ainsi dans le cas d'un réseau plus complexe, voir Jehiel et De Palma, 1996), et optimise le coût total de transport, avec des effets diversifiés sur la répartition et une réduction du temps de trajet total qui est moindre que lorsqu'il n'y a qu'une seule catégorie d'usagers.

Delons (1997) a élaboré un modèle (évoqué au chapitre IV) qui reproduit l'effet des modulations tarifaires mises en place au cours de 1995-1996 sur les autoroutes du réseau Cofiroute à l'entrée de Paris. Il met en évidence que la diversité des comportements et des situations des usagers quant aux origines et destinations réduit l'efficacité de la modulation tarifaire par rapport à ce qu'indiquent les modèles théoriques simples ne prenant pas en compte cette diversité.

Un autre moyen d'atteindre l'optimum réside dans l'information fournie aux automobilistes, une voie qui fait actuellement l'objet de nombreuses recherches en raison des possibilités de mise en œuvre offertes par le développement de l'informatique et des télécommunications. L'efficacité de ce moyen dépend de la manière dont les usagers répondent aux informations qui leur sont diffusées. Leur comportement fait l'objet de nombreuses études (par exemple celles de Khattak *et al.* (1997) et Polydoro-Poulou (1997)). Dans ces deux études, les comportements sont étudiés à travers un mélange de préférences révélées (comment les usagers répondent aux informations fournies par le système actuel ?) et de préférences déclarées, à travers des interviews où est demandé aux interviewés ce qu'ils feraient en présence de système d'information hypothétiques, du type de ceux qu'on peut imaginer pour le futur. Les analyses font apparaître les types d'informations auxquelles les usagers sont les plus sensibles : des informations quantifiées, des prescriptions de comportement. Il apparaît ainsi que les comportements peuvent être sensiblement modifiés à travers l'information ; et ces modifications peuvent porter aussi bien sur l'heure de départ et sur le choix des itinéraires.

Le taux de changement des comportements est un élément essentiel pour évaluer l'efficacité des informations fournies, car, contrairement à ce que l'on pourrait penser, une information parfaite n'est pas gage d'optimum, comme l'a montré De Palma (1992) pour un large éventail de situations concernant à la fois les coûts des itinéraires et les taux de réponses des usagers à l'information.

• Transport aérien

Le principe du modèle de goulot de Arnott *et al.* (1994) s'applique au transport aérien : c'est pour des raisons similaires d'heure d'arrivée ou de départ désirée qu'apparaît à certaines périodes de la journée la congestion des pistes. Mais la perte économique est plus faible que dans le cas routier car il n'y a pas de file d'attente. Il existe une règle qui fixe l'ordre d'arrivée des avions et qui évite les temps d'attente à l'atterrissage ; elle consiste dans la plupart des cas à reconduire chaque année entre les compagnies les allocations de slots de l'année précédente (le système des « droits du grand-père »), alors que l'optimum économique exigerait que les slots les mieux placés par rapport à l'horaire désiré soient alloués aux vols qui présentent l'utilité sociale la plus grande, c'est-à-dire en gros aux vols qui emportent le plus grand nombre de passagers. Un moyen d'assurer l'optimalité serait d'établir une taxe d'atterrissage dépendant de l'écart entre l'heure désirée et l'heure réelle d'atterrissage. Mais cette taxe devrait traire la perte que chaque avion occasionne aux autres par sa présence ; elle devrait être liée à l'emport des autres avions et d'autant plus forte qu'on est prêt de l'heure désirée, cette information est insuffisante pour fixer la taxe.

Un moyen de révéler les vrais valeurs des slots est de les mettre aux enchères. C'est ce qui est fait dans plusieurs aéroports des USA (Starkie,

1994). Le marché qui en résulte n'est pas complètement libre. Il y a en général deux marchés séparés, l'un pour les gros et l'autre pour les petits avions sans transfert possible des slots de l'un à l'autre. Ceci a pour but de protéger les lignes peu fréquentées qui desservent des régions de faible densité et qui, étant indispensables à la survie de l'activité économique de ces régions, assurent une sorte de service public. De toute façon une certaine rupture avec le système des droits du grand-père est nécessaire là où l'on veut ouvrir le marché du transport aérien ; le minimum en ce domaine semble être ce qui est en vigueur en Europe à la suite des directives bruxelloises : les slots non utilisés peuvent être réalloués, et les nouveaux entrants ont des droits prioritaires sur une partie des slots réalloués. Notons enfin que la pleine efficacité de la mise aux enchères n'est obtenue que si les enchères sont coordonnées entre aéroports origines et destinations.

- Pertes de temps ferroviaires

La congestion ferroviaire nuit également à la réalisation de l'optimum, non par l'apparition de temps d'attente, puisque, comme dans le secteur aérien, il existe des règles fixant la priorité. Mais rien n'assure que ces règles aboutissent à un optimum. Cet optimum pourrait être obtenu par la tarification, mais le calcul des tarifs est complexe : il faut évaluer ce que la circulation envisagée fait perdre aux autres circulations par sa présence, c'est-à-dire construire une grille avec la circulation en cause, une autre sans elle, et imputer à cette circulation les différences de recettes entre les deux situations ; l'opération est à la rigueur réalisable pour un opérateur intégré, qui dispose de toutes les informations nécessaires ; elle est beaucoup plus difficile à mettre en œuvre pour un régulateur extérieur qui ne dispose pas de toutes les informations nécessaires.

La complication essentielle de l'établissement d'une tarification des sillons horaires résulte de l'imbrication de trains de vitesses différentes, par exemple trains de marchandises, trains de voyageurs à longue distance et trains de banlieue ; de ce fait la caractérisation d'une circulation dépend de deux paramètres, la position dans le temps et la vitesse, et non comme pour le transport aérien d'un seul paramètre, la position dans le temps. Cette complication est réduite par l'établissement de règles qui donnent une priorité à certains types de trains dans certaines tranches horaires ; par exemple on donne priorité aux trains de banlieue durant les heures de début et de fin de travail. Mais on pourrait aussi envisager des mises aux enchères de circulation, un moyen sûr de dégager les priorités des différents candidats. La difficulté est que l'objet mis aux enchères comporte deux dimensions de localisation, dans l'espace et dans le temps, et qu'il y a des liens entre certains de ces objets (par exemple le trajet Paris-Lyon doit associer un sillon Paris-Mâcon et un sillon Mâcon-Lyon).

Il faut également noter que, en situation de congestion, le retard d'un train par rapport à l'horaire prévu entraîne des perturbations dans le trajet des autres trains et crée des externalités négatives. Pour les corriger, il convient de prévoir des pénalités de retard ou des bonus de régularité.

2.1.3. L'environnement

On a vu au chapitre V les possibilités de traduire en termes monétaires les effets sur l'environnement. On pourrait alors s'attendre à ce que leur correction s'effectue à travers un système de taxes. En fait les taxes sont peu utilisées pour les combattre, si ce n'est au travers de timides modulations des licences annuelles, et c'est l'outil réglementaire qui assure l'essentiel des efforts de corrections de ces externalités, à travers la réglementation des normes d'émission des véhicules et la composition des carburants. Pourtant l'outil tarifaire présente des avantages importants en termes d'efficacité, comme le montre l'encadré 7-2.

Le choix entre les deux instruments dépend aussi du degré de précision avec lequel sont connus les quantités et coûts correspondant au niveau de pollution optimale. Si les deux sont parfaitement connus, le choix est théori-

ENCADRÉ 7-2

Réglementation, taxation et décentralisation des décisions

Soit la situation où n entreprises ($i = 1, n$) émettent chacune une quantité K_i de polluant à l'état naturel, et peuvent dépolluer, le coût marginal de dépollution de chacune variant avec le niveau d'émission q_i selon des fonctions décroissantes différentes d'une entreprise à l'autre :

$$c_i(q_i)$$

Un même niveau de pollution totale Q peut être obtenu :

– soit en fixant à chaque entreprise i un niveau d'émission G_i tel que :

$$\sum G_i = Q$$

– soit en fixant une taxe t sur chaque unité de pollution émise. Alors chaque entreprise minimisera la somme du coût de dépollution et de la taxe payée sur la pollution restante, soit :

$$\min_{q_i} \int_{q_i}^{K_i} c_i(t) dt + tq_i \quad \text{soit :} \quad t = c_i(q_i)$$

Les deux procédures aboutissent au même résultat si :

$$q_i = G_i$$

mais la détermination des G_i par la puissance publique exige, surtout si n est grand, beaucoup plus d'informations que celle de t , qui peut en outre être ajustée par tâtonnement.

Par ailleurs la fixation d'une taxe constitue une plus forte incitation au progrès dans les techniques de dépollution, car alors l'entreprise a un plus grand intérêt à l'amélioration de ces techniques, qui en outre la conduisent à dépolluer davantage, comme on peut le voir sur le schéma ci-dessous : une réglementation fixée à N_1 a le même résultat qu'une taxe t dans la situation initiale ; mais si un progrès technique se produit, la procédure de taxation conduit automatiquement à une plus forte réduction de pollution que la procédure de réglementation : on aboutit à N_2 .

De même, la perte économique résultant de la réglementation, égale à l'aire grisée claire, sera :

$$P_t = \varepsilon^2(b+c)(b/c)^2 = P_A(b/c)^2$$

Ainsi, si les bénéfices sont à peu près indépendants du niveau de pollution, la réglementation est préférable, car une erreur sur les coûts aura peu d'influence sur l'optimum, et c'est l'inverse si les bénéfices varient beaucoup avec le niveau de pollution (par exemple en raison des seuils).

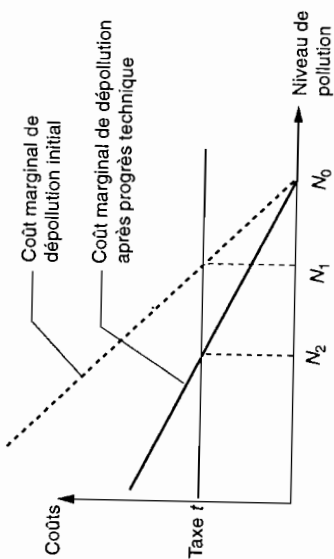
Les résultats seraient inverses en cas d'incertitude non plus sur les coûts, mais sur les bénéfices de la dépollution.

On peut illustrer ces principes dans le cas de la pollution atmosphérique : il est facile d'appliquer une réglementation ou une taxation aux constructeurs automobiles, dont les technologies sont similaires ; une réglementation fine de l'usage de même qu'une taxation de l'usage est difficile à envisager, car elle nécessiterait pour être optimale, une multiplicité de prescriptions et de contrôles : le degré de pollution émis dépend de nombreux paramètres difficiles à contrôler. Quant au choix, pour le contrôle des constructeurs automobiles entre la fixation des normes ou une taxation, il doit être dicté par le fait que l'incertitude sur les bénéfices à attendre d'une dépollution est beaucoup plus forte que l'incertitude sur les coûts, et que les coûts varient beaucoup plus que les bénéfices avec le niveau de pollution. Ces derniers faits plaident en faveur d'un contrôle des émissions des véhicules par une taxation assise par exemple sur le niveau d'émission (taxe à l'achat ou vignette payée annuellement, éventuellement modulée selon l'année).

Le système des marchés de droits ou permis négociables, recommandé par l'analyse économique, est peu utilisé dans les transports. Deux exemples s'en rapprochent. En Californie les constructeurs automobile se voient imposer une moyenne d'émission par véhicule produit et restent libres de répartir cette moyenne entre leur production. Le système de contrôle du trafic automobile à Singapour (Chin et Smith, 1997) comporte aussi un dispositif de permis négociables. Ce contrôle est fondé d'abord sur une maîtrise de l'usage, à travers un système de vignettes de zone destinées à être remplacées par le péage électronique ; et ensuite sur la maîtrise de la possession des véhicules, assurée initialement par des taxes à l'achat et maintenant par un système de permis négociables émis en quantités limitées. Les auteurs analysent l'efficacité de cette mesure en termes de minimisation de la perte économique due à une erreur d'appréciation des paramètres : les pentes respectives des courbes de demande et d'offre montre qu'on est bien dans le cas où le contrôle des quantités est préférable au contrôle par les prix.

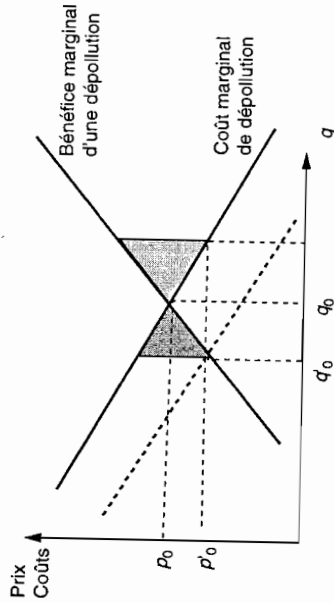
2.1.4. L'insécurité

La prise en compte des externalités d'insécurité pose des problèmes particuliers qui ont été analysés en particulier par Jansson (1994). Consi-



quement, de ce point de vue, indifférent. Si en revanche, une incertitude les affecte, alors, comme le montre l'encadré 7-3, c'est la taxation ou la réglementation qu'il faut préférer selon les pentes des fonctions de coût ou de bénéfice. Ce choix dépend enfin des possibilités pratiques respectives d'établissement d'une taxe ou d'un règlement et du contrôle de leur respect.

ENCADRÉ 7-3 Précision des informations, et choix entre taxation et réglementation



Supposons que les vrais coûts et bénéfices unitaires de pollution soient représentés par la courbe en traits pleins. L'optimum vrai est défini par q_0 correspondant à un coût p_0 . Supposons que les coûts soient mal évalués, la courbe de coût constatée est la courbe pointillée. Alors la perte économique moyenne résultant de la taxation est représentée par la moitié de l'aire grisée en sombre. Si, en valeurs absolues, b est la pente de la courbe des bénéfices c celle des courbes des coûts, si ε est l'incertitude sur les coûts (pour une même abscisse la distance entre les courbes pointillées et la courbe vraie est de ε) alors cette perte vaut :

$$P_t = \varepsilon^2(b+c)$$

dérivons d'abord la situation où il y a un seul type de trafic. Soient :

Q : le débit,

$r(Q)$: le risque d'accidents, qui dépend du débit,

a : la disposition à payer de l'automobile moyen pour annuler le risque,

c : le coût d'un accident supporté par le reste de la collectivité (perte de production, frais éventuels ambulance, de soins... s'ils ne sont pas supportés par l'automobiliste).

Alors le coût total des accidents est :

$$TC = (a + c) rQ$$

Il convient de faire supporter à chaque usager la différence entre le coût marginal pour la collectivité et le coût privé, et une tarification optimale devrait donc comporter deux termes :

- les frais liés à l'insécurité et supportés par la collectivité soit c .
- le terme : $(a + c) Q \partial r / \partial Q$ qui dépend de la variation du taux d'accidents en fonction du trafic Q . L'analyse des statistiques d'accident ne fait pas apparaître de variation nette du taux d'accident en fonction du trafic, et le taux r est souvent considéré comme constant, même si certains auteurs considèrent qu'il serait normal que le nombre d'accidents varie comme le carré du trafic (c 'est ainsi que varie la probabilité de rencontres entre véhicules dans une hypothèse de mouvement aléatoire) et donc que r soit proportionnel au trafic.

L'analyse est plus complexe lorsque le trafic est composite. Jansson présente le modèle suivant, dans lequel les accidents proviennent simplement de collisions entre voitures et bicyclettes (les accidents entre voitures obéissent à la loi précédente, indépendamment de ce qui se passe entre voitures et bicyclettes ; il n'y a pas d'accidents entre bicyclettes) ; soient :

Q : le trafic automobile

M : le trafic deux roues

$X(Q, M)$: le nombre d'accidents moyen résultant des trafics Q et M . Ces accidents sont subis par les bicyclettes seulement. Le coût total des accidents sera en moyenne :

$$TC = aX(Q, M) + cX(Q, M)$$

a représentant la valeur de la vie humaine,

c représentant les coûts sociaux non payés par les accidentés.

La tarification à appliquer est toujours la différence entre le coût marginal collectif et le coût privé (le coût privé est nul pour les automobilistes) ; elle diffère selon qu'il s'agit des automobilistes ou des bicyclettes et sera :

$$P(\text{auto}) = dTC/dQ = (a + c) dX/dQ$$

$$P(\text{bicy}) = dTC/dM - aX/M = (a + c) dX/dM - aX/M$$

L'auteur montre ensuite que selon les valeurs de a et c , et selon la fonction $X(Q, M)$, on peut trouver une variété de situation entre le cas où les automobilistes ne paient rien et seules les bicyclettes paient, et celui où les bicyclettes devraient être subventionnées, leur nombre réduisant les accidents. Quant au revenu issu de ces péages, il peut lui aussi couvrir plus ou moins la somme des coûts publics cX , ou la dépasser.

2.2. Tarification dans un environnement non-optimal

Dans les modèles présentés jusqu'ici, seul un mode était considéré, les autres modes étant implicitement supposés gérés à l'optimum. Le tableau change lorsque des imperfections existent dans le reste de l'économie, par exemple si les modes concurrents ne sont pas tarifés à l'optimum ou si l'exploitant est tenu à une contrainte d'équilibre financier.

2.2.1. Imparfaite tarification des modes concurrents

Alors la tarification optimale n'est plus la tarification au coût marginal, mais doit tenir compte de l'imparfaite tarification du mode concurrent. On aboutit à une relation simple explicitée dans l'encadré joint. Si par exemple le mode concurrent est sous tarifé, le mode sous analyse devra l'être aussi, dans une proportion qui dépend de la taille des marchés des deux modes et de l'élasticité de substitution existant entre ces deux marchés.

ENCADRÉ 7-4

Tarification d'un mode dans le cas d'une imparfaite tarification du mode concurrent

Soient q_1 et q_2 les deux modes substitués, c étant le bien courant de l'économie de prix unité. Soient p_1 et p_2 les tarifs des deux modes.

Les consommateurs sont supposés avoir la même fonction d'utilité $U(q_1, q_2, c)$ et le même revenu, ce qui conduit aux égalités classiques sur la proportionnalité des utilités marginales aux prix et aux fonctions de demande.

On suppose par ailleurs que la production de $q_1(q_2)$ consomme $a_1 q_1(a_2 q_2)$ unités de c , dont la ressource totale est C , et que les marchés sont concurrentiels. L'optimum collectif est atteint pour :

$$U(q_1, q_2, c) \text{ Maximum}$$

avec :

$$q_1 = q_1(p_1, p_2, r)$$

$$q_2 = q_2(p_1, p_2, r)$$

$$c = c(p_1, p_2, r)$$

$$c = r - p_1 q_1 - p_2 q_2$$

et la contrainte :

$$a_1 q_1 + a_2 q_2 + c = C$$

En appelant λ la variable duale de cette contrainte, et en développant le lagrangien, on obtient finalement :

$$q_1 = \lambda [q_1 + (p_1 - a_1) (\partial q_1 / \partial p_1) + (p_2 - a_2) (\partial q_2 / \partial p_2)]$$

$$q_2 = \lambda [q_2 + (p_1 - a_1) (\partial q_1 / \partial p_1) + (p_2 - a_2) (\partial q_2 / \partial p_2)]$$

$$1 = \lambda [1 - (p_1 - a_1) (\partial q_1 / \partial r) - (p_2 - a_2) (\partial q_2 / \partial r)]$$

On vérifie que :

$$\lambda = 1$$

$$p_1 = a_1$$

$$p_2 = a_2$$

est une solution ; c 'est la tarification de premier rang.

Si p_1 est imposé à une valeur différente de a_1 , alors la première équation du système précédent est supprimée, seules subsistent les deux dernières, dont on tire en éliminant λ :

$$(p_1 - a_1) [(\partial q_1 / \partial p_2) + q_2 (\partial q_1 / \partial r)] + (p_2 - a_2) [(\partial q_2 / \partial p_2) + q_2 (\partial q_2 / \partial r)] = 0$$

En introduisant les élasticités directes et croisées :

$$e_1 = (\partial q_1 / \partial r) / (q_1 / r)$$

$$e_{22} = (\partial q_2 / \partial p_2) / (q_2 / p_2)$$

$$e_{12} = (\partial q_1 / \partial p_2) / (q_1 / p_2)$$

et en notant les écarts de tarifs par rapport à l'optimum de premier rang :

$$\delta x_1 = p_1 - a_1$$

$$\delta x_2 = p_2 - a_2$$

on obtient :

$$q_1 \delta x_1 [e_{12} + (p_2 q_2 / r) e_1] + q_2 \delta x_2 [e_{22} + (p_2 q_2 / r) e_2] = 0$$

Si $(p_2 q_2) / r$ est faible (place du mode 2 faible dans l'ensemble de l'économie), alors :

$$(q_2 \delta x_2) / (q_1 \delta x_1) = - (e_{12} / e_{22})$$

Cette formule est généralisable lorsqu'il y a plus de deux modes, et lorsque le transport est un bien intermédiaire (Quinet, 1992).

2.2.2. Contrainte d'équilibre financier de l'exploitant

Le modèle présenté ci-dessus, qui ne prend en compte que les dépenses de construction, généralise au cas des infrastructures de transport le résultat classique selon lequel, dans une activité à rendement constant, la tarification au coût marginal couvre les dépenses : c'est le théorème de Mohring (Mohring et Harwitz, 1962).

Ce résultat sera présenté ici en suivant une formulation due à Bernard (1985). Soit une société composée de n usagers identiques dont la fonction d'utilité est :

$$U(q, T, l) = U(q, T, \bar{t} - tT)$$

avec

q : bien courant de prix,

T : nombre de déplacements effectués, chacun à un coût unitaire,

\bar{t} : temps total disponible,

t : temps disponible pour des activités de consommation,

t : temps passé en transport,

la contrainte de revenu est :

$$q + cT + \gamma T - r \leq 0$$

expression dans laquelle γ est la taxe qui frappe le transport.

On suppose en outre que le temps de transport t est fonction du trafic T et des caractéristiques de l'infrastructure :

$$f(nT, K) - t \leq 0$$

Ce temps de trajet varie avec le trafic total, mais est une donnée pour chaque individu, dans la mesure où n est grand.

Par ailleurs, le coût annuel de fonctionnement de l'équipement K est ρK , recouvrant les dépenses fixes d'intérêt et d'entretien, le stock disponible de bien courant est F vérifiant :

$$F = nq + ncT + \rho K$$

Dans ces conditions, à γ et t donnés, chaque usager maximise son utilité, compte tenu de sa contrainte de revenu et de la taxe qui frappe le transport. La puissance publique détermine γ et K , compte tenu de la contrainte de ressources totales, pour que cette utilité soit maximale.

Le lagrangien définissant le comportement du consommateur s'écrit :

$$L = U(q, T, \bar{t} - tT) - \lambda(q + cT + \gamma T - r)$$

Sa maximisation par rapport à q et T donne :

$$\frac{\partial U}{\partial q} = \lambda$$

$$\frac{\partial U}{\partial T} - t \frac{\partial U}{\partial t} = \lambda(\gamma + c) = \frac{\partial U}{\partial q}(\gamma + c)$$

Quant à l'optimisation que réalise la puissance publique, elle est représentée par le programme suivant

$$\max_{q, T, t, K} U(q, T, \bar{t} - tT)$$

$$nq + ncT + \rho K - F \leq 0$$

$$f(nT, K) - t \leq 0$$

$$\frac{\partial U}{\partial T} - t \frac{\partial U}{\partial t} = \frac{\partial U}{\partial q}(\gamma + c)$$

En introduisant pour les 2 premières contraintes les variables de Lagrange μ et v on aboutit à :

$$\frac{\partial U}{\partial T} - t \frac{\partial U}{\partial t} - \mu nc - v \left(n \frac{\partial f}{\partial(nT)} \right) = 0$$

$$\mu \rho + v \frac{\partial f}{\partial K} = 0$$

$$-T \frac{\partial U}{\partial t} + v = 0$$

On en tire, en tenant compte de la dernière contrainte du programme de la puissance publique :

$$\gamma \frac{\partial U}{\partial q} - (nT) \frac{\partial U}{\partial t} \frac{\partial f}{\partial(nT)} = 0$$

Or $\left(\frac{\partial U}{\partial t} / \frac{\partial U}{\partial q} \right)$ représente la valeur du temps h (c'est en effet le taux marginal de substitution entre le temps et l'argent). D'où :

$$\gamma = h(nT) \frac{\partial f}{\partial(nT)}$$

c'est le coût marginal de congestion, présenté d'une autre manière que précédemment, et :

$$\frac{\partial U}{\partial q} \frac{\rho}{n} + T \frac{\partial U}{\partial l} \frac{\partial f}{\partial K} = 0$$

ou :

$$\rho = -h(nT) \frac{\partial f}{\partial K}$$

Si la fonction f est homogène de degré 0, ce qui correspond à une situation de rendements constants (quand on double la dépense d'infrastructure, le trafic qu'on écoute à vitesse donnée est doublé), alors :

$$K \frac{\partial f}{\partial K} + (nT) \frac{\partial f}{\partial (nT)} = 0$$

La recette issue de la tarification est :

$$\gamma(nT) = h(nT) \frac{\partial f}{\partial (nT)} = -h(nT) K \frac{\partial f}{\partial K} = \rho K$$

Elle couvre exactement la dépense d'infrastructure.

Ce résultat peut être étendu à une situation où le trafic varie au cours des années, comme on peut le voir à travers une modélisation simplifiée :

Soient :

- $P_t(T)$ la fonction de demande inverse de l'année t ,
- T_t le trafic de l'année t ,
- $AC(T_t, K)$ le coût moyen de transport avec une infrastructure de coût K et le trafic T_t ,
- θ la durée de vie, supposée donnée, de l'infrastructure,
- j le taux d'actualisation.

La gestion optimale correspond à la maximisation du surplus collectif :

$$\max_{K, T_t} \left[\sum_{t=1}^{t=\theta} \left(\int_0^{T_t} P_t(u) du - T_t AC(T_t, K) \right) \exp(-jt) - K \right]$$

Les conditions sont :

$$1) P_t(T_t) = AC(T_t, K) + T_t \frac{\partial AC}{\partial T_t}$$

Cela exprime qu'il faut imposer une taxe égale au coût marginal social.

$$2) \sum_{t=1}^{t=\theta} T_t \frac{\partial AC}{\partial K} \exp(-jt) + 1 = 0$$

On vérifie que, si la fonction de coût de l'infrastructure est à rendements constants, c'est-à-dire si :

$$AC(AT, \lambda K) = AC(T, K)$$

alors la somme actualisée des recettes issues du coût marginal est égal à K . Ce résultat ne tient plus lorsque l'investissement varie de façon discontinue, voir Oum et Zhang (1990) : alors la couverture des dépenses dépend de la chronologie des trafics.

Comme on l'a vu au chapitre V sur les fonctions de coût des infrastructures, on voit que les cas où cette fonction de coût est à rendements constants ou décroissants se produisent essentiellement en matière routière et pour les infrastructures urbaines ; alors la recette issue de la tarification au coût marginal social équilibrera ou même dépassera les dépenses de construction. Dans les autres cas les rendements croissants entraîneront un déficit.

Newbery (1988) a montré que, pour les opérations de renouvellement de chaussées, la situation de déficit était fréquente. Il fait l'hypothèse que les dommages occasionnés à une chaussée dépendent de deux paramètres : le climat qui, même sans trafic, finit avec le temps par détruire les routes, et le passage des essieux, une chaussée d'un type donné ne pouvant supporter qu'une quantité maximale d'essieux. Dans le modèle, le passage d'un essieu avance la date de reféction de la chaussée, ce qui, par l'actualisation, accroît la dépense actualisée. Si le trafic ne croît pas et si les conditions météorologiques ne jouent pas, alors, dans un état stationnaire où les sections de route du réseau sont d'ancienneté également répartie, la recette issue du coût marginal couvre exactement les dépenses de réparation de la chaussée. Dans la situation courante où le trafic croît et où les conditions climatiques jouent un rôle, alors la recette issue du coût marginal est inférieure aux dépenses, dans des proportions qui semblent en général modérées.

La situation où la tarification au coût marginal engendre un déficit au regard des dépenses du gérant d'infrastructure est fréquente. Ce déficit présente de nombreux inconvénients. Il doit être comblé par des impôts, et ces impôts engendrent des pertes d'efficacité et ont un coût psychologique. D'autres raisons d'efficacité seront développées au chapitre IX et résident dans l'intérêt qu'il y a à soumettre le gestionnaire de l'infrastructure à une contrainte budgétaire pour l'inciter à une meilleure gestion. Enfin, des considérations d'équité et de juste répartition peuvent conduire à juger souhaitable que les dépenses routières soient payées par ceux qui en profitent.

Pour toutes ces raisons, on peut souhaiter substituer à la tarification au coût marginal une tarification qui, tout en visant à maximiser le bien-être collectif, respecte une contrainte d'équilibre budgétaire. La tarification dite de Ramsey-Boiteux (Ramsey, 1927 ; Boiteux, 1956) répond à cet objectif. Elle aboutit (voir encadré) à un prix défini par la formule :

$$\frac{P_i - c_i}{P_i} = \frac{\lambda}{\epsilon_i}$$

formule dans laquelle :

l'indice i recouvre les biens produits,

c_i est le coût marginal de production du bien i ,

P_i est le tarif cherché,

ϵ_i est l'élasticité du bien i à son prix,

λ est un paramètre d'ajustement.

La tarification de Ramsey-Boiteux dans un cadre d'analyse partielle

Soit une entreprise qui produit i (i varie de 1 à n) biens, dont les demandes sont indépendantes, et les fonctions inverses de demande sont :

$$p_i(q_i)$$

le coût de production est :

$$C(q_1, \dots, q_n)$$

Alors le surplus de Dupuit résultant d'un vecteur de production donné sera :

$$SC = \sum_i \int_0^{q_i} p_i(u) du - C(q_1, \dots, q_n)$$

la contrainte budgétaire est :

$$\sum_i p_i(q_i) q_i - C(q_1, \dots, q_n) = 0$$

En appelant μ la variable duale correspondante et c_i la dérivée de C par rapport à q_i la maximisation de SC donne :

$$p_i(q_i) - c_i - \mu p_i - \mu q_i \frac{dp_i}{dq_i} + \mu c_i = 0$$

En réarrangeant les termes, et en introduisant l'élasticité

$$\varepsilon_i = \frac{p_i}{q_i} \frac{dq_i}{dp_i}$$

on aboutit à :

$$\frac{p_i - c_i}{p_i} = \frac{\mu}{1 - \mu} \varepsilon_i$$

Le paramètre μ est déterminé de façon à ce que la contrainte budgétaire soit respectée.

Si les produits i engendrent des externalités la formule doit être complétée. Appelons $CE(q_1, \dots, q_n)$ le coût total externe. Alors le programme devient :

$$\max_{q_i} SC = \sum_i \int_0^{q_i} p_i(u) du - C(q_1, \dots, q_n) - CE(q_1, \dots, q_n)$$

avec :

$$\sum_i p_i q_i - C(q_1, \dots, q_n)$$

soit :

$$\frac{p_i - c_i - \left(\frac{1}{1 - \mu} \right) ce_i}{p_i} = \frac{\mu}{1 - \mu} \varepsilon_i$$

Dans cette formule, ce_i est le coût marginal externe, et μ la variable duale assurant la contrainte budgétaire.

On peut étendre cette analyse au cas où les fonds publics ont un coût (par exemple les impôts qui les financent ne sont pas conformes aux critères de l'optimum).

Alors, le programme à résoudre est, dans le cas de deux produits :

$$\max SC = \sum_{i=1}^{i=2} p_i(u) du - C(q_1, q_2) + \lambda \left(\sum_i p_i q_i - C \right)$$

Ce qui donne, tous calculs faits :

$$\frac{p_i - \frac{\partial C}{\partial q_i}}{p_i(q_i)} = \frac{\lambda}{(1 + \lambda) \varepsilon_i}$$

ε étant l'opposé de l'élasticité de la demande du bien i .

La formule est un peu compliquée si les biens sont substitués ou complémentaires :

$$\frac{p_i - \frac{\partial C}{\partial q_i}}{p_i} = \frac{\lambda}{\lambda + 1} \left(\frac{q_1}{p_1} \frac{\partial p_1}{\partial q_1} + \frac{q_2}{p_1} \frac{\partial p_2}{\partial q_1} \right)$$

qui se transforme en :

$$\frac{p_i - \frac{\partial C}{\partial q_1}}{p_i} = \frac{\lambda}{1 + \lambda} \frac{1}{\varepsilon_1} \frac{\varepsilon_1 \varepsilon_2 - \varepsilon_1 \varepsilon_{12}}{\varepsilon_1 \varepsilon_2 - \varepsilon_{12} \varepsilon_{21}} = \frac{\lambda}{1 + \lambda} \frac{1}{\eta}$$

η étant appelée super-élasticité.

Oum et Tretheway (1988) ont étendu cette relation au cas où les biens ou services produits engendrent des externalités, ce qui est le cas général dans les transports (voir encadré).

Les formules présentées permettent de saisir la philosophie de ce type de taxation : le prix s'écarte du coût marginal d'autant plus que la demande est rigide, peu sensible au prix et d'autant plus, bien sûr, que le déficit du coût marginal est important. Par ailleurs, les externalités jouent un rôle important quand le déficit est « faible » (variable duale faible en valeur absolue) : elles comptent presque autant que le coût marginal privé ; en revanche quand le déficit est important, leur incidence s'estompe au regard du coût marginal privé.

2.3. La tarification des infrastructures : de la théorie à la réalité

Dans la pratique, la tarification des infrastructures est bien différente de ce que préconise la théorie. Elle est en fait beaucoup plus fondée sur des considérations de financement que sur des considérations d'orientation des usagers. Examinons ce qui se passe pour chaque mode.

En matière routière, les principaux outils de tarification sont la taxe sur les carburants, les péages et les vignettes et le stationnement payant. Aucun

cation des aéroports est établie dans la plupart des pays, et notamment en France, de façon à équilibrer les recettes et les dépenses, mais la répartition selon les aéronaves ne suit pas les recommandations de la théorie économique ; elle est fondée sur la taille des avions et non sur l'encombrement qu'ils provoquent. On peut la rattacher à une disposition à payer des compagnies, mais non à une responsabilité de ces compagnies dans la gêne aux autres utilisateurs.

3. Les choix d'investissement¹

Les choix d'investissements constituent un domaine d'application privilégiée de la théorie des surplus. Les modèles développés précédemment à la section 2.2.2 fournissent une règle de choix, mais dans un cadre trop théorique pour déboucher sur une méthode pratique d'évaluation.

Les méthodes pratiques sont fondées sur le calcul d'un surplus de Dupuit qui vise à prendre en compte tous les effets de l'investissement. Examinons successivement la construction de ce critère, appelé le bénéfice actualisé, puis les conditions de sa validité, ensuite ses lacunes, enfin les modalités de sa mise en œuvre dans quelques cas particuliers.

3.1. Le critère du bénéfice actualisé

Celui-ci prend en compte les effets sur les usagers, gains de temps et variation de frais de fonctionnement. Ces effets se répercutent sur l'ensemble de l'économie. Si le reste de l'économie était gérée à l'optimum, l'analyse partielle résumerait les effets globaux et le surplus de Dupuit limité au secteur des transports, le surplus de l'usager, mesurerait le surplus collectif complet.

En fait il y a des non-optimalités évidentes liées aux infrastructures de transport, et il faut intégrer des termes complémentaires correspondant aux sources de non-optimalités, comme on l'a vu en 1-3 :

- les effets externes liés à l'environnement,
 - les rendements croissants, notamment ceux localisés « près » (au sens économique) de l'investissement, c'est-à-dire dans les organismes gérant les infrastructures,
 - enfin il faut tenir compte de la présence de l'État : les transferts auxquels il procède sont modifiés par la réalisation de l'investissement.
- Examinons rapidement chacun de ces éléments, en prenant comme exemple celui d'une autoroute à péage, qui met en jeu la plupart des éléments intervenants dans les calculs de rentabilité.

1. Les résultats et propositions contenus dans cette section sont étroitement inspirés des réflexions contenues dans Boiteux (1995).

de ces outils ne présente la variabilité dans le temps et dans l'espace pour reproduire, même de façon imparfaite, les coûts de congestion ou d'environnement. Newbery (1995) a déterminé les coûts marginaux routiers en Angleterre pour diverses situations. Les coûts auquel il arrive varient entre près de 10F par km dans les centres urbains à l'heure de pointe et quelques centimes par km sur les routes rurales. Plusieurs autres études ont permis de calculer une moyenne du coût marginal en agglomération (de Palma et Marchal, 1997 ; Newbery, 1998 ; Nicolas, 1998) ; ces études analysent comment varie le coût total de transport de l'agglomération quand la matrice *O*. *D* se déforme homothétiquement ; les résultats montrent que le coût marginal de congestion est de l'ordre du coût de circulation ; ceci conduirait à des tarifs de l'ordre de quelques francs par kilomètre.

En comparaison, les taxes sur l'essence rapportent au km entre 30 et 50 centimes, quelles que soient la route et l'heure. L'environnement n'a jamais fait l'objet de taxes, sauf la taxe sur le bruit autour des aéroports ; mais cette dernière n'est pas fondée sur des évaluations de coûts marginaux. Dans tous les pays, les péages ont eu pour objectif de rassembler les ressources nécessaires au financement et non d'orienter le choix des usagers, à par quelques essais de modulation horaire, ils présentent en général le défaut de taxer l'infrastructure non congestionnée et de laisser libre de péage l'infrastructure congestionnée. Quant aux stationnements payants, ils sont en général plus élevés aux endroits et aux périodes où la congestion est la plus élevée ; mais ils sont loin de varier dans les proportions voulues et leur effet sur le trafic est ambigu : ils contribuent à réduire les séjours des véhicules aux centres-villes, et donc à accroître la mobilité. Le système qui permettrait de recouvrir aussi exactement que l'on veut le coût marginal existe : c'est le péage électronique, où des cartes électroniques sont débitées au passage de bornes, à la manière de cartes de téléphone. Mais les essais pour implanter ce système ont toujours échoué pour des raisons politiques (une expérience de deux ans à Hong-Kong a finalement été abandonnée). Les péages de cordon implantés autour de certaines villes scandinaves ont des objectifs financiers. Enfin les combinaisons de moyens tels que taxes sur l'essence, licences et vignettes, réglementations du stationnement sont rarement coordonnés pour répondre à une politique de maîtrise du trafic : le cas de Singapour est une exception, liée à l'organisation particulière de cet État.

En matière de transports routiers de marchandises, les calculs de tarification montrent que les gros poids lourds devraient payer des taxes très élevées, en raison des dégradations qu'ils font subir aux chaussées et qui croissent avec la puissance quatrième du poids des essieux. Une tarification correspondant à ces caractéristiques avait été mise en place en France en 1968, avec la taxe à l'essieu, mais cette taxe n'a pas été révisée depuis cette date, et la taxe sur le gazole a été décrochée de la taxe sur l'essence et est maintenant beaucoup plus faible que cette dernière.

La situation n'est pas meilleure dans les autres modes. La tarification des chemins de fer ne tient pas compte des coûts de congestion. La tarifi-

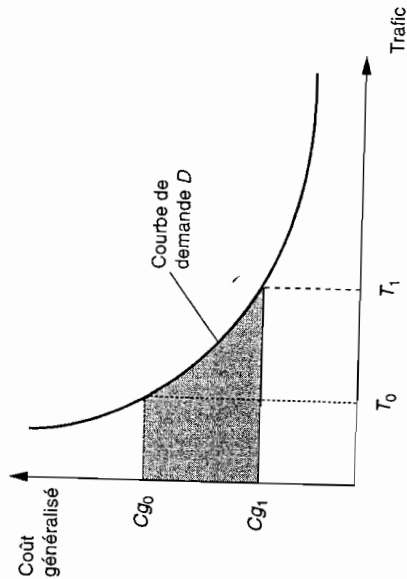
3.1.1. Le surplus des usagers

Dans son acception la plus simple (figure 7-5), le surplus des usagers est calculé par la variation de l'aire sous la courbe de demande (qui définit le trafic en fonction du coût généralisé de circulation) lorsque, à la suite de l'investissement, le coût généralisé passe de Cg_0 à Cg_1 . C'est l'application directe de la formule 7-3. Cette aire est un trapèze curviligne qui, compte tenu de l'imprécision des chiffres, peut en général être approché par un trapèze rectiligne :

$$SC = T_0(Cg_0 - Cg_1) + (T_1 - T_0) (Cg_1 - Cg_0)/2$$

Les valeurs de T_0 et T_1 sont fournies par l'étude de trafic ; celles des coûts généralisés intègrent les coûts monétaires de circulation, les temps passés, les éventuels termes de confort, ou de pénalités pour décalage horaire.

FIGURE 7-5
Surplus de l'usager



Un autre calcul du surplus des usagers consiste à calculer celui-ci sur la loi de demande qui se trouve contenue dans le modèle de trafic. Ainsi, dans le cas de modèles de choix discrets, Small et Rosen (1981) ont établi la formule suivante pour l'expression du bénéfice des usagers :

$$SC = T \int_{U_0}^{U_1} \sum_{i=1}^n \pi_i(u) du$$

expression dans laquelle :

i : représente les alternatives disponibles,

U : représente l'utilité indirecte indexée 0 dans la situation de départ, et dans la situation finale,

$\pi_i(u)$: est la probabilité de choisir le mode i lorsque l'utilité de ce mode est u .

Si le modèle de choix est du type logit :

$$SC = (T/\lambda) \Delta \left[\log \left(\sum_i \exp(-\lambda u_i) \right) \right]$$

C'est la formule du « logsum » présentée au chapitre IV.

Si la valeur du temps prise dans l'expression élémentaire du surplus de Dupuit est la même que celle qui ressort du modèle de trafic, et plus généralement si le coût généralisé a la même expression que l'utilité du modèle de trafic, alors (Jara-Diaz, 1990) les deux méthodes de calcul du surplus aboutissent à peu près au même résultat (exactement en l'absence de transfert modal).

Ceci suppose que les utilités des modèles de trafic représentent bien la valeur sociale du transport. Des considérations d'équité ou de service public peuvent conduire à établir des différences entre les valeurs individuelles et les valeurs collectives. Ainsi plusieurs auteurs considèrent que, dans les transports domicile-travail, il convient de choisir de donner au temps de transport de tous les usagers la même valeur unitaire.

On peut aussi s'interroger sur la précision de la détermination des valeurs du temps (et plus généralement des paramètres de qualité de service) déduites des modèles. Il est facile de voir que des imprécisions dans cette détermination peuvent être sans influence sur les résultats de l'étude de trafic, mais avoir au contraire des conséquences très importantes sur le surplus (voir encadré 7-6). Dans ces conditions, on pourra préférer la méthode élémentaire de calcul du surplus, en choisissant une valeur ou un éventail de valeurs du temps (ou de qualité de service) résultant de la confrontation de plusieurs méthodes d'évaluation.

Une mention spéciale doit être faite concernant l'insécurité. Le problème essentiel y est celui de l'information sur les conditions d'accidents : les individus, qui sont prêts à payer pour réduire le risque qu'ils courent, connaissent-ils bien ce risque ? Si la réponse est oui, il faut introduire un terme de sécurité dans le coût généralisé de circulation ou dans l'utilité du moyen de transport. Si la réponse est non, il convient de ne pas compter la sécurité dans le coût généralisé, mais d'introduire la variation du coût total de l'insécurité dans le calcul des effets de l'investissement, et surtout de mieux informer les usagers.

ENCADRÉ 7-6

Sensibilité du surplus aux valeurs du temps

L'exemple pédagogique ci-dessous peut facilement être transposé à des situations plus réalistes :

Soit un ensemble d'usagers constitué de deux groupes en nombre égal. Pour le premier, la valeur du temps est $h = 100$; pour le second elle est $h' ; h'$ est mal connu et se situe entre 200 et 300.

Dans la situation initiale, il n'y a qu'un mode de transport, emprunté donc par tous les usagers, caractérisé par un prix $p = 100$ et un temps de trajet $t = 1$. Supposons qu'on réalise un autre mode plus rapide et plus cher, dont le prix du billet est $p = 200$ et le temps de trajet $t' = 0,3$. Le coût du trajet pour les usagers de la première catégorie est de :

$$\begin{array}{ll} 100 + 1 \times 100 = 200 & \text{dans le mode ancien} \\ 200 + 0,3 \times 100 = 230 & \text{dans le mode nouveau} \end{array}$$

Donc les usagers de la catégorie 1 ne changent pas de mode. Le coût du trajet pour les usagers de la seconde catégorie est frappé d'incertitude.

Dans le mode ancien, il vaut entre

$$100 + 1 \times 200 = 300 \text{ et } 100 + 1 \times 300 = 400$$

Dans le mode nouveau, il vaut entre

$$200 + 0,3 \times 200 = 260 \text{ et } 200 + 0,3 \times 300 = 290$$

Donc, l'incertitude sur la valeur du temps des usagers de la catégorie 2 n'a pas d'effet sur l'étude de trafic : de toute façon ces usagers changent de mode. Elle en a en revanche sur le calcul du surplus qui, par usager, varie entre :

$$300 - 260 = 40$$

et :

$$400 - 290 = 110$$

3.1.2. Les externalités d'environnement

Leur prise ne compte ne pose pas de problème de principe particulier, à partir du moment où l'on a déterminé une batterie de coûts pour les atteintes à l'environnement. Il suffit d'évaluer ces atteintes dans la situation initiale et dans la situation finale (quantité de polluants émis, nombre de décibels...), et de multiplier ces quantités physiques par les coûts unitaires. Son ainsi pris en compte les effets sur le bruit et sur la pollution de l'air, locale ou globale. En revanche, il n'existe pas actuellement de méthode pour prendre en compte dans les bilans monétaires les autres atteintes à l'environnement telles que :

- effets de coupure
- pollution des eaux et des sols
- conséquences esthétiques.

3.1.3. Les organismes gérant les transports

Le gestionnaire de l'infrastructure nouvelle et les gestionnaires des infrastructures vont voir leurs dépenses d'exploitation et leurs redevances se modifier. Si ces opérateurs sont à rendements constants et tarifient au coût marginal (alors égal au coût moyen), les dépenses et recettes supplémentaires s'annulent. Si, comme c'est le cas le plus fréquent, ils sont à rendements croissants, et/ou ne tarifient pas au coût marginal, il faut intégrer au surplus collectif leurs variations de profit (cf. encadré 7-1).

3.1.4. Les contribuables et l'État

Le surplus est mesuré en prix à la consommation et inclus donc les taxes indirectes : dans le cadre d'un budget public équilibré, hypothèse cohérente avec l'idée de comparaison des états d'équilibre qu'implique la théorie des surplus, une variation des taxes perçues par l'État à la suite de

la réalisation d'une nouvelle infrastructure entraîne une variation de même montant (et de signe opposé) des autres taxes payées par les contribuables, dont des surplus individuels.

Si toutes les taxes indirectes étaient au même taux, cet effet aurait pour seule conséquence de multiplier par un même coefficient les surplus individuels. Mais les biens ne sont pas frappés au même taux, et, parmi les taux exceptionnels figurent les taxes sur les carburants, frappés à des taux très sensiblement plus élevés que le taux normal. Il convient alors d'ajouter au surplus la variation des revenus des taxes sur les carburants (dans leur part qui excède le taux normal de TVA, voir l'encadré pour l'expression mathématique correspondante) qui sont engendrés par la réalisation de l'investissement de transport.

ENCADRÉ 7-7

La prise en compte des taxes sur les carburants

Si un bien a est taxé à un taux différent du taux θ_a normal de taxe indirecte θ , comme par exemple c'est le cas pour les carburants routiers, l'équilibre de l'entreprise qui le produit, supposée opérer dans un marché concurrentiel, donc vendre au coût marginal, s'écrit ainsi, en utilisant l'indice i pour les autres biens de l'économie, et en appelant p_i les prix à la consommation taxes comprises :

$$\sum_i p_i(1 - \theta) dq_i + p_a(1 - \theta_a) dq_a = 0$$

Si l'investissement routier entraîne une variation dq_a de la consommation du bien a , le terme du surplus auquel il aurait donné lieu aurait été, avec une taxation normale ;

$$dU = \sum_i p_i dq_i = -p_a dq_a$$

Compte tenu du taux de taxation particulier, il sera :

$$dU = -\frac{\theta - \theta_a}{1 - \theta} p_a dq_a$$

Par rapport au taux de droit commun θ , l'existence d'un taux exceptionnel procure à l'État un supplément (algébrique) de recette égal à $(\theta - \theta_a)$ fois le prix hors taxe soit :

$$\frac{\theta - \theta_a}{1 - \theta} p_a$$

par unité du bien a consommée (ou économisée).

3.1.5. Synthèse

L'ensemble des éléments qui viennent d'être dégagés doit être combiné avec le jeu de l'actualisation. Le même bien considéré à deux points de temps différents représente en fait deux biens de valeurs différentes, le

bien présent ayant en général une valeur supérieure au même bien disponible plus tard (Varian, 1992). C'est la préférence pour le présent, que traduit le taux d'intérêt réel, dégagé des influences perturbatrices de l'inflation : si le taux d'intérêt nominal annuel est de $x\%$, si le taux d'inflation est de p le taux d'intérêt réel est approximativement : $(x - p)$.

Chacun des éléments de surplus qui vient d'être analysé doit donc être actualisé de sa date d'apparition à la date actuelle. Si, à la fin de la durée de vie de l'investissement celui-ci garde une valeur résiduelle, il convient de l'intégrer dans le calcul. La formule qui exprime le critère ainsi obtenu, qu'on appelle le bénéfice actualisé, s'exprime sous la forme :

$$B = - \frac{I}{(1+i)^{t_0}} + \sum_{t=t_0+1}^{t=T} \frac{A(t) - r(t)}{(1+i)^t} + \frac{VR}{(1+i)^T}$$

Dans cette expression :

i : est le taux d'actualisation,

I : est le coût de l'investissement. Si la dépense d'investissement est étalée sur plusieurs années, I ne représente pas la somme des dépenses annuelles, mais la somme actualisée à l'année de mise en service.

t_0 : est l'année de mise en service,

$r(t)$: représente les variations de dépenses d'entretien et d'exploitation de l'année t pour les gestionnaires des transports,

VR : la valeur résiduelle de l'équipement,

$A(t)$: rassemble les variations de surplus des catégories précédemment recensées :

- surplus des usagers
- surplus des riverains
- surplus des exploitants d'infrastructures
- surplus de l'État (contribuables).

3.1.6. L'utilisation de ce critère pour la programmation

Comment ce critère peut-il être mis en œuvre pour les choix d'investissement et l'élaboration des programmes ?

L'utilisation du bénéfice actualisé permet d'abord de définir la date de réalisation d'un investissement donné : c'est la valeur de t_0 qui maximise $B(t_0)$.

Cette maximisation dépend de l'ensemble des paramètres intervenant dans la formule. Mais il y a une situation, assez représentative des situations les plus fréquentes, où cette maximisation aboutit à une relation remarquable : c'est celle où la durée de vie de l'investissement est infinie, et où les avantages vont croissant dans le temps (ceci se réalise lorsque le trafic augmente dans le temps). Alors la date optimale de mise en service est l'année t_0 qui vérifie (cf. encadré 7-8) :

$$\frac{A(t_0)}{I} = i$$

Détermination de la date optimale de mise en service

En raisonnant en temps continu, et en appelant :

- θ l'année de mise en service,
 - t l'année courante,
 - $a(t, \theta)$ l'avantage à l'année t lorsque la mise en service est à l'année θ ,
 - I le coût de l'investissement,
 - r le taux d'actualisation,
 - B le bénéfice actualisé,
- on a :

$$B(\theta) = \int_{t=\theta}^{\infty} a(t, \theta) \exp(-rt) dt - I \exp(-r\theta)$$

Si la fonction $a()$ ne dépend pas de θ et est croissante avec t , alors le maximum de B est obtenu pour θ vérifiant :

$$a(\theta) = rI$$

La rentabilité immédiate de l'investissement considéré doit être égale au taux d'actualisation.

Le critère du bénéfice actualisé permet aussi de comparer des projets incompatibles, par exemple deux variantes de tracé d'un même projet : on détermine pour chacun sa date optimale et on choisit celui qui, placé à sa date optimale, rapporte le bénéfice le plus élevé. Les décisions concernant des projets indépendants (c'est-à-dire tels que la rentabilité de l'un n'est pas affecté par l'existence ou la non-existence des autres, et réciproquement) sont également simples à l'aide des critères précédents : on réalise chaque projet à sa date optimale. Lorsque les projets ne sont pas indépendants il faut énumérer la combinaison des séquences possibles, pour chacune d'entre elles déterminer la date de réalisation de chaque projet, en déduire le bénéfice actualisé de la séquence, et retenir finalement la séquence qui procure le bénéfice actualisé le plus élevé.

Il faut bien voir les limites des enseignements à tirer du bénéfice actualisé, qui n'est pas une optimisation, mais simplement une comparaison entre le projet envisagé et la situation de référence à laquelle il est comparé. Le fait que le bénéfice actualisé soit positif prouve simplement que le projet est préférable à la situation de référence. D'où l'importance à la fois de bien choisir la situation de référence et d'ouvrir au maximum l'éventail des projets dont on compare le bénéfice actualisé. En ce qui concerne la situation de référence, l'idéal serait qu'elle soit la meilleure décision possible en l'absence du projet qu'on évalue. On ne peut jamais être sûr qu'il en sera ainsi, car cela supposerait en quelque sorte d'avoir résolu le problème de la hiérarchisation des projets, qui est précisément celui qu'on cherche à résoudre. Mais un biais fréquent des analystes qui veulent faire passer un projet est de choisir une situation de référence

La règle de choix d'investissement en situation de tarification non optimale (inspiré de Small, 1992)

Soit une route sur laquelle chaque usager met un temps $t(q, k)$ qui dépend du trafic q et de la largeur de la route K ; la valeur du temps commune à tous les usagers est h ; le coût annuel de la route est ρK (amortissement + entretien). Enfin, la loi de demande est $p(q)$. Le surplus collectif est :

$$SC(q, K) = \int_0^q p(u) du - hqt(q, K) - \rho K$$

on voit facilement que la maximisation par rapport à q et K conduit aux règles habituelles :

- de tarification :

$$p(q) = ht(q, K) + hq \frac{\partial t}{\partial q}$$

- de choix d'investissement :

$$\rho = -hq \frac{\partial t}{\partial K}$$

Supposons maintenant que la tarification ne soit pas optimale, et plus précisément qu'elle soit inférieure à l'optimum. Cela se traduira par une contrainte :

$$p(q) - ht(q, K) - p_0 \leq 0$$

où p_0 est la tarification imposée, inférieure à l'optimum. Si λ est la variable duale qui lui est associée, on voit que :

$$\rho = -h \frac{\partial t}{\partial K} (q - \lambda)$$

La règle de choix doit être modifiée : on calcule la rentabilité en appliquant le gain de temps permis par l'investissement, non pas au trafic q , mais au trafic $(q - \lambda)$.

voie entraîne une faible réduction du coût généralisé de circulation et l'investissement présente une faible rentabilité.

En revanche, l'instauration d'une tarification de la congestion aurait des effets beaucoup plus considérables sur le coût généralisé, sur le trafic, et sur le surplus collectif. Ces considérations montrent l'importance d'une bonne modélisation des trafics, et en particulier du trafic induit, comme l'illustre la figure 7-7.

Jansson (1993), en présentant deux graphiques de ce type, soutient que celui de gauche est représentatif des conditions interurbaines, et celui de droite des conditions urbaines. Goodwin (1996) développe ce point en soutenant, par l'intermédiaire d'études avant-après, que le trafic induit est en général sous-estimé en zone urbaine, et en déduit que les rentabilités sont surestimées.

mauvaise, pour accroître d'autant la rentabilité du projet proposé. Un autre impératif est de balayer le plus largement possible le champ des variantes, sans oublier qu'une alternative à un projet peut être une mesure d'exploitation ou de tarification, ou peut concerner un autre mode, ou même se rapporter à un autre domaine que celui des transports.

3.1.7. Les conditions de validité du critère du bénéfice actualisé

Les conditions essentielles qui assurent la validité du critère du bénéfice actualisé sont, comme on l'a vu :

- le caractère marginal des transformations analysées,
- l'optimalité du reste de l'économie (essentiellement entreprises se comportant en *price-takers*),
- la mise entre parenthèses des problèmes de répartition. Cette ignorance peut être justifiée soit parce que la répartition est jugée optimale, soit parce que les transferts nécessaires seront faits.

Illustrons par quelques exemples les conséquences du non respect de ces hypothèses.

3.1.7.a) Cas de transformations non marginales

Comme on l'a vu plus haut, le surplus de Dupuit présente alors une dose d'arbitraire, et des mesures plus appropriées semblent être la variation équivalente et la variation-compensatoire Kanemoto et Méra (1985) ont établi des formules approchées permettant de passer de la situation de transformations marginales à celle de situations finies. Morisugi (1997) a effectué des évaluations du surplus de Dupuit et de la variation équivalente résultant de la réalisation du réseau ferroviaire japonais depuis environ un siècle. Il ressort de ses calculs que la différence entre les deux évaluations est faible, certainement inférieure aux erreurs de mesure qui affectent ces évaluations.

3.1.7.b) Effet de l'imparfaite tarification dans le système de transports

L'hypothèse que le reste de l'économie est à l'optimum implique en particulier que l'infrastructure nouvelle soit tarifée au coût marginal social. S'il n'en est pas ainsi, la règle de choix des investissements ne s'applique pas, et doit être modifiée comme le montre l'encadré 7-9 sur un exemple simple (voir Nilsson, 1992 pour une présentation plus complète en présence de plusieurs modes substitués).

On voit en particulier qu'en situation de sous-tarification la règle de choix usuelle surestime la rentabilité des investissements. Les distorsions dans les décisions d'investissement entraînées par une imparfaite tarification sont nombreuses ; elles ont été illustrées par le paradoxe de Braess et celui de Villé, présentés au chapitre IV, qui fournissent des exemples où un investissement, même gratuit, peut être néfaste.

Un autre exemple est celui où la demande de transport sur un axe est très élastique. Alors (figure 7-6) une augmentation de la capacité de la

cation sans discernement de l'analyse classique en présence d'un monopole qui suit comme règle de gestion la maximisation du profit.

FIGURE 7-6
Conséquences d'une demande de transport élastique

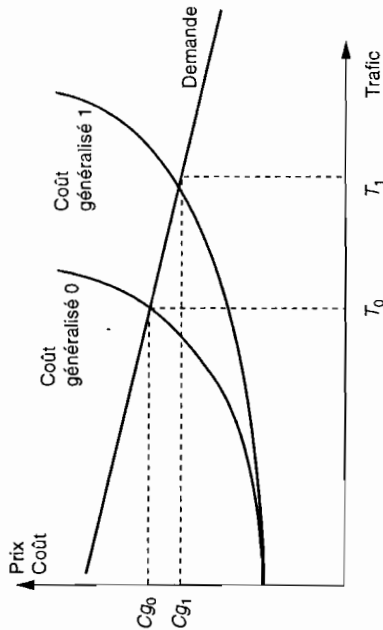
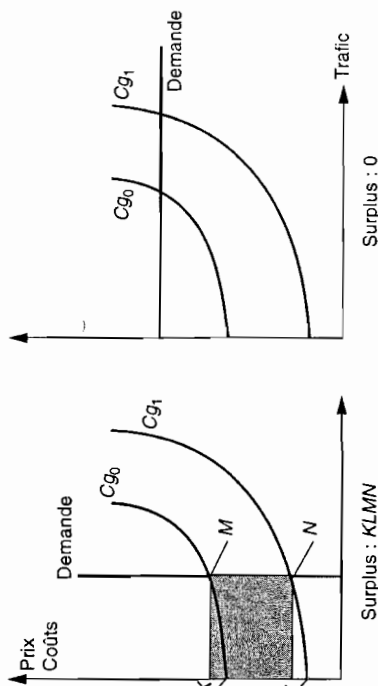


FIGURE 7-7
Conséquence de différences dans l'appréciation du trafic induit



L'élasticité de la demande de transport et l'effet de demande latente qui en résulte s'est bien constaté au début de la période de développement des transports publics urbains en France, dans les années 70. Le développement de l'offre en transport public ne s'est pas traduit par un désengorgement de la circulation routière aussi élevé qu'on l'espérait. En effet la place libérée par ceux qui avaient ripé de l'automobile vers le transport public était en partie reprise par de nouveaux automobilistes qui auparavant ne se déplaçaient pas.

3.1.7.c) Non-optimalités extérieures au système des transports

Les non-optimalités peuvent aussi se produire en dehors du secteur des transports. L'encadré joint fournit un exemple des erreurs d'une appli-

ENCADRÉ 7-10

Le calcul de rentabilité en présence de monopoles

Soit un univers économique dans lequel sont produits :

- un bien courant q : produit par une activité gérée de façon concurrentielle qui n'emploie que de la ressource naturelle en quantité q ,
- un bien s : produit par un monopole à un coût unitaire c constant en ressources naturelles, et vendu au prix unitaire p .

Le transport entre les logements et le monopole t : produit par chaque ménage à un coût unitaire constant, chaque transport servant à transporter une unité du bien s produit par le monopole. Le trafic sur la route est s . Les individus sont munis d'une fonction d'utilité : $U(q, s) = q + u(s)$. La ressource naturelle est en quantité limitée r et on a donc :

$$q + cs + ts = r$$

Alors le comportement des consommateurs est décrit par :

$$t + p = u'(s)$$

Celui du monopole par :

$$\max_s [(u'(s) - t - c) s]$$

La quantité produite par le monopole, s^* , vérifie :

$$u'' - t - c + su'' = 0$$

Supposons maintenant que le coût du transport passe de t à $t + dt$. Le calcul usuel du surplus associé à cette transformation fournit la valeur :

$$dSC_u = -s dt$$

C'est le surplus de l'utilisateur. Cette expression peut se transformer en remarquant que la quantité de ressources naturelles r ne change pas, donc que :

$$dr = dq + c ds + t ds + s dt$$

On voit alors que :

$$dSC_u = dq + (c + t) ds$$

En fait, le vrai surplus est égal à la variation d'utilité des ménages :

$$dSC = \frac{\partial U}{\partial q} dq + \frac{\partial U}{\partial s} ds = dq + u'(s) ds = dq + (p + t) ds$$

Comme p est différent de c , dSC diffère de dSC_u

On voit ainsi qu'on obtient la valeur exacte du surplus en ajoutant au surplus de l'utilisateur dSC_u , la variation de du profit à prix constant du monopole i . En effet :

$$\delta \Pi = (p - c) ds$$

et :

$$dSC = dSC_u + \delta \Pi$$

1. La variation de profit total du monopole est : $d\Pi = [(p - c) s] = (p - c) ds + s dp = \delta \Pi + s dp$.

3.1.7.d) Déséquilibres macro-économiques

Dans la période actuelle, un déséquilibre macro-économique majeur, source d'écart par rapport à l'optimum collectif¹, affecte nos économies, c'est le chômage. La prise en compte du chômage a fait l'objet dans le passé de deux groupes de travail du Commissariat Général du Plan, élaborés d'ailleurs à une période où le chômage s'accompagnait d'un autre déséquilibre, celui du commerce extérieur, à l'époque chroniquement déficitaire (CGP, 1981 et CGP, 1983). Les conclusions très prudentes de ces deux rapports recommandaient deux méthodes alternatives pour la prise en compte du chômage : la première consistait à effectuer les calculs économiques en prenant comme prix du travail un chiffre inférieur au salaire du marché, pour traduire le fait que ce prix du marché ne reflète pas le coût social du travail. Le taux de réflexion proposé était de l'ordre de 0,85. La deuxième méthode proposée consistait à évaluer le nombre d'emplois créés par l'action envisagée et à fournir cette indication quantitative en guise de critère annexe au critère monétarisé classique des surplus. La difficulté de mise en œuvre de cette deuxième méthode réside dans le calcul des emplois créés par l'investissement. Ce calcul nécessite de pouvoir calculer tous les effets, même les plus éloignés de l'action initiale. Cela nécessite une description complète de l'économie et des mécanismes auxquels elle obéit. Or, on ne dispose pas d'une telle maquette intégrant tous les effets, et endogénéisant toutes les variables économiques.

3.1.7.e) Coût des fonds publics

On suppose souvent que les fonds publics ont un coût collectif dans la mesure où les taxes qui les fournissent ne sont pas optimales. Alors un franc public n'est pas égal à un franc privé, il vaut $(1 + \lambda)$ francs privés. Les méthodes précédentes restent valables à condition de multiplier les fonds publics par le coefficient $(1 + \lambda)$. Abraham et Laure (1959) ont démontré que cette procédure était aussi celle qui résultait de l'optimisation d'un programme d'investissement dans le cadre d'une contrainte de crédit imposée de l'extérieur.

3.1.8. Les lacunes du critère du bénéfice actualisé

Le critère du bénéfice actualisé fait l'impasse sur certains faits qui accompagnent les choix d'investissement et problèmes que se pose le décideur : les effets sur les équilibres écologiques ; les conséquences esthétiques ; la présence d'incertitudes ; les incidences sur le développement économique ; les contraintes de financement ; les conséquences sur la répartition.

3.1.8.a) Les effets sur les équilibres écologiques et les conséquences esthétiques

Ces deux catégories d'effets présentent la caractéristique commune de ne pas pouvoir, au moins actuellement être traduits en termes monétaires. Les conséquences sur les équilibres écologiques ont une mesure simpliste qui est celle des changements dans la biodiversité : une nouvelle infrastructure, rompant l'équilibre écologique local, entraîne la disparition de certaines espèces, disparition soit locale, soit éventuellement mondiale. On sait à peu près prévoir - c'est l'affaire des écologues - les conséquences en termes de biodiversité de la réalisation d'une nouvelle infrastructure ; mais comment les apprécier monétairement ? Les conséquences de la biodiversité cumulent en effet toutes les difficultés de l'évaluation des biens non marchands : effets à long terme mettant en jeu les générations futures, conséquences mal connues, valeurs d'existence et d'option probablement élevées par rapport à la valeur d'usage. L'analyse contingente fournirait des résultats très incertains et dispersés en raison du manque d'information des interviewés ; l'évaluation par le coût des dommages est impossible, au moins à l'heure actuelle. L'économiste a actuellement peu de choses à dire sur la biodiversité. Il n'en va pas de même, au moins en théorie, pour les conséquences esthétiques. On pourrait interroger les citoyens, au moyen d'analyses contingentes, sur la qualité esthétique du projet et en déduire une valorisation collective du projet susceptible d'être introduite dans le calcul du bénéfice actualisé. Mais ces analyses statistiques sont coûteuses, leurs résultats incertains, et jusqu'ici elles n'ont guère été utilisées.

La prise en compte de ces deux types de conséquences ne peut pas s'effectuer à travers la monétarisation, mais par fourniture d'informations au décideur. En ce qui concerne les équilibres écologiques, on décrit les effets sur la survie des espèces végétales et animales ; pour les effets esthétiques, des maquettes et des photomontages donnent une idée de ce que deviendra le paysage, et permettent donc aux décideurs - et plus généralement aux acteurs intervenant dans le processus de décision - de se faire leur opinion.

3.1.8.b) L'incertitude

Deux catégories d'incertitude entrent en jeu : la première ne sera pas réduite avec le temps ; elle touche par exemple les aléas sur les coûts de construction, dans la mesure où toutes les études nécessaires ont été faites, et où aucune étude supplémentaire ne pourrait venir réduire l'incertitude ; l'autre catégorie est celle qui sera partiellement levée dans l'avenir, par exemple celle qui concerne la croissance économique future : on en saura plus sur elle si on attend. Examinons successivement ces deux sortes d'incertitude en utilisant les outils classiques du comportement en présence de risque, notamment l'attitude d'aversion au risque. L'expression la plus simple en est de supposer qu'un agent confronté à une loterie dont l'issue est une somme de monnaie de moyenne \bar{g} et d'écart-type σ en retire une utilité non pas égale à \bar{g} , mais à : $(\bar{g} - r\sigma^2/2)$. Dans

1. En effet, en situation de chômage, le théorème établissant l'équivalence entre analyse partielle et analyse globale n'est plus valide (cf. encadré 7-1).

cette expression r est un coefficient représentatif de l'aversion au risque de l'agent (si $r = 0$ il n'y a pas d'aversion au risque). L'annexe 6 détaille les outils d'analyse de l'incertitude.

- L'incertitude « irréductible »

Avec ces outils d'analyse, il n'est pas compliqué en principe d'introduire l'incertitude dans les calculs de rentabilité, à partir du moment où on peut définir des probabilités pour les événements possibles. Lorsqu'il n'y a ni aversion ni gout du risque, et que les lois reliant les aléas aux résultats sont linéaires, alors le calcul de rentabilité classique effectué à partir des espérances mathématiques conduit à l'espérance mathématique de l'utilité et fournit donc le bon critère. Ceci devient erroné lorsqu'il y a aversion au risque ou lorsque les lois reliant l'aléa au résultat ne sont pas linéaires. Deux cas particuliers illustrent les différences introduites alors. Le premier concerne le calcul des avantages issus d'un investissement lorsque le trafic, prévu T est frappé d'une incertitude irréductible. Alors le calcul de l'avantage est de la forme :

$$A = hTt(T)$$

h étant la valeur du temps, T le trafic et $t(T)$, temps de trajet présente une forte courbure quand on s'approche de la saturation. En raison de cette courbure, si T est une variable aléatoire, $E(A)$ peut être fort différente de $A(T^*)$, et les instructions pour le calcul des rentabilités d'investissement corrigent en général cet effet.

Le deuxième exemple est lié à l'attitude vis-à-vis de l'aversion au risque ; cette attitude peut être fort différente d'un agent à l'autre ; elle dépend du tempérament de l'agent et aussi de sa surface financière : une loterie où l'on a une chance sur deux de perdre 100 et une chance sur deux de gagner 100 vaut à peu près 0 pour une entité qui dispose d'un avoir de 1 000 000, très élevé par rapport au risque encouru ; en revanche pour une entité qui dispose d'un avoir de 100, la loterie en cause sera très dangereuse (elle implique un risque de ruine non négligeable), et l'aversion au risque sera alors beaucoup plus forte. Ainsi dans les décisions d'investissement, les pouvoirs publics, et en premier lieu l'État, ont en général une aversion au risque beaucoup plus faible que les investisseurs privés. Cela se traduit dans la pratique des choix : les décideurs publics font presque seulement attention à la rentabilité moyenne sans se préoccuper des risques ; en revanche les investisseurs privés, et surtout les banques qui ont traditionnellement une forte aversion au risque, font très attention au risque. On peut cependant critiquer les pratiques courantes ainsi définies : d'abord pour ce qui concerne les pouvoirs publics, leur inattention au risque est excessive et entraîne des conséquences concrètes néfastes qu'on analysera plus bas ; ensuite en ce qui concerne les opérateurs privés, leurs analyses du risque se borne en général à explorer des scénarios dits contrastés, dont on sait rarement s'il sont vraiment contrastés, et dont on ne définit pas la probabilité d'apparition, livrant les résultats de l'analyse à une grande subjectivité d'interprétation.

- L'incertitude « réductible » et l'irréversibilité

Le schéma précédent d'une incertitude irréductible est un cas extrême. En fait la plupart des incertitudes se réduisent avec le temps : on en saura plus l'an prochain qu'on en sait aujourd'hui. Cette incertitude pose des problèmes particuliers lorsqu'elle se présente dans des situations marquées par l'irréversibilité des décisions, comme c'est souvent le cas en matière d'investissement d'infrastructure où, une fois l'infrastructure réalisée, on ne peut pas en changer l'usage si, à l'expérience, celui-ci s'avère insuffisamment profitable. L'investissement constitue alors un fonds perdu (« *sunk-cost* »). Ce trait se retrouve dans d'autres situations ; ainsi les dépenses de publicité, et, à un moindre degré, les investissements en matériel roulant, revendables mais à perte.

Voyons sur un exemple simple, inspiré de Pindyck (1991), les changements qu'entraîne cette situation par rapport aux règles de décision classiques :

Soit un investissement de coût 800 susceptible de rapporter la première année une valeur certaine de 100 et ensuite chaque année jusqu'à l'infini la somme de 50 avec une probabilité 1/2 et 150 avec une probabilité 1/2. Dans le cadre d'une incertitude irréductible, du type de celle analysée à la section précédente, la valeur actuelle nette de l'investissement serait en espérance mathématique :

$$VAN = -800 + \sum_{t=0}^{\infty} \frac{100}{(1,1)^t} = 300$$

en supposant que le taux d'actualisation est de 10 % et l'investissement doit être réalisé dès maintenant.

Les choses sont différentes ici, puisque dans un an on en saura plus que maintenant. N'a-t-on pas intérêt à attendre un an ? Pour le voir calculons l'espérance mathématique de la valeur actuelle nette en $t = 1$: si la somme recueillie est de 50, il est facile de voir qu'il n'y a pas intérêt à investir, car alors la valeur actuelle nette est :

$$VAN(p = 50) = -\frac{800}{1,1} + \sum_{t=1}^{\infty} \frac{50}{(1,1)^t} < 0$$

Si en revanche le profit annuel est de 150, alors il faut investir et la VAN est

$$VAN(p = 150) = -\frac{800}{1,1} + \sum_{t=1}^{\infty} \frac{150}{(1,1)^t} = 772$$

Au total l'espérance de VAN en cas de décision en $t = 1$ est :

$$(1/2)0 + (1/2)772 = 386$$

Elle est supérieure à la VAN de la décision d'investissement immédiat ; il vaut mieux attendre un an, et l'acquisition d'information accroît la valeur de la décision future. La différence entre la VAN de la réalisation optimale et celle de la réalisation immédiate est appelée « valeur d'op-

tion » : c'est le supplément qu'un investisseur est prêt à payer pour avoir le droit d'attendre et de choisir le moment de réalisation, c'est en quelque sorte la valeur de la flexibilité.

Le modèle présenté est simpliste dans la mesure où l'incertitude se trouve dénouée totalement à une date déterminée. La réalité est mieux rendue par un modèle où l'incertitude se dévoile peu à peu, comme celui présenté en Annexe 6. Il en découle qu'en présence d'incertitude de ce type, la règle de l'égalisation de la rentabilité immédiate au taux d'actualisation doit être modifiée : l'investissement doit être réalisé pour une valeur plus élevée du taux de rentabilité, l'écart dépendant de l'importance des aléas. Le paradoxe de cet effet est que, conduisant à reculer les décisions (pour atteindre une rentabilité supérieure), il aboutit à des investissements qui, devant répondre à un trafic supérieur, sont également plus capitalistiques : là où en avenir certain on aurait fait dès maintenant une route à 4 voies, on fera dans 5 ans une autoroute (Quinet, 1974).

3.1.8.b) Les conséquences sur la répartition

En principe des compensations devraient résoudre les objectifs de répartition exprimés par la volonté politique. Dans la pratique, il n'en est pas ainsi, et les compensations auxquelles invite l'analyse théorique ne sont jamais opérées. D'ailleurs on ne dispose pas des outils qui permettent de les calculer. En effet, dans le calcul usuel du surplus, le surplus de l'utilisateur n'est qu'un intermédiaire de calcul, comme le montre bien l'encadré 7-1. En fait le surplus créé par l'investissement ne reste pas aux mains des usagers, mais se diffuse dans l'économie. Pour déterminer les vrais bénéficiaires de l'investissement, il faudrait disposer d'une modélisation complète de l'économie. Certains modèles visent à ce but, comme les modèles transport-utilisation des sols ; on a vu par exemple comment ces modèles permettraient d'appréhender les effets sur les PIB régionaux. Mais on peut être intéressé par d'autres décompositions que les découpages géographiques.

De toute façon les critères de compensation ne sont pas dépourvus de paradoxes et d'incohérences, surtout lorsque les transformations ne sont pas marginales ; enfin, les actions de répartition ont des effets sur l'efficacité, et, en toute rigueur, on ne peut pas séparer les deux problèmes, optimiser d'abord l'efficacité, et ensuite répartir les produits.

3.1.8.c) Les conséquences sur le développement économique

Les conséquences des améliorations des transports sur le développement économique peuvent être vues au plan national ou au plan régional.

Au plan régional elles recourent les préoccupations d'aménagement du territoire, expression d'une politique volontaire de répartition des activités sur le territoire. C'est ainsi que la France veut éviter la désertification des régions de montagne et assurer une certaine uniformité dans les densités de peuplement. De même l'Union Européenne affiche un désir de

resserrer les liens entre les pays centraux et les zones périphériques. Ces objectifs politiques passent par des moyens au rang desquels figurent les transports. La foi dans l'amélioration des transports pour fixer les activités ou mieux les attirer est très forte parmi les décideurs politiques ; elle donne lieu à l'élaboration de programmes d'infrastructure ambitieux tels que le plan routier breton, le plan routier Massif Central en France, la politique des Réseaux Trans-Européens au niveau de l'Union Européenne. On a vu aux chapitres II et III précédents les mécanismes alors à l'œuvre : l'amélioration des communications induit une polarisation des activités et est susceptible de créer des pôles de développement à partir d'investissements, qui concernent d'ailleurs non pas seulement les transports, mais l'ensemble des infrastructures publiques. Toutefois, cet effet est aléatoire, encore mal connu, et aussi, autant qu'on puisse en juger, très différencié selon la nature des activités et selon les modes de transport. Dans le cadre d'un projet particulier, on ne peut envisager de l'appréhender que sous forme qualitative, à l'aide par exemple de scénarios dont l'élaboration utilisera le savoir expérimental et théorique accumulé en ce domaine. En tout cas l'argument du développement économique local doit être pris avec une grande prudence, d'autant que lorsque, dans le cadre d'un programme national d'une ampleur donnée, on favorise une région, c'est au détriment d'une autre région. Lorsque l'on donne une priorité aux zones défavorisées on réduit les dotations des pôles de développement déjà existants, où elles auraient probablement en des conséquences plus fortes et plus assurées sur le développement économique. Il faudrait donc chiffrer le coût de ces objectifs politiques, ce qui ne serait pas aisé, et qui, en tout cas, n'est jamais fait.

En dehors de ces effets de localisation, les infrastructures de transport ont un effet économique national d'amélioration générale de la productivité, dont on a vu l'essence au cours du chapitre I. Est-il raisonnable, est-il possible de le prendre en compte dans les évaluations de projet ? Il n'est pas possible, en l'état actuel des connaissances, de le quantifier. Certes la quasi-totalité des évaluations macro-économiques ex-post des investissements publics, notamment de transport, font apparaître une rentabilité systématiquement supérieure à celle calculée ex ante au niveau micro-économique. Mais les écarts sont très variables d'une étude à l'autre, et insuffisamment précis pour permettre d'avancer un chiffre. Il semble bien, d'ailleurs, que l'importance de ces effets dépende du mode, et soit différente selon les secteurs économiques. Enfin on peut s'interroger sur le point de savoir si cet effet qui est par nature une externalité positive, doit être traduit prioritairement à travers la tarification ou à travers les choix d'investissements. Comme la tarification des transports est déjà, en général, inférieure au niveau optimal déterminé en l'absence de cet effet sur la productivité, la question reste ouverte de savoir s'il faut aller encore au-delà, et si oui, de celui des deux instruments qu'il faut privilégier : infrastructures ou tarifs.

3.1.8.d) Pour quelle collectivité calculer ?

Cette question devient de plus en plus pertinente au fur et à mesure que se développent les liens entre certains pays, par exemple au sein de l'Union Européenne. Les trafics qui circulent dans un pays comprennent une proportion croissante d'usagers étrangers qui créent des externalités dont certaines restent localisées dans le pays où elles sont créées (congestion, insuffisant recouvrement des coûts d'infrastructures), et d'autres reviennent dans le pays d'origine du trafic (pollution transfrontière). Courcelle *et al.* (1998) ont étudié ce problème d'un point de vue théorique, à partir d'un modèle à deux pays sur les territoires desquels circulent des trafics domestiques et internationaux. L'analyse du modèle fait apparaître plusieurs interdépendances. Par exemple, un changement de prix de transport dans le pays 1 a un effet direct sur l'utilité des consommateurs de la région 1 ; mais il a aussi un effet à travers les modifications des effets externes locaux ainsi qu'à travers les variations de recettes de l'État. Sur les consommateurs du pays 2, les effets passent par la pollution transfrontière, et bien sûr par le paiement des taxes proprement dit. Il en résulte des politiques d'investissement et de tarification différentes selon le degré de coopération entre les deux pays et selon la possibilité ou non de discriminer le trafic étranger en matière de tarification. En l'absence de coopération, si chaque pays peut discriminer, il se comportera différemment selon qu'il est plus soucieux de l'environnement (alors il y aura des taxes très élevées sur le trafic étranger pour l'éliminer, c'est la politique de la Suisse vis-à-vis du trafic de transit), ou de recettes fiscales (alors le pays a un comportement du type monopole à l'égard du trafic étranger, les taxes sont définies de manière à maximiser la recette). Si la discrimination des trafics n'est pas possible, le niveau de la taxe va être un compromis entre la taxation du trafic intérieur (conforme à l'optimum économique), et celle du trafic international. Quant au niveau d'investissement mis en œuvre par les pays, il ne prendra en compte que les avantages des nationaux.

Dans une optique de coopération, ce niveau d'investissement devra aussi prendre en compte les avantages qu'en retirent les étrangers. Il sera donc plus élevé. La tarification sera plus basse, car le comportement de monopole n'a plus lieu d'être. Mais la discrimination des taxes entre nationaux et étrangers reste souhaitable. L'optimisation globale entraîne l'existence de 4 taxes différentes, selon que l'usager est national ou étranger, et selon qu'il est dans le pays 1 ou 2. Les différences de taxes dépendent des différences dans les goûts et dans les fonctions de production de chaque pays.

3.1.9. Le champ de l'analyse multicritère

On argue souvent des lacunes du critère du bénéfice actualisé pour prôner l'utilisation de l'analyse multicritère, qui consiste à présenter au décideur les différentes conséquences de l'investissement projeté sous forme non agrégée, par exemple à travers des notes ou des indicateurs quantitatifs

sur chacun des postes, en laissant au décideur le soin de pondérer entre les postes retenus.

La pondération des conséquences de l'investissement est choisie par le décideur au lieu de l'être par les agents économiques en cause. En dehors des cas de mauvaise information ou des situations d'effets externes, ou de concertement collectif, il ne devrait pas y avoir de différence si le processus démocratique fonctionne correctement.

Par ailleurs l'analyse multicritère ne supplée pas à notre mauvaise connaissance des phénomènes, raison de la plupart des lacunes recensées plus haut (conséquences macro-économiques, effets sur le développement régional, impact sur la répartition) : il n'y a pas de raison que le décideur politique en sache plus que l'analyste sur ces sujets.

3.2. Les conditions de mise en œuvre des méthodes de choix des investissements

Ces conditions de mise en œuvre peuvent être appréciées à deux niveaux : quelle est la pratique des études ? Les études présentent-elles sur les décisions ? On analysera successivement ces deux points, en terminant par l'analyse des distorsions et modifications qu'entraînent les modalités de financement.

3.2.1. La pratique des études de rentabilité

Elle est très diverse d'un pays à l'autre et d'un mode à l'autre. C'est en matière routière qu'elle connaît la plus grande diffusion et le plus d'uniformité. Elle donne lieu à des circulaires de mise en œuvre très détaillées et précises, qui assurent l'uniformité des procédures de calcul, et donc la comparabilité des résultats d'un projet à l'autre, permettant leur classement. Des circulaires de ce type existent dans la plupart des pays de l'OCDE, notamment pour les investissements interurbains, les plus soumis aux calculs de rentabilité.

En France un effort de réflexion entrepris sous l'égide de Marcel Boiteux a abouti à un rapport général (Boiteux, 1996), qui définit les grands axes de la mise en œuvre des méthodes de calcul économique, la nature des indicateurs à calculer, et les principes de leur détermination. Sur ces bases le Ministère des Transports élabore des directives précisant pour chaque mode les procédures de calcul à mettre en œuvre. Ces directives sont établies de façon à tenir compte des spécificités propres à chaque mode et à rendre les méthodes suffisamment similaires pour que leurs résultats permettent des comparaisons intermodales, permettent par exemple de comparer la rentabilité d'un train avec celle d'une autoroute qui assure le même type de liaison.

Le rapport précité pose un certain nombre de grands principes, comme la préférence à accorder à un critère unique de choix plutôt qu'un recours à des analyses multicritères, source de subjectivité ; l'attention à porter au choix de la solution de référence ; la nécessaire adoption de valeurs uni-

taires de comportement de préférence aux valeurs tutélaires ; l'intégration des effets sur l'environnement sous forme monétaire. Le rapport fournit la liste des critères à retenir pour comparer les projets : bénéfice actualisé, taux de rentabilité... ; il fournit aussi les plus importantes des valeurs unitaires à utiliser, notamment la valeur du taux d'actualisation (8 %), celle des coûts d'environnement (bruit et pollution), de la sécurité (valeur du mort et du blessé), et des indications pour la détermination des valeurs du temps.

Les études économiques ont une place différente dans le processus de décision selon l'étape du processus. Traditionnellement on peut distinguer 3 étapes dans ce processus :

- Celle du plan à long terme, par exemple des Schémas Directeurs. Il s'agit alors de tracer l'avenir lointain, en définissant des axes stratégiques tels que par exemple l'importance donnée aux différents modes, le volume total de crédits à consacrer aux infrastructures. On conçoit que les calculs de rentabilité ne fournissent en ces domaines qu'un élément de la décision, qui peut obéir à d'autres considérations ; les résultats du calcul micro-économique de la rentabilité individuelle des projets doivent être contrôlés par des considérations plus macro-économiques sur les capacités de financement des budgets publics, sur les atteintes globales à l'environnement (conséquences globales sur le réchauffement planétaire), sur les conséquences en matière de développement régional ou de consommation de ressources rares telles que l'énergie.

- Celui de la planification à moyen terme. On connaît alors, par la phase précédente, l'état final des axes de transport de la zone étudiée. Ce qu'il faut alors c'est fixer leur date de réalisation et leur nature précise (tracé, largeur de voie...). C'est à cette phase que le calcul économique doit prendre la place la plus importante ; c'est d'ailleurs cette phase qui en général marque de façon irréversible la décision de réaliser.

- Au-delà, se présente la phase de réalisation, avec la mise au point du projet, et les enquêtes publiques. C'est alors le dialogue et la négociation avec les autorités locales et associations de riverains directement concernés par le projet qui prennent le dessus et qui interviennent dans les décisions. Le calcul économique joue un rôle plus réduit : le principe du projet a déjà été décidé et sera rarement remis en cause, même si les négociations le changent sensiblement et en réduisent la rentabilité : le coût politique, psychologique, parfois même économique (lorsque des études coûteuses ont été engagées ou lorsque des projets complémentaires sont déjà en cours de réalisation), de l'abandon serait plus élevé que l'achèvement d'un projet à la rentabilité désormais insuffisante.

Souvent, les modifications impliquées par cette dernière phase relèvent du syndrome Nimby (*not in my backyard*) selon lequel les riverains veulent bien d'une infrastructure, mais à condition qu'elle ne les gêne pas ; d'où des coûts liés à l'enterrement des voies, aux protections en faveur de l'environnement ou aux détournements de tracés. Ces coûts sont rarement prévus initialement, à l'étape de planification, car les études techniques et surtout les études d'insertion sont alors embryonnaires. Ces déviations

ENCADRÉ 7-11a

Les valeurs unitaires proposées par le rapport « Transport : pour un meilleur choix des investissements »

« L'évolution au fil des années des valeurs unitaires des avantages de qualité de service et de temps sera, conformément aux études et recommandations courantes, prise égale à celle de la consommation finale des ménages par tête. »

1. Valeur de la vie humaine et coût des blessés

Tué :	3 600 000 F
Blessé grave :	370 000 F
Blessé moyen :	200 000 F
Blessé léger :	79 000 F

2. Valorisation de l'effet de serre

Modes	Consommation énergétique en Kep (1)	Consommation en kg de carbone pour les carburants : (2) = 0,85 × (1)	Coût lié à l'effet de serre, en centimes (3) = (2) × $\frac{50 \times 100}{1 000}$
<i>Voyageur × km</i>			
Automobile en milieu urbain	0,062	0,053	2,4
Automobile en rase campagne	0,030	0,026	1,1
Train rapide ou express électrique	0,018	ε	0
TGV	0,012	ε	0
Avion	0,051	0,043	2,0
<i>Tonne × km utile</i>			
Route > 3 T de CU	0,062	0,053	2,4
dont maxicode	0,017	0,015	0,7
Train complet électrique	0,008	ε	0
Voie d'eau	0,010	0,009	0,4

pourraient probablement être évitées si la définition précise du projet et les processus de négociation avec les populations concernées étaient entrepris plus tôt dans la vie du projet, au moment de la phase de planification, et si la décision de réalisation n'était prise qu'à leur issue. Mais cette anticipation aurait un coût.

3. Valorisation de la pollution locale, en centimes

Modes	Rase campagne	Milieu urbain
<i>Par voyageur × km</i>		
Auto	3,8	8,0
Rail électrique	0,12	ε
Rail Diesel	0,8	1,1
Avion	1,6	/
<i>Par véhicule × km</i>		
Voiture	7,50	10,0
<i>Par tonne × km utile</i>		
Route	5,6	7,5
Rail électrique	0,09	0,09
Rail Diesel	0,6	0,8

Une mention spéciale doit être faite des investissements en milieu urbain. Les calculs de rentabilité y sont beaucoup moins fréquents, et peut-être faut-il en trouver l'origine dans le fait que les décisions sont très largement décentralisées auprès des collectivités locales. Le pouvoir central n'a son mot à dire que pour les rares voiries à niveau national (et encore, même là, les décisions et financements sont de fait partagées avec les collectivités locales) et pour l'octroi de subventions à certaines grosses opérations de transport collectif. Mais de toute façon les projets en milieu urbain présentent des spécificités particulières (Quinet, 1997), qui justifient que les méthodes générales du calcul économique utilisées en rase campagne ne leur soient appliquées qu'après adaptation. Il y a à cela deux catégories de raisons :

- des raisons d'ordre économique tenant à ce que les hypothèses de cette analyse sont moins bien satisfaites en milieu urbain où l'on rencontre :
 - des rendements croissants et des externalités d'agglomération,
 - des externalités négatives d'environnement fortes,
 - des effets de répartition importants.
- des interactions fortes entre les transports, l'utilisation des sols et l'évolution des systèmes urbains ; les effets se diffusent plus largement dans l'économie (cf. mécanisme de route foncière), et ce, à travers des mécanismes marqués par des effets de seuils, des non-linéarités, des irréversibilités et par des distorsions (tarifications, subventions imparfaites, biens publics gérés de façon non optimale).
- des raisons d'ordre institutionnel, tenant à ce que le processus de décision urbain met en jeu de nombreux acteurs munis de valeurs et d'objectifs sociétaux différents, et différents d'une agglomération à l'autre, impliquant des priorités différentes concernant les politiques d'urbanisme et les redistributions et répartitions entre citoyens. L'évaluation doit donc

être vue comme un instrument de dialogue favorisant l'éclairage du débat public, plus que comme un processus de calcul dont le résultat dicte la solution à adopter.

Ces considérations justifient que le calcul économique usuel soit, en milieu urbain, complété par une analyse des liens entre transports et développement urbain, notamment à travers l'effet des transports sur les valeurs foncières. Pour cela il est nécessaire de développer des modèles transports-utilisation des sols fondés sur une analyse économique de la rente foncière et des analyses de scénarios nourris de raisonnements économiques simples partant d'une identification des acteurs, avec capitalisation du savoir en matière d'interaction transport-urbanisme.

Enfin les calculs de rentabilité ne s'appliquent pas uniquement aux investissements d'infrastructure. Ils doivent aussi permettre d'évaluer l'intérêt d'opérations d'exploitation. Frybourg et Orselli (1996) ont analysé la rentabilité de l'opération Sirius, qui consiste en une action d'information des usagers des voies rapides parisiennes par des panneaux à messages variables indiquant les temps de parcours sur les itinéraires en cause. Des comptages particuliers et des enquêtes auprès des automobilistes ont permis d'évaluer les reports (essentiellement d'itinéraires) engendrés par ce système, et les gains de temps et de sécurité obtenus. Il ressort que la rentabilité de l'opération tient essentiellement aux gains de temps obtenus par les usagers qui se maintiennent sur l'itinéraire et qui paradoxalement n'utilisent pas l'information fournie. À noter que l'étude n'intègre pas les gains de fiabilité ni les effets d'apaisement des comportements. Le taux de rentabilité immédiate de Sirius est de l'ordre de 100 %. Les analyses d'expériences étrangères (voir Perrett *et al.*, 1996) aboutissent à des valeurs du même ordre, très supérieures à la rentabilité des investissements, ce qui indique que l'on a trop mis l'accent sur ces derniers et pas assez sur l'exploitation.

3.2.2. Les recommandations des études sont-elles suivies ?

Si les études sont effectuées dans les proportions inégales selon les secteurs et modes, leurs conclusions sont suivies d'effet dans des proportions plus inégales encore. En témoignent par exemple les taux de rentabilité des projets d'autoroutes ou de trains à grande vitesse des derniers programmes approuvés en France.

Nilsson (1991) a analysé l'usage fait des études coût-bénéfice en Suède. S'appuyant sur la comparaison ex ante- ex post de plus de 200 projets routiers, l'auteur conclut que l'analyse coût-bénéfice semble avoir eu un impact très faible sur la hiérarchisation des projets ; l'étude explore aussi l'incidence sur les décisions d'autres effets comme par exemple les conséquences sur l'emploi, et là non plus l'analyse statistique ne fait pas apparaître d'incidence. L'auteur avance quelques hypothèses expliquant cette situation : l'existence de fonds dédiés à certains types de travaux ; l'importance dans les classements des collectivités locales des critères d'équité et d'égalité de répartition des crédits ; enfin l'hypothèse que l'analyse coût-bénéfice peut avoir agi à titre préventif pour éliminer les projets les moins rentables.

ENCADRÉ 7-12
Les taux de rentabilité de récents programmes de TGV

Les 16 projets TGV (dont premières phases possibles)	Km de ligne nouvelle	Trafic en millions de voyageurs		Rentabilité (1)		
		Avant	Après		Finan- cière	Socio-éco- nomique
Aquitaine	480	14,7	20,1	5,4	7,5	10
dont Tours-Bordeaux	361	(14,7)	(19,7)	(5,0)	9,5	12,3
Auvergne	130	2,9	3,9	1,0	3,1	6,7
Bretagne	156	9,1	12,2	3,1	7,4	13,6
Est	460	8,4	14,5	6,1	4,3	8,8
Grand Sud	70	3,7	5,3	1,6	5	12
Interconnexion Sud	49	12,6	13,4	0,8	3,2	9,6
Liaison Transalpine	251	(11,4)	(19,1)	(7,7)	6	10
dont Lyon-Montmélian	107	8,2	11,8	3,6	8,5	14,7
Limousin	174	3,3	4,0	0,7	2,4	4,4
Provence	219	16,6	20,9	4,3	9,8	13
Côte d'Azur	132	5,1	7,8	2,7	8,4	11
Languedoc-Roussillon	290	5,8	9,5	3,7	6,1	9
dont Avignon-Montpellier	75	(5,8)	(7,1)	(1,3)	7,8	10,5
Midi-Pyrénées	184	2,3	3,3	1	5,5	6,5
Normandie	169	5,5	7,1	1,6	0,1	3
Pays de la Loire	78	7,5	8,4	0,9	5,4	7,7
Picardie	165	13,5	14,4	0,9	4,8	5
Rhin-Rhône	425	9,5	15,3	5,8	5,9	10,7
dont Mulhouse-Dole	190	(9,6)	(13,8)	(4,2)	8,4	13,9

(1) Ces taux résultent d'une première approche des projets et doivent être affinés avec le tracé définitif.

Source : Lettre de la SNCF, déc. 1992.

L'auteur remarque aussi que les coûts de construction ex ante étaient fortement sous-estimés, et ce résultat rejoint les résultats de Bruzelius *et al.* (1996), qui, analysant l'histoire de grands projets européens récents, constatent des sous-estimations de coût se situant fréquemment dans un intervalle de 50 % à 100 %, accompagnées de surestimation des trafics dans des proportions un peu inférieures, mais élevées aussi. Ils procèdent alors à une critique du processus usuel de décision concernant ces grands projets qui implique la plupart du temps un mélange de financement public et privé :

- absence d'une étude suffisamment précise avant la décision, entraînant des engagements prématurés et difficilement réversibles.
- accent mis trop exclusivement sur les aspects techniques des projets, au détriment des aspects économiques.

- préoccupation trop tardive à l'égard des effets externes des projets.
- Les personnes et groupes lésés sont trop tardivement impliqués dans le projet.

- pas d'analyse de risque.

- le contrôle du projet et les arrangements institutionnels permettant de l'assurer ne sont guère évoqués au cours de la préparation du projet.

Cette situation résulte, selon les auteurs, de ce que l'administration joue différents rôles dans la plupart des projets ; elles est à la fois le promoteur du projet, le garant des problèmes de financement et l'acteur de la satisfaction de l'intérêt général, notamment des problèmes d'environnement et de sécurité. Les conflits entre ces rôles diluent la responsabilité, et permettent à des groupes d'intérêt particuliers de capturer le processus. Pour y remédier, il faudrait développer des mécanismes renforçant la responsabilité (fixer des objectifs, des récompenses et des pénalités ; définir des institutions chargées de son exercice) et accroître l'intervention du secteur privé, notamment pour profiter de sa meilleure capacité à gérer les risques.

3.2.3. Le financement privé des infrastructures

C'est en général pour une autre raison que le financement privé des infrastructures est mis en place : l'augmentation des ressources consacrées aux investissements et la réduction du recours aux budgets publics, toujours en mal d'équilibre. Le principe en est simple : la concession de l'infrastructure est accordée à un opérateur extérieur. Le contrat passé avec l'autorité publique prévoit que le concessionnaire construit l'infrastructure, puis, pour la durée du contrat (qui varie entre 20 et 50 ans), l'entretient et l'exploite, moyennant un péage perçu sur les utilisateurs, et dont le produit sert à payer les dépenses de fonctionnement, à rembourser les emprunts émis pour la construction, et pour le reste à rémunérer les fonds propres de l'opérateur.

Les variations autour de ce principe sont nombreuses. Elles portent sur le statut de l'opérateur, qui peut être soit semi-public (cas de la majorité des Sociétés d'Autoroutes en France), soit privé (cas d'Eurotunnel, concessionnaire du Tunnel sous la Manche) ; elles portent aussi sur les éventuelles aides de l'État, qui peut participer à l'opération soit par des apports en nature (apport des terrains ou d'une partie des travaux, comme ce fut le cas pour plusieurs concessions de chemins de fer au 19^e siècle), soit par des subventions (cas des Sociétés d'Autoroutes en France dans les années 70), soit encore par des bonifications d'intérêt ou des garanties d'emprunt (utilisé pour certaines sociétés autoroutières françaises), soit enfin par des garanties de recettes ; elles portent enfin sur les modalités d'attribution de la concession, qui peut intervenir soit de gré à gré, soit à la suite d'un appel d'offre.

Le mécanisme du financement privé présente deux sortes de conséquences qui seront analysées successivement : les premières sont relatives aux conditions de réalisation des infrastructures en terme de coûts, délais... ; les secondes portent sur les programmes d'infrastructures eux-mêmes.

3.2.3.a) Les conditions de réalisation

La concession donne lieu à une série de contrats. Le principal relie le concessionnaire à l'autorité concédante, et les autres relient le concessionnaire aux banques, aux sociétés d'assurance, aux entreprises de construction de l'ouvrage, aux organismes chargés de certifier les travaux. Leur caractéristique commune est que les matières dont ils traitent sont soumises au risque et à l'asymétrie d'information. On distingue usuellement 3 sortes de risques, contre lesquels les contrats ne peuvent totalement se prémunir, car ils sont forcément incomplets, ne pouvant pas prévoir toutes les éventualités susceptibles de se produire au cours de la longue durée de la concession, nécessaire pour amortir les dépenses de premier établissement :

- le risque politique, tenant essentiellement aux décisions futures de la puissance publique : changement de réglementation (notamment des taux de péage), décisions concernant les infrastructures concurrentes, la fiscalité, les taux de changes pour les opérations à l'étranger,
- le risque technique : coût des travaux, niveau de qualité (sécurité, environnement) de l'infrastructure et de son exploitation,
- le risque commercial : risque sur le trafic, dont on a vu que son évaluation restait entachée d'incertitude, et donc sur les recettes.

Ces risques existent aussi bien sûr dans les procédures publiques, mais ils sont en quelque sorte occultés par l'intégration de toutes les opérations de programmation, financement, construction et exploitation, au sein de l'autorité publique, et par la faiblesse des contrôles et des sanctions, ce qui entraîne des phénomènes importants de dépassement des coûts et de surévaluation des trafics.

On peut espérer d'une gestion privée qu'elle exerce un contrôle plus rigoureux sur l'ensemble de l'opération, dans la mesure où les partenaires qui supporteront en dernier lieu les risques sont à la fois ceux qui sont les mieux à même d'en apprécier les conséquences et d'agir pour les maîtriser, et dans la mesure où les contrats seront bâtis de façon à les inciter à les maîtriser efficacement.

Ainsi les pouvoirs publics devraient être responsables des risques politiques et supporter tout ou partie des conséquences que leurs décisions peuvent impliquer. De même, les risques de coûts seraient du ressort des constructeurs, et les risques de marché sont le propre du concessionnaire.

Même si chacun de ces risques était totalement dans la main de l'acteur correspondant, il ne serait pas logique que cet acteur le supporte en totalité ; c'est certes le moyen permettant d'obtenir la meilleure incitation à l'effort. Mais il faudrait alors payer à cet acteur une prime de risque trop élevée liée à son aversion au risque et à l'incertitude.

En matière de coût de construction par exemple, certains risques peuvent être à peu près totalement maîtrisés par l'entreprise par exemple ceux qui concernent les parties hors sol ; mais d'autres, par exemple ceux qui concernent les ouvrages de fondation, comportent une part d'incertitude inévitable. Comme on le voit dans l'encadré, les premiers peuvent

être gérés par un contrat forfaitaire, donc sans risque pour le concédant ; pour les seconds en revanche, le concessionnaire devra supporter une part du risque (cf. encadré 7-13).

ENCADRÉ 7-13

Incitation et couverture du risque sur un modèle simple

Soit une entreprise, appelée dans la suite « agent », qui doit fournir à un commanditaire public, appelé dans la suite « principal », un service, dont la fabrication coûte :

$$C = Cp - e + \varepsilon$$

- ε étant une variable aléatoire de moyenne nulle et d'écart-type σ , dont la valeur est inconnue lorsque les deux parties signent le contrat,
- e étant l'effort fait par l'agent, que le principal ne peut pas observer,
- Cp étant une constante.

Le principal ne constate que C et ne peut distinguer dans sa réalisation ni e ni ε .

L'agent est rémunéré par le principal selon une formule du type :

$$R = \delta + \gamma C$$

L'objectif de l'agent est de maximiser son utilité U :

$$U = R - C - \psi(e)$$

- $\psi(e)$ étant le coût de l'effort pour l'agent. On suppose que : $\psi' > 0$ et que ψ'' est une constante positive.

Cette utilité est aléatoire, dépendant de la réalisation de l'aléa ε , et de moyenne :

$$\bar{U} = \delta + (\gamma - 1)(Cp - e) - \psi(e)$$

Mais l'agent étant par hypothèse averse au risque, avec un taux d'aversion de r , son espérance d'utilité ex-ante n'est pas \bar{U} , mais :

$$E(U) = \delta + (\gamma - 1)(Cp - e) - \psi(e) - (\gamma - 1)r\sigma^2/2$$

L'utilité du principal est :

$$Up = W - R$$

W étant la valeur accordée au service. Si le principal n'a pas d'aversion au risque, ce qu'on supposera, alors :

$$E(U_p) = W - E(R)$$

Le comportement de l'agent consistera à maximiser son utilité ex-ante, ce qui conduit à :

$$\psi'(e) = 1 - \gamma$$

et fixe donc e dès lors que le principal a fixé γ .

Le comportement du principal sera de maximiser son utilité, compte tenu de ce comportement de l'agent, soit :

$$\max E(U_p) = \max (W - \delta - \gamma(Cp - e))$$

Les variables à la disposition du principal étant : δ et γ , e étant déterminé par :

$$\psi(e) = 1 - \gamma$$

et avec la contrainte :

$$E(U) \geq 0$$

qui exprime que l'agent est disposé à participer. Le principal a intérêt à ce que cette utilité soit nulle, bien évidemment, et l'inégalité précédente peut être remplacée par une égalité.

Le résultat de cette optimisation est :

$$\gamma = 1 - 1/(1 + \sigma^2 r \psi')$$

$$\delta = (1 - \gamma)(Cp - e) + \psi(e) + (\gamma - 1)^2 r \sigma^2 / 2$$

et bien sûr :

$$\psi(e) = 1 - \gamma$$

On peut examiner le cas particulier où soit l'agent est neutre au risque ($r = 0$) soit il n'y a pas de risque ($\sigma = 0$). Alors,

$$\gamma = 0$$

$$\psi(e) = 1$$

$$E(U_p) = W - (Cp - e) - \psi(e)$$

On remarque que c'est alors l'agent qui supporte tout le risque. L'effort qu'il fournit est optimum d'un point de vue collectif, c'est celui qui maximise : $(E(U) + E(U_p))$. L'utilité espérée de l'agent est nulle.

Par comparaison, lorsque l'agent est averse au risque, l'effort est inférieur à l'optimum, le principal supporte une part du risque d'autant plus forte que l'agent y est averse, mais l'utilité espérée de l'agent reste nulle.

Le principal peut réduire l'écart par rapport à l'optimum s'il connaît une quantité y corrélée avec l'aléa. Alors il remplace R par :

$$\delta + \gamma C + \alpha y$$

et choisit α de façon à minimiser l'écart-type de cette nouvelle rémunération, ce qui aboutit à

$$\alpha = -\rho \sigma' \sigma y$$

relation dans laquelle :

σy est l'écart-type de y

ρ est le coefficient de corrélation entre y et e

Pour des raisons évidentes liées à l'incomplétude des contrats (impossibilité de prévoir toutes les éventualités), la couverture du risque politique n'est pas non plus possible en totalité.

Celle du risque commercial est encore plus difficile ; on a vu l'incertitude des études de trafic, et si le concessionnaire doit normalement être le mieux placé pour l'apprécier, il n'en reste pas moins que l'autorité publique, qui a souvent mené les études préliminaires, et qui a suscité la concession, en porte aussi une responsabilité. En outre, il n'est guère pos-

sible de s'assurer contre ce risque, si ce n'est parfois par un contrat d'achat à long terme de l'usage de l'infrastructure ; mais cette possibilité, qui se présente par exemple dans le secteur de l'électricité, n'est guère utilisable en matière de transport où la clientèle de l'infrastructure est atomisée¹ ; on devrait alors penser à certaines forme de garanties de recettes, fréquentes dans les concessions d'exploitation (par exemple dans les concessions de transports publics urbains), mais peu utilisées par les concessions d'infrastructures.

Les bénéfices que l'on peut tirer du financement privé en terme de conduite de l'opération peuvent diminuer ou disparaître si les contrats sont mal rédigés et s'ils ne gèrent pas le risque et le système de sanctions-récompenses de façon convenable. Donnons en quelques exemples.

Tout d'abord, la coopération public-privé n'est pas sans danger si elle est mal organisée. Ainsi d'un concessionnaire privé d'une infrastructure qui réalise lui-même l'investissement mais en sous-traite l'exploitation à une société où participerait l'autorité concédante ; le montage peut séduire dans la mesure où l'exploitation de l'ouvrage serait mieux coordonnée avec celle du reste du réseau de transport. Mais il en résulte une confusion de responsabilités génératrice d'inefficacité.

Les promoteurs du projet peuvent chercher à échapper au risque commercial en se retirant très tôt. Bien souvent ces promoteurs sont des associations d'entreprises de travaux publics et de banques, et les pertes éventuelles qu'ils pourraient subir sur la vente de leurs actions sont susceptibles d'être compensées par les profits qu'ils auront retirés de l'exécution des travaux et du placement des emprunts. Ceci conduit à penser qu'il est souhaitable, pour la bonne conduite de l'opération, que les promoteurs restent engagés durablement dans l'opération et ce, pour un capital propre représentant une fraction significative de l'actif (en général, dans les opérations de financement privé, les capitaux propres représentent de l'ordre de 10 % à 30 % de l'actif).

La présence de ces risques fait que le coût du financement privé est plus élevé que celui du financement public. Dans ce dernier cas la garantie de l'État permet d'obtenir des emprunts à taux minimum, comme c'est le cas pour la SNCF, entreprise qui serait en faillite selon les critères privés, mais qui dispose d'un excellent *rating* auprès des banques. Dans le financement privé, les emprunts sont effectués à un taux plus élevé, et le capital propre réclame une rémunération moyenne encore plus forte, en raison des risques plus élevés qu'il couvre en tant que dernier servi.

Le concessionnaire privé est en revanche plus efficace sur les coûts de construction. Ainsi, en France, il est apparu que ceux-ci étaient dans l'ensemble plus faibles de 5 % à 10 % que ceux obtenus par l'administration, la réduction de coût venant d'ailleurs pour une bonne part de la réduction

1. Toutefois, dans un contexte voisin, celui d'Eurotunnel, il est certain que le contrat d'achat de sillons par la SNCF et les ex-British Rail a constitué pour Eurotunnel une garantie de recette qui a facilité le placement des emprunts.

Les mécanismes du financement privé

Le financement privé met en jeu divers partenaires qui se distinguent par la priorité accordée à la rémunération de leurs fonds et par la nature de cette rémunération. Pour simplifier on distinguera les banques fournisseuses de fonds rémunérés selon un taux fixe et remboursés selon un rythme prédéterminé, et les actionnaires qui apportent un capital dont la rémunération, variable, est le cash-flow restant une fois que les banques sont servies.

Pour se prémunir contre le risque, les banques exigent que les recettes escomptées soient supérieures aux prêts qu'elles accordent, dans une proportion (variant d'habitude entre 1,5 et 2) qui dépend de l'incertitude qu'elles attribuent au projet.

Les ratios calculés pour évaluer cette proportion sont :

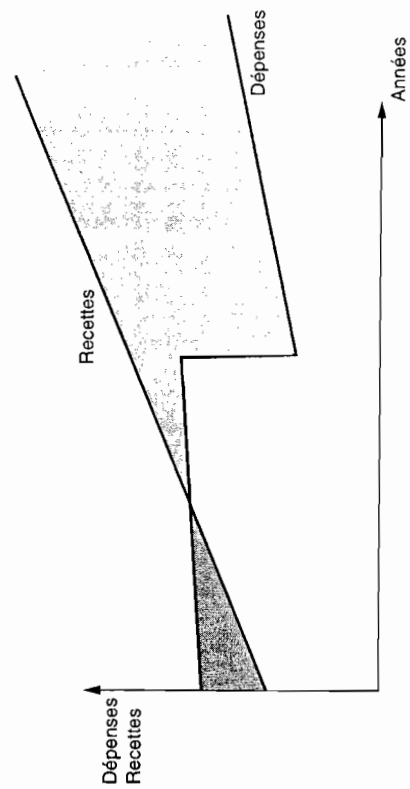
- Le « *Debt-Service Cover Ratio* » (DSCR) :
(*Cash-flow* de l'année *t*)/(Remboursement de l'année *t*)
- Le « *Loan Life Service Cover Ratio* » (LLSCR), calculé sur la durée des emprunts :

(Somme actualisée des *cash-flow* futur)/(Somme actualisée des remboursements futurs)

- Le « *Project Life Service Cover Ratio* » (PLSCR), analogue au précédent, mais calculé sur la durée de la concession.

Ces contraintes imposées par les banques influent sur la structure du financement, limitant le recours à l'emprunt, et rendent la rémunération du capital très sensible au niveau des recettes, à celui du DSCR exigé par les banques, et aux aléas.

FIGURE 7-8
Évolution de la trésorerie d'une concession



des dimensions de premier établissement, compensée ultérieurement par des dépenses d'entretien plus élevées ; ces dépenses interviennent alors à un moment où le concessionnaire a une trésorerie plus à l'aise.

En effet, la caractérisation d'une concession d'infrastructure est de procurer une trésorerie difficile au début de la concession et plus à l'aise, parfois même très abondante vers la fin, comme le montre la représentation schématique jointe.

Les recettes augmentent dans le temps comme le produit du trafic par les péages ; quant aux dépenses, elle sont assez stables¹ pendant toute la durée des emprunts, puis chutent brutalement, et il ne reste plus que les coûts d'entretien, qui sont peu importants comparés aux annuités d'amortissement.

Les risques dont il a été fait mention font que le bilan prévisionnel de trésorerie est entaché d'une grande incertitude, qui d'ailleurs ne se dénoue pas d'un seul coup à la mise en service : la connaissance des coûts s'améliore progressivement déjà pendant la construction, et les coûts d'exploitation ne se révèlent pas complètement à la mise en service. De même les trafics mettent un certain temps à s'ajuster à leur niveau définitif. Il y a ainsi une période de plusieurs années autour de la mise en service, où le bilan se révèle progressivement et où en général des besoins de financement nouveaux apparaissent. Leur satisfaction est le résultat d'un jeu entre le concessionnaire, les banques et les pouvoirs publics, le concessionnaire a besoin d'argent, et ni les banques ni les pouvoirs publics ne le lui refusent. Pour les banques, la faillite serait en général plus coûteuse en remboursement des dettes que le refinancement, d'autant que les besoins de refinancement arrivent non pas d'un seul coup, mais par petits paquets. Quant aux pouvoirs publics ils ont un intérêt évident au bon fonctionnement de l'infrastructure, qui engendre des surplus collectifs auxquels les opérateurs privés ne sont pas sensibles, sans parler des aspects politiques et psychologiques qu'aurait l'arrêt de son exploitation ou de sa construction ou des conséquences de la faillite d'un grand ouvrage. Leur action consistera par exemple à prolonger la durée de la concession. Ces mesures d'aide sont parfois insuffisantes ; cela peut se produire tôt dans la vie de la concession, lorsque les risques de coût et de marché sont trop défavorables. Cela peut aussi se produire plus tard, à l'occasion d'une crise macro-économique ou sectorielle (cas des chemins de fer à la fin du 19^e siècle, et dans les années 30). Alors l'issue est en général la nationalisation.

3.2.3.b) Le financement privé et les programmes d'investissement

De nombreuses considérations interviennent dans l'évaluation du financement privé et de ses effets sur les programmes d'investissement. D'abord, le péage qui frappe les infrastructures nouvelles est antiécono-

1. Sous réserve des conditions de remboursement du principal et de paiement des intérêts.

mique ; dans le cas le plus général, celui des routes, où les infrastructures nouvelles sont faites pour soulager des infrastructures anciennes saturées et sous tarifées, le péage les empêche alors de remplir cette fonction¹. Il faut aussi considérer le coût de perception du péage, qui est loin d'être négligeable. Dans le cas des autoroutes interurbaines en France, ce coût représente environ 15 % de la recette de péage.

Le financement privé a d'autres conséquences sur les programmes d'investissement à travers les attitudes qu'il induit de la part des concessionnaires existants ou potentiels vis-à-vis des pouvoirs publics. Tout d'abord, selon une problématique étudiée par Szymanski (1991), le financement privé incitera à modifier la date de réalisation du projet. Le financeur privé va en effet proposer une date de réalisation qui maximise son bénéfice actualisé, lequel diffère du bénéfice actualisé collectif, et lui est en général inférieur². Sous ce seul effet la date de réalisation serait repoussée. Mais si plusieurs financeurs privés sont en compétition, il y aura entre eux une course de vitesse dont il est assez intuitif que l'aboutissement sera une réalisation du projet lorsque la valeur actuelle nette privée est, non plus maximale, mais nulle. Alors, dans la plupart des cas, la date de réalisation sera avancée.

Enfin et surtout, le financement privé a pour effet de distordre la programmation des infrastructures. Cette conséquence est particulièrement claire lorsque les premières réalisations sont sorties de la période critique et procurent une *cash-flow* important. Ce *cash-flow* pousse alors, par sa disponibilité, à la réalisation d'autres infrastructures à péage dont la période critique initiale est ainsi facilitée. L'incitation au réinvestissement des *cash-flows* est d'autant plus forte que sont ainsi réduites les tentations que pourrait avoir la puissance publique de faire main-basse sur eux (le risque de *hold-up* de la littérature économique). On peut ainsi avoir une sorte d'effet de boule de neige qui permettrait de développer durablement un réseau d'infrastructure avec une soutenabilité financière satisfaisante, même si ces infrastructures n'ont pas une rentabilité suffisante.

3.2.3.c) Quelques exemples de financement privé

On présentera ici les cas d'Eurotunnel, des sociétés d'autoroutes interurbaines en France, et quelques formules originales de financement des infrastructures.

• Eurotunnel

L'ouvrage que constitue le tunnel sous la Manche est exceptionnel par son ampleur et remarquable par les modalités financières de sa réalisation.

1. Cet effet dépend de la répartition de propriété. Si le même opérateur possède et peut tarifier l'infrastructure concurrente, la structure du péage sera moins éloignée de l'optimum (voir Mills, 1995).

2. Pour la plupart des projets, la rentabilité collective est à peu près 1,5 à 2 fois égale à la rentabilité privée. Pour une analyse des relations entre les deux (voir Charmell, 1991).

tion. Il est, on le sait, constitué d'un double tunnel ferroviaire permettant de faire passer tantôt des navettes pour les passagers et les véhicules automobiles, tantôt des trains aux couleurs de la SNCF et des chemins de fer britanniques (Eurostar). L'originalité du projet tient à la nature de son financement, purement privé. Le concessionnaire a été d'abord constitué des promoteurs du projet (entreprises de travaux publics et banques françaises et britanniques), qui se sont ensuite effacés pour laisser place aux actionnaires particuliers pur de temps après la signature de la concession, en 1987.

Le besoin de financement total, estimé au départ à environ 50 GF (dont 28 GF pour les travaux proprement dits, le reste pour les frais généraux et les intérêts intercalaires) s'est progressivement élevé jusqu'à environ 105 GF, liés essentiellement à l'augmentation des coûts et à un retard dans la mise en service complète. Il a été assuré par des capitaux propres à hauteur de 26 GF et des crédits bancaires pour le reste.

L'augmentation des coûts s'est jointe à une moindre rentrée de recettes, liée à une mise en service retardée de plusieurs mois, à une montée en régime lente, due aux difficultés d'approvisionnement du matériel roulant, et aux problèmes de circulation des trains en Grande-Bretagne ; il en est résulté des ruptures de trésorerie en 1996, qui se sont résolues par un rééchelonnement de la dette. Auparavant des compléments de financement intervenus entre 1987 et 1994 avaient été facilités par une prolongation de 10 ans de la concession, dont la durée avait été initialement fixée à 58 ans.

D'un point de vue organisationnel, Eurotunnel, société franco-anglaise, est reliée par une série de contrats à de nombreuses entités :

- Aux États, par le traité de concession de mars 1986. Celui-ci comporte la réalisation et la bonne exploitation de l'ouvrage et ces conditions sont contrôlées par une Commission Inter-Gouvernementale.

- Aux entreprises de travaux publics par le contrat de construction (août 1986).

- Aux banques regroupées en un syndicat bancaire international comprenant environ 130 établissements de crédits par des conventions de crédits dont la première remonte à novembre 1987 et qui ont ensuite été modifiée à plusieurs reprises.

- Aux utilisateurs ferroviaires : SNCF et British Rail pour le passage des trains.

On peut tirer de nombreuses leçons du déroulement et de la mise en œuvre de cette expérience unique (Ponsolle, 1996) :

- Tout d'abord, la confirmation des aléas élevés qui frappent un projet de cette taille. Dans le cas présent ces aléas ont été accrues par l'insuffisance des études préalables, notamment en ce qui concerne le matériel roulant : difficultés de mise au point du matériel, exigences de sécurité imposées par la CIG et non prévues au départ, problèmes d'adaptation du matériel aux normes des pays.

- Ensuite, la nécessité d'un concessionnaire fort, suivant le projet dès le départ. Ici le concessionnaire a été mis en place par les promoteurs du

projet auxquels il s'est substitué pour diriger et contrôler les travaux et le financement, alors que ces promoteurs avaient déjà signé les contrats qui les liaient à l'État d'une part, aux constructeurs et aux banques d'autre part : le concessionnaire se trouvait vis-à-vis d'eux dans une situation de faiblesse évidente.

— Enfin, le caractère exceptionnel du financement de ce genre de projet, par son ampleur et par sa durée : les travaux durent de 7 à 8 ans, leur remboursement est lent, et un prêt bancaire usuel pour un financement de projet, de l'ordre de 15 ans, est inadapté. Le financement est facilité par une garantie de recettes, et de ce point de vue le contrat de longue durée avec les futurs exploitants ferroviaires était indispensable (ce contrat prévoyait à travers un montage complexe, une rétribution fixe à l'année et un péage en fonction des circulations de trains effectives).

- Le système des autoroutes interurbaines à péage en France

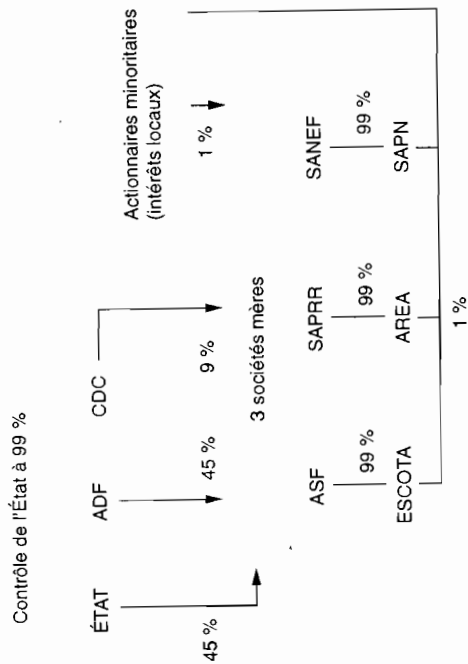
C'est au début des années 60 que sont apparues les premières autoroutes à péage, dont le mode de financement avait pour but de suppléer aux insuffisances des recettes budgétaires (« mieux vaut une autoroute libre qu'une autoroute à péage ; mais mieux vaut une autoroute à péage que pas d'autoroute du tout »). Le système s'est développé d'abord sous la forme de 5 Sociétés d'Économie Mixte qui se partageaient le territoire, et dont au début l'autonomie était très faible. C'était en fait de faux-nez de l'État, destinées à permettre la perception de péages pour le remboursement des emprunts, une opération interdite à l'État pour des raisons juridiques. Au début des années 70, le système fut libéralisé, c'est-à-dire que davantage d'autonomie fut accordée aux SEM, et que des concours pour des concessions nouvelles furent ouverts à des sociétés privées. Trois nouveaux concessionnaires privés apparurent ainsi. Deux tomberont rapidement en faillite à la suite d'erreurs dans les prévisions de trafic et à la suite du renchérissement des conditions de crédit ; elles furent reprises par des SEM qui, déjà anciennes, avaient passé leur phase critique initiale et disposaient d'un *cash-flow* suffisant. La troisième, Cofiroute, a passé la phase initiale, son réseau s'est étendu, comme celui des SEM. Le système des sociétés semble maintenant stabilisé sous la forme de 5 monopoles régionaux disposant de vastes ressources et d'un programme d'investissement mobilisant ces ressources, à travers un Schéma Directeur des autoroutes ambitieux.

Parallèlement le système des SEM s'est structuré en groupe à travers l'établissement public « Autoroutes de France » qui dispose de 45 % des parts des SEM (le reste des parts des SEM est détenu par l'État et par des collectivités locales), et qui a pour fonction de répartir les ressources de trésorerie entre les SEM et de procéder à une politique d'harmonisation des péages. Les SEM elles-mêmes sont regroupées en trois pôles. Un bureau d'étude commun, Sectauroute, agit comme maître d'œuvre des SEM, tout en développant ses activités sur le marché libre.

Les péages restent contrôlés par l'État. Leur taux, variable selon les sociétés et les sections, augmente un peu plus que l'inflation. Il est actuel-

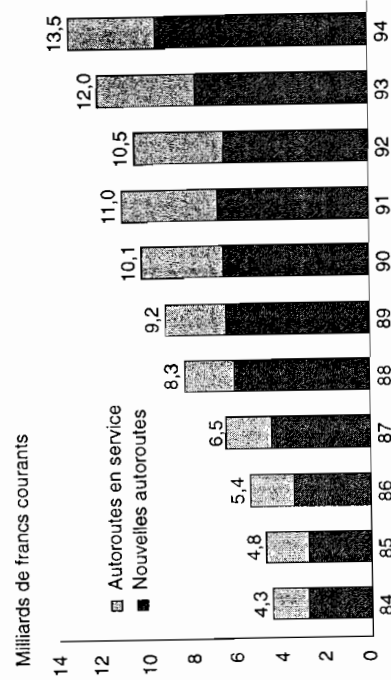
lement de l'ordre de 0,30 à 0,50 F par km pour les automobiles, et à peu près le double pour les poids lourds. La figure 7-9 présente la structure du système autoroutier à péage national et l'évolution des investissements en F et du réseau en longueur.

FIGURE 7-9
Structure du système autoroutier français
Organisation du système autoroutier

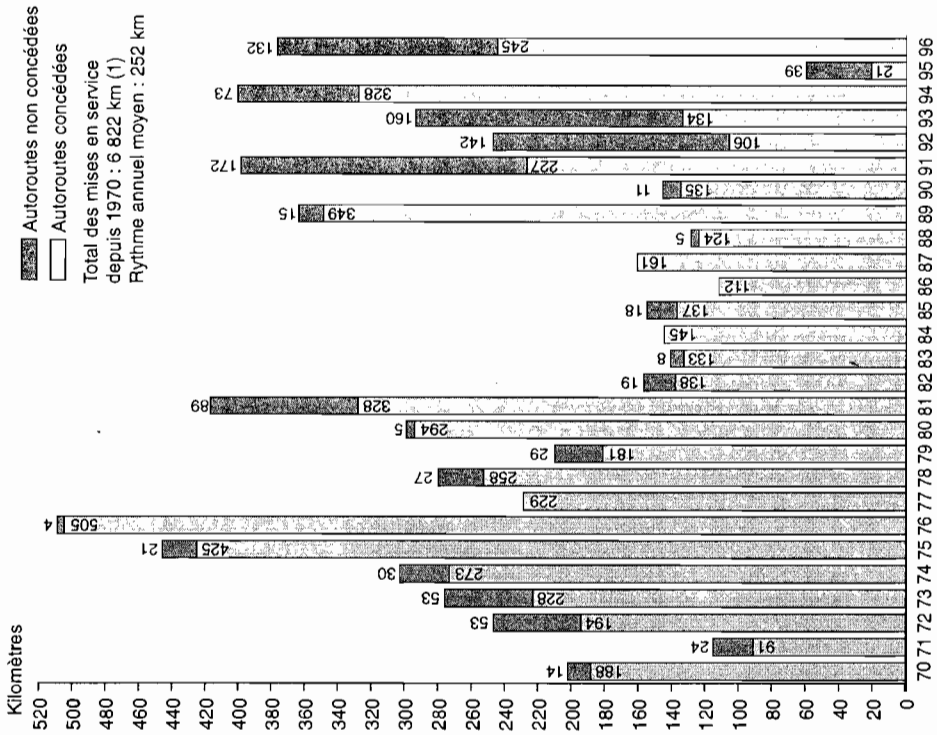


CDC : Caisse des Dépôts et Consignations
ADF : Établissement public « Autoroutes de France »

Investissements annuels des sociétés d'autoroutes



Mises en service des autoroutes de liaison depuis 1970



(1) Autoroutes de liaison au 1^{er} janvier 1996 = 7 396 km

• Autres formes de financement privé

À côté du système autoroutier interurbain, le financement privé se développe en France en zone urbaine sous des maîtres d'ouvrage locaux ou centrales (tunnel Prado-Carenage à Marseille, boulevard périphérique Nord de Lyon, A86 Ouest et A14 en région Ile-de-France...).

À l'étranger citons le système des péages fictifs, mis en œuvre il y a quelques décennies en Belgique, et qui réapparaît en Grande-Bretagne : l'autoroute est concédée soit de gré à gré, soit à travers un système d'appel d'offre ; aucun péage n'est installé, mais le concédant verse au conces-

sionnaire une redevance assise sur le trafic constaté multiplié par un taux de péage fictif.

Signalons enfin une idée voisine du financement privé par péage, qui consiste à trouver les ressources de financement dans les plus values foncières engendrées par la réalisation de l'infrastructure ; le système connaît deux modalités de réalisation. Dans la première c'est l'État ou la collectivité locale qui établit un impôt sur les plus values, sous la forme soit une taxe annuelle, soit d'une taxe à l'occasion des transactions. Ce système est utilisé aux USA et fournit un apport non négligeable pour le financement des projets. Dans la seconde modalité, c'est l'opérateur de transport lui-même qui est autorisé à effectuer les opérations immobilières et commerciales lui permettant de profiter des plus values créées par son infrastructure. Ce système est pratiqué notamment au Japon et fournit des ressources substantielles aux groupes dont dépendent les opérateurs privés. C'est aussi selon un principe analogue que les aéroports fonctionnent en bénéficiant de recettes commerciales qui constituent une part importante de leurs profits, et parfois subventionnent l'activité aéronautique, déficitaire.

4. Conclusion : les décisions publiques dans la pratique

Le corps de doctrine élaboré autour du concept de décision publique optimale est impressionnant par son élégance, sa cohérence, et par l'étendue des questions auxquelles il vise à répondre. Mais au-delà de ces vertus intellectuelles, la mise en œuvre dans le concret de ses recommandations reste incomplète. On a vu, pour chacun des deux grands domaines d'application que sont l'exploitation et le choix des investissements, combien la réalité des décisions était parfois éloignée des recommandations de la théorie.

On doit s'interroger sur les raisons de cet écart, soit pour que les économistes affermissent et précisent leurs recommandations, soit pour qu'ils orientent différemment leurs réflexions et leurs recherches.

Une première raison concerne la difficulté pratique de mettre en œuvre ces recommandations, notamment celles qui concernent la tarification : le coût marginal social présente une telle variabilité dans le temps et l'espace que son application rigoureuse est quasiment impossible, ou plutôt l'était tant que le péage électronique n'existait pas. Maintenant que ce système technique est de plus en plus au point, il devient d'autant plus facile de répercuter sur l'utilisateur la diversité des situations de coût, et on constate effectivement des débuts, lents, mais prometteurs, de différenciation des tarifs.

Une autre raison tient ici ce que les études d'optimum économique font appel à des concepts et à des calculs difficiles à comprendre et peu aptes à entraîner l'adhésion des décideurs, en général des hommes politiques généralistes, non spécialistes de la théorie économique et redevable des intérêts d'électeurs devant lesquels ils auront à se justifier.

Ensuite, ces études sont manipulables. Leurs résultats dépendent d'un grand nombre de paramètres connus avec incertitude, notamment les valorisations de biens non marchands : la valeur de temps, la valeur de la vie, le coût des externalités. Les quantités qui interviennent dans ces calculs ne sont pas non plus exemptes d'aléas ; qu'on songe aux incertitudes qui frappent les prévisions de trafic. Dans ces conditions, ce n'est même pas une intention consciente, c'est un instinct naturel qui poussera chaque fabricant d'étude à infléchir les paramètres de ces modèles dans le sens qui convient à ses intérêts, ce qu'il pourra faire en restant dans les limites du vraisemblable.

La manipulabilité est souvent accrue par les tendances qui consistent à admettre l'argument du développement économique local comme justification des projets. On a vu la portée et les limites de cet argument : il est vrai que, globalement, les infrastructures ont un effet sur le développement économique ; mais cet effet est tenu et variable d'un projet à l'autre. En outre, il mélange deux sortes d'effets : les effets d'efficacité annoncés par les promoteurs d'un projet arguant du fait que la nouvelle infrastructure contribue à la productivité de la collectivité ; mais l'effet réel du projet est en fait surtout un effet de répartition : il attire l'activité économique, au détriment éventuel de la zone voisine, beaucoup plus qu'il n'en crée ex nihilo.

Enfin, les processus de décisions en matière d'infrastructures répondent de moins en moins à la situation d'un despote bienveillant qui prendrait en toute connaissance de cause la décision optimale, situation qui inspire les recommandations du calcul économique ; elle est de plus en plus le fruit d'un processus de décision complexe où les acteurs multiples négocient le résultat. Les conflits d'intérêts et jeux stratégiques s'engouffrent dans l'espace de liberté que leur ouvre la manipulabilité des calculs.

On le voit bien à l'examen des difficultés qui se manifestent pour les changements de taxation : ainsi la taxe sur le gazole traîne désespérément derrière la taxe sur l'essence ; la taxe à l'essieu n'a pas été réévaluée depuis près de 30 ans, le péage urbain connaît des résistances considérables.

Mais les méthodes d'évaluation des projets n'ont pas non plus suivi les évolutions dans la nature du processus de décision. Elles restent trop exclusivement fondées sur l'analyse de l'efficacité et la recherche de l'optimum alors qu'elles devraient davantage se centrer sur un certain nombre d'aspects qui tournent autour de l'idée de répartition.

Il ne suffit plus en effet de savoir qu'un projet est globalement rentable, en supposant que le principe de compensabilité permet d'ignorer les pertes des uns à partir du moment où elles sont plus que compensées par les gains des autres. Il faut être capable d'identifier ces perdants, de savoir de combien ils perdent pour pouvoir éventuellement les compenser de façon équitable et désarmer leur éventuelle opposition, pour assurer l'acceptabilité sociale du projet, pour pouvoir si nécessaire mener en connaissance de cause les négociations que la bonne fin du projet implique. À côté

de ces besoins purement instrumentaux, le souci d'équité, de plus en plus fort dans nos sociétés, est un motif supplémentaire pour porter davantage l'attention sur les aspects de répartition¹.

Quand on parle du jeu des intérêts, il faut bien sûr comprendre non seulement ceux des personnes physiques, mais aussi ceux des personnes morales, c'est-à-dire des institutions, en général porteuses d'intérêts particuliers. C'est ainsi que les administrations responsables des différents modes de transports, qui présentent bien des spécificités, ont développé chacun des méthodologies d'évaluation des projets qui leur sont propres en adaptant à leur profit les principes théoriques, et ne manifestent aucun enthousiasme à plier leurs procédures à une méthodologie commune.

Ces particularismes sont accentués lorsque les modes participent chacun à des sources de financement spécifiques ; alors les calculs de rentabilité ne permettent que de hiérarchiser les projets à l'intérieur de chaque programme, mais ne servent plus du tout à fixer les volumes respectifs de ces programmes qui résultent de la plus ou moins grande abondance des fonds qui les alimentent.

Il faut peut-être voir là une des difficultés de mise en œuvre de solutions d'optimum premier appuyées sur la tarification au coût marginal social et sur le choix des investissements selon la rentabilité collective. Il n'y a guère d'institutions qui en assure la pérennité, à part celles qui seraient composées d'agents omniscients dévoués à l'intérêt général ; un organisme qui aurait pour tâche d'équilibrer recettes et dépenses pourrait théoriquement atteindre cet objectif en s'en tenant strictement à ces bases de gestion ; mais ce serait tellement plus facile pour lui de respecter l'équilibre budgétaire auquel il serait astreint en sous-investissant, ou, profitant de l'asymétrie d'information, de demander des subventions publiques en faisant état de la mauvaise rentabilité de ses investissements.

Un autre biais institutionnel est celui qui a été éprouvé notamment aux États-Unis et qui provient de l'attribution de subventions : les subventions d'investissements parfois très importantes dont bénéficiaient certaines collectivités locales les ont conduit à négliger l'entretien des ouvrages, leur réflexion leur coûtant moins cher que leur entretien.

Le financement privé est une dernière illustration d'implication des structures institutionnelles dans les choix. Conçu dès l'origine comme un moyen de soulager les finances publiques et de faire profiter les ouvrages publics de la créativité et de la productivité traditionnellement reconnues à la gestion privée, il se trouve souvent dévié de ces seuls objectifs par la pression des intérêts qu'il met en jeu, comme on l'a vu à l'occasion des exemples qui en ont été donnés.

1. Verhoef *et al.* (1994) ont étudié les arbitrages entre efficacité et acceptabilité sociale. Les conclusions en sont la reconnaissance des conflits entre les deux et la possibilité d'améliorer les situations actuelles par l'utilisation de permis négociables, la combinaison de mesures de « second best » et la mise en œuvre de taxes affectées.

Dans les situations les moins favorables comme celles que représentent Orlyval, le Tunnel sous la Manche ou certaines concessions autoroutières, le financement apparaît comme une loterie quant aux résultats financiers, en raison de l'incertitude sur les trafics. N'est ce pas au fond une forme d'impôt caché, spécialement indolore, sur le rôle duquel s'empres- sent de s'inscrire ceux qui croient en profiter, et qui finalement le paieront.

La conséquence de ces mécanismes de déviance que la théorie ne prend pas en considération, c'est bien sûr une perte économique, qui se manifeste sous plusieurs formes. La première est l'écart entre la rentabilité de l'investissement qu'on s'apprête à réaliser et la norme de rentabilité : une rentabilité insuffisante et on perd, car on anticipe indûment l'opération ; une rentabilité excessive et on a perdu dans le passé en la retardant par erreur.

La seconde, plus difficile à percevoir, est plus grave dans ses conséquences. C'est le mauvais choix de l'instrument d'intervention. De nombreux exemples peuvent en être trouvés, comme le faible recours à la tarification pour remédier aux effets externes, ou le choix de l'investissements pour pallier une demande élevée que l'on traiterait à moindre coût et plus efficacement par la tarification.

Comment éviter ce genre de situation et arriver à des décisions plus rationnelles ? Il faut réduire au maximum les manipulations auxquelles peuvent donner lieu les calculs économiques, et pour cela d'abord traiter explicitement l'incertitude alors qu'usuellement les calculs sont menés comme si tout était certain ; multiplier les contrôles et les audits de la part d'organismes indépendants, à la fois avant et après la décision.

Il faut aussi que le langage du calcul économique soit davantage admis par les décideurs, et pour cela que les procédures de calcul soient élaborées avec eux. Il faudrait que les calculs s'attachent davantage à évaluer les effets de répartition et à identifier les gagnants et les perdants. Il faudrait enfin porter plus d'attention aux dispositifs institutionnels à l'intérieur desquels sont utilisés les formules de calcul qu'on élabore avec un si grand soin, en ne se préoccupant pas toujours de savoir si la succession d'acteurs qui allaient ensuite les utiliser n'en dénatureront pas le sens.

Mais de toute façon, les décisions publiques d'infrastructures portent en elles une contradiction fondamentale, leur caractère irréversible et la permanence de leurs effets, effets dont on ne peut prévoir que ceux qui s'exercent à court terme, et dont notre perception devient de plus en plus floue à mesure que l'horizon s'éloigne.

BIBLIOGRAPHIE

- C. Abraham et A. Laure (1959), *Étude des programmes d'investissement routier*, Annales des Ponts et Chaussées, novembre.
- R. Arnott, A. de Palma et R. Lindsey (1994), « Welfare effects of congestion tolls with heterogeneous commuters », *Journal of Transport Economics and Policy*, May.
- A. Bernard (1983), *La congestion peut-elle constituer un mode de régulation ?*, mimeo, CGPC, Paris.
- M. Boiteux (1956), « Sur la tarification des monopoles publics astreints à l'équilibre budgétaire », *Econométrica*, p. 24.
- M. Boiteux (1995), *Transports : pour un meilleur choix des investissements*, rapport d'un groupe de travail présidé par M. Boiteux sous l'égide du Commissariat Général du Plan. La Documentation Française, Paris.
- D. Braess (1968), *Über ein paradoxon der verkehrsplannung*, Unternehmensforschung, vol. 12.
- N. Bruzelius, B. Flyvbjerg et W. Rothengatten (1996), *Big decisions, big risks*, Draft. Aalborg University.
- C. Charmeil (1991), « Le calcul de la rentabilité financière d'un projet constitue-t-il une bonne approche de la rentabilité publique ? Quelques réflexions sur l'exemple du tunnel sous la Manche », *Économie appliquée*, n° 2.
- A. Chin et P. Smith (1997), *Automobile ownership and government policy : the economics of Singapore vehicle quota scheme*, *Transportation Research*, part A, n° 2.
- Commissariat Général du Plan (1981), *Les choix d'infrastructure en période de croissance ralentie*, Rapport d'un groupe présidé par E. Malinvaud.
- Commissariat Général du Plan (1983), *Calcul économique et résorption des déséquilibres*, Rapport d'un groupe présidé par E. Malinvaud.
- C. Courcelle, B. de Borger et D. Swysen (1998), « Optimal pricing of transport externalities in a federal system : a theoretical analysis », *Communication au WTCR (Anvers)*.
- A. de Palma et F. Marchal (1998), « Présentation du projet METEOR », Université de Cergy-Pontoise.
- J.A. Dupuit, « De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, vol. 8, 1844.
- M. Frybourg et J. Orselli (1997), *Évaluation technique, économique et sociale de SIRIUS*, Rapport Ministère des Transports, Paris.
- X. Galègue (1996), « Irreversibilité de l'investissement et valeur d'option », *Revue d'économie politique*, sept.-oct.
- P. Goodwin (1996), « La mobilité induite par les infrastructures », Table ronde CEMT, n° 105.
- G. Hubbard (1994), « Investment under uncertainty », *Journal of Economic Literature*, Dec.
- S. Jara-Diaz (1990), « Consumer's surplus and the value of travel time savings », *Transportation Research B*, vol. 24.
- J.O. Jansson (1993), « Government and transport infrastructure Investment », in *European Transport Economics*, édité par J. Polack et A. Heertje, Blackwell.
- J.O. Jansson (1994), « Accident externality charges », *Journal of Transport Economics and Policy*, January.
- Y. Kanemoto, K. Mera (1985), « General equilibrium analysis of the benefits of large transportation improvements », *Regional Science and Urban Economics* 15, 343-363.
- J. Lesourne (1969), *Le calcul économique*, Dunod, Paris.
- G. Mills (1995), « Welfare and profit divergence for a tolled link in a road network », *Journal of Transport Economics and Policy*, 29.

- H. Morhing, *Transportation Economics*, Ballinger Press, Cambridge Mass., 1976.
- H. Morhing et M. Harwitz (1962), *Highway benefits: an analytical framework*, Evanston, III Transportation Center, Northwestern University.
- H. Morisugi (1983), « A basic definition of transport benefits : advocating equivalent variation », Proceedings of the WCTR Hambourg.
- H. Morisugi et Y. Hayashiyama (1997), « Post-evaluation of the Japanese railway network ; 1875-1940 », in Quinet et Vickerman, *The econometrics of major transport infrastructures*, MacMillan.
- D. Newbery (1988), « Road damage externalities and road user's charges », *Econometrica*, March.
- D. Newbery (1995), « Royal Commission Report on Transport and the Environment - Economic Effects of Recommendations », *The economic Journal*, September.
- D. Newbery (1995), « Reforming road taxation », Automobile Association.
- J.P. Nicolas (1998), « Circulation routière et évolution des nuisances sonores », Communication au WCTR, Anvers.
- J.E. Nilson (1992), « Second best problems in railways infrastructure pricing and investment », *Journal of Transport Economics and Policy*, ???
- J.F. Nilson (1991), « Investment decisions in a public bureaucracy », *Journal of Transport Economics and Policy*, May.
- T.H. Oum et M. Tretheway (1988), « Ramsey pricing in the presence of externality costs », *Journal of Transport Economics and Statistics*, September.
- T. Oum et Y. Zhang (1990), « Airport pricing : congestion tolls, lumpy investments and cost recovery », *Journal of Public Economics*, n° 43.
- A. de Palma (1992), « A game theoretic approach to the analysis of simple congested networks », *Transportation Economics*, vol. 82, n° 2.
- F. Papon (1996), *Méthodes innovantes de financement des routes de France*, Rapport INRETS, 1997.
- K.E. Perrett et A. Stevens (1996), « Review of the potential benefits of Road transport telematics », *TRL Report*, n° 220.
- Pindyck (1991), « Irreversibility, uncertainty and investment », *Journal of Economic Literature*, September.
- A. Polydoropoulou, M. Ben akiva, A. Khattak et G. Lauprete (1997), « Modeling revealed and stated en route travel response to advance traveler information systems », *Transportation Research Record*, 1537.
- P. Ponsolle (1996), « Le financement des grands projets d'infrastructure : l'expérience d'Eurotunnel », *Revue Transports*, sept.-oct.
- E. Quinet (1992), *Infrastructures de transport et croissance*, Paris, Economica.
- E. Quinet (1974), « L'incertitude dans les calculs et rentabilité routiers », *Annales des Ponts et Chaussées*.
- E. Quinet (1997), « Évaluation économique des projets de transports urbains », Rapport d'un groupe animé par E. Quinet ; Ministère de l'Équipement et ENPC-CERAS, Document ronéoté, Paris.
- E. Ramsey (1927), « A contribution to the theory of taxation », *Economic Journal*, vol. 37, March.
- K. Small et M. Rosen (1981), « Applied welfare economics with discrete choice models », *Econometrica*, 49.

- K. Small (1992), *Urban Transportation Economics*, Harwood Academic Publishers.
- D. Starkie (1994), « The US market in airport slots », *Journal of Transport Economics and Policy*, September.
- H. Varian (1992), *Microeconomic Analysis*, Norton International.
- E. Verhoef, P. Nijkamp et P. Rietveld (1994), « The trade-off between efficiency, effectiveness and social feasibility of regulating road transport externalities », note, Free University, Amsterdam, December.
- W.S. Vickrey « Pricing urban and suburban transport », *American Economic Review*, 59, 1963.
- S. Szymanski (1991), « The optimal timing of infrastructure investment », *Journal of Transport Economics and Policy*, September.
- H. Yang et H. Huang (1998), « Principle of marginal cost pricing : how does it work in a general road network ? », *Transportation Research*, vol. 32, n° 1.

Modalités et effets de la concurrence entre opérateurs

Les marchés des transports présentent une grande diversité de situation, qu'on peut analyser à travers les formes qu'y revêt la compétition entre opérateurs et les conséquences qui en résultent, à la fois en termes d'efficacité et d'équité.

1. Les formes de la concurrence

La concurrence s'exerce à des degrés divers à tous les stades de l'activité des transports. Mais c'est surtout aux stades situés en aval de la gestion des infrastructures qu'elle est développée, et ce sont donc ces stades, dits d'exploitation, qu'on prendra pour référence.

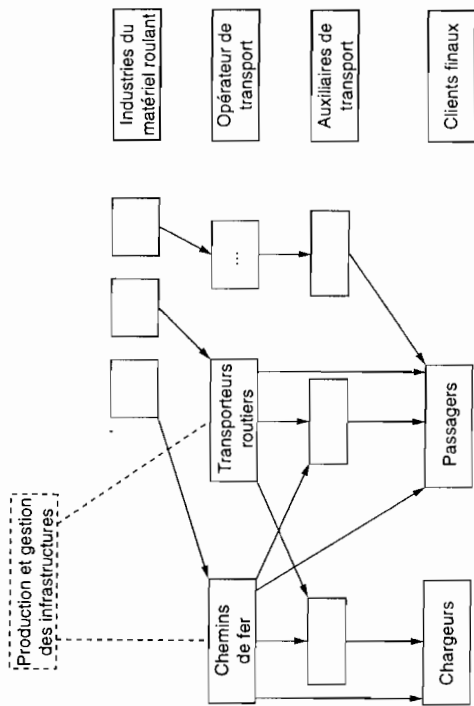
En développant et particulierisant le schéma des couches successives de production présenté au chapitre VI, on peut analyser le processus de production des services de transport comme une chaîne plus ou moins développée suivant les secteurs, selon un schéma du type de celui de la figure 8-1.

Comme le montre ce graphique, l'exploitation des transports fait intervenir plusieurs types de marchés :

- le plus important est celui qui relie les opérateurs de transports aux clients finaux, chargeurs pour les marchandises, et passagers ;
- mais entre aussi en jeu en amont le marché du matériel roulant ;
- en aval, les rapports entre opérateurs de transports et clients peuvent s'effectuer soit directement, soit par l'intermédiaire d'auxiliaires.

On analysera d'abord succinctement le marché du matériel et l'intervention des auxiliaires, avant de se concentrer sur le marché principal, celui qui relie les opérateurs de transports aux clients finaux.

FIGURE 8-1
Processus de production des services de transport



1.1. Le marché du matériel

Ce marché ne peut être considéré comme purement concurrentiel que pour les voitures et camions.

Dans le domaine de la construction aéronautique d'avions de plus de 100 places la concurrence est caractérisée par la présence de trois grands constructeurs : Boeing, qui représente plus de la moitié du marché, Mac Donnell-Douglas dont la présence est plus forte sur le marché militaire que sur le marché civil, et Airbus Industrie. La croissance des rendements, dus à l'importance des coûts de recherche et de développement, rend l'entrée sur le marché difficile, d'autant que l'entreprise installée bénéficie de l'effet d'apprentissage.

L'activité est soumise aussi à de forts frais de structures, correspondant aux réseaux commerciaux et aux services après-vente, et contribuant, comme le savoir-faire technique, à l'apparition d'économie d'échelle. Aussi une firme du secteur a intérêt à être présente sur l'ensemble du marché. La fragilité de Mac-Donnell-Douglas est liée à sa relativement faible taille et à l'étroitesse de son créneau. Le succès d'Airbus vient de ce qu'il a bénéficié d'un créneau (moyen courrier à fuselage large) laissé inoccupé par Boeing (Haudeville, 1994). Il vient aussi de choix technologiques hardies et judicieux dont Boeing n'avait pas pris le risque, et d'un soutien politique dont d'ailleurs son concurrent n'est pas dénué.

Enfin, les trois firmes développent des partenariats avec de nombreuses entreprises pour répartir les coûts fixes de développement. Les conséquences stratégiques de ces accords sont importantes : d'une part ils comportent un risque de transfert technologique permettant au parte-

naire d'entrer ultérieurement sur le marché ; d'autre part, ils jouent un rôle de prescription vis-à-vis des concurrents.

Golob (1995) analyse les effets de la déréglementation du transport aérien sur les politiques des constructeurs. Il conclut que les constructeurs ont répondu aux besoins nouveaux en changement rapide par un développement des procédures financières de leasing et en supportant une part accrue des risques, à travers des relations de partenariat.

Le marché du matériel de transport ferroviaire se présente sous des aspects différents en Europe et en Amérique du Nord.

L'Amérique du Nord constitue un marché unique, vaste, sur lequel opèrent plusieurs constructeurs, et où les acheteurs, c'est-à-dire les compagnies de chemin de fer du nouveau continent sont nombreux. Le marché a alors les caractéristiques d'un marché concurrentiel ; les produits proposés sont relativement peu différenciés, ils bénéficient des abaissements de coûts qui caractérisent les économies d'échelle et les courbes d'expériences de longues séries.

En Europe au contraire, au moins jusqu'à l'époque actuelle – les choses vont changer à la vitesse où s'installera la concurrence et où le marché européen s'unifiera –, les rapports entre constructeurs et opérateurs de chemin de fer s'effectuent sur une base nationale. Dans chacun des grands pays, on trouve face à face – ou plutôt côte à côte – un constructeur de matériel et un opérateur : Siemens et la D.B. en Allemagne, Ansaldo et les F.S. en Italie, GEC-Alsthom et la SNCF en France. On se trouve alors en présence de deux monopoles en chaîne. On sait que, dans un marché libre, la présence de deux monopoles en chaîne multiplie par deux les inconvénients du monopole et la perte économique résultant de la pratique de prix excessifs. Certes, dans le cas présent, le monopole de l'opérateur de transport est étroitement contrôlé par l'État, et les effets de rançonnement des monopoles ne s'exercent pas à plein, mais on ne peut pas s'empêcher de penser que l'absence de concurrence a pour effet d'augmenter les prix et les coûts. D'autres inconvénients apparaissent ; il y a sur le plan technique une très forte intégration entre les deux firmes ; cette intégration est certes justifiée par les fortes interactions entre le matériel roulant et les dispositifs constituant l'infrastructure : dispositifs de sécurité, alimentation en énergie, communication sol-mobile... Mais elle se traduit aussi par la possibilité et la tendance, de la part des techniciens, à intégrer très vite dans les productions les progrès techniques et les résultats de la recherche-développement. C'est un facteur de progrès technique, et effectivement les chemins de fer européens utilisent des technologies supérieures à celle des chemins de fer américains. Mais le rythme rapide du progrès technique ne permet pas de bénéficier des économies d'échelle et des courbes d'expérience, et cela d'autant plus que le marché est morcelé, chaque pays étant protégé non seulement par la préférence nationale accordée au constructeur local, mais aussi par les normes techniques bâties au fil des évolutions historiques, et qui portent sur la largeur des voies (supérieure dans la péninsule ibérique), sur les gabarits des

trains (largeur, hauteur totale, hauteur des caisses), sur les alimentations (voltage, fréquence), et sur les systèmes de communication et de sécurité.

On se trouve alors devant le dilemme soit de construire du matériel de petite série susceptible de circuler uniquement sur des réseaux de taille limitée, soit de prévoir du matériel capable d'emprunter des réseaux de normes différentes, donc plus coûteux, même si les séries sont plus longues. Il y a là aussi autant de raisons pour que les coûts soient augmentés au regard de ce qu'un marché concurrentiel et unifié quant aux normes, permettrait d'obtenir. L'Union Européenne essaie de combattre cette situation, et entend pour cela (voir Quinet, 1997) d'assurer l'interopérabilité des réseaux européens. Cette interopérabilité est spécialement souhaitable pour le réseau à grande vitesse et c'est surtout pour lui que se développent les actions de l'Union. Pour cela des programmes sont lancés, conjointement avec les compagnies de chemins de fer et les constructeurs de matériel :

- ERTMS (European Railways Traffic Management System),
- EUROCCAB,
- EUROBALISE,
- EURORADIO.

Cette interopérabilité est une condition pour le développement du trafic international, la réduction des coûts et l'amélioration de la qualité de service. Mais sa réalisation se heurte à de nombreux obstacles : le poids de l'équipement existant dont la durée de vie est longue, les spécificités des cultures des compagnies des différents pays, et enfin le fait que les différences de normes constituent une protection contre la concurrence étrangère pour les entreprises de matériel.

La fixation des normes est un équilibre délicat entre trop d'uniformité et trop de diversification. Elles devraient résulter d'une coopération entre les partenaires, mais cette coopération est difficile à obtenir en raison de la multiplicité des partenaires et de la diversité de leurs intérêts.

1.2. L'intervention des auxiliaires de transport

Les auxiliaires de transport effectuent des tâches multiples.

Pour les passagers, les auxiliaires, qui sont les agences de voyage, fournissent à leur client les modes à utiliser, les horaires, les tarifs permettant à leur client de satisfaire au mieux leur besoin particulier.

Pour les marchandises, sous des noms divers, les auxiliaires prennent en charge :

- le groupage des envois dans les véhicules
- l'organisation de la chaîne de transport et de la succession des modes utilisés

- les formalités de passage en douane pour le transport international. Ces auxiliaires peuvent être des opérateurs de transport qui effectuent eux-mêmes une part importante du transport. C'est le cas par exemple des grands transporteurs routiers, qui assurent le transport complet et se char-

gent de trouver la compagnie maritime éventuellement nécessaire, ou de traiter avec le chemin de fer. Ce peut être aussi des intermédiaires purs, ne disposant d'aucun moyen technique.

Le cas particulier du transport combiné est intéressant, puisque l'on y rencontre plusieurs types d'intermédiaires (Savy, 1997) :

- Novatrans, qui transporte des semi-remorques et des caisses mobiles. Ses clients sont des transporteurs routiers. Ce sont aussi ses actionnaires majoritaires, directement et par l'intermédiaire de la FNTR. Le réseau de Novatrans est concentré sur quelques grands chantiers. Novatrans opère à la fois sur le marché national et sur le marché international, et possède un peu moins de la moitié du marché total.
- La CNC, filiale de la SCETA et donc sous filiale de la SNCF, qui fait le transport des conteneurs internationaux et des caisses mobiles. Elle offre une prestation complète à ses clients qui sont les chargeurs ; jusqu'à il y a peu, CNC n'opérait que sur le marché national ; CNC dessert un réseau géographiquement plus étendu que Novatrans et gère un plus grand nombre de chantiers.
- Intercontainer, société internationale dont la CNC est l'actionnaire et le correspondant français, spécialisée dans le trafic international.

- La SNCF est sous-traitante de ces opérateurs pour la partie ferroviaire du transport ; à ce titre, elle gère et finance jusqu'à 50 % des investissements, sauf pour les engins de manutention. Les autres 50 % sont subventionnés par l'État. Quant aux engins de manutention, ils sont financés par les opérateurs.

Selon les clauses de la directive communautaire 91-440, le transport combiné international doit être ouvert à la concurrence, clause susceptible de concerner environ les 2/3 du transport combiné total en France.

Parmi ce total, le trafic de transit, qui représente à peu près la moitié du transport combiné total dans notre pays, semble le plus vulnérable : c'est probablement la partie du marché la plus rentable. La concurrence pourrait être le fait de regroupements de chargeurs ou de transporteurs routiers étrangers.

Le transport combiné, sur lequel beaucoup fondent de grands espoirs, est affligé de nombreux handicaps.

Handicaps techniques entraînant des coûts élevés. Le transport combiné cumule en effet les inconvénients des deux modes auxquels il fait appel. Ses coûts fixes spécifiques (20 % du coût total) s'ajoutent à ceux du chemin de fer, et la complexité de l'organisation qu'il nécessite en rend l'utilisation plus rigide encore. Du transport routier, il ne retient que la forme la plus coûteuse, celle de transports terminaux à courte distance et sans fret de retour dans des zones ayant un certain niveau d'urbanisation et particulièrement sensibles à l'environnement et à la congestion.

Ces caractéristiques expliquent que, compte tenu de la concurrence des autres modes, le marché du transport combiné soit très étroit : le transport combiné n'est compétitif que dans des conditions limitatives :

- transport à longue distance (plus de 500 km),
- transport par trains entiers, ou par train bloc avec un minimum de tri,
- envois massifs et réguliers,
- origines et destinations proches des chantiers.

Sa part de marché pourrait être plus grande si son développement n'était pas obéré par l'organisation du marché. Celui-ci met en jeu en effet une cascade de partenaires : l'opérateur ferroviaire, l'opérateur de transport combiné, le transporteur routier, éventuellement l'armateur.

Plusieurs niveaux de cette chaîne sont constitués de monopoles ou d'oligopoles, une situation particulièrement défavorable au développement des activités. En outre, ces monopoles et oligopoles n'échangent pas entre eux un produit bien défini ; chaque transaction est l'objet d'un marchandage, dans lequel chaque partenaire ne sait pas exactement les coûts et projet de l'autre ; il arrive dans cette situation que certaines transactions mutuellement avantageuses n'aboutissent pas. Enfin, les partenaires dans l'échange sont aussi concurrents (la SNCF cliente de l'opérateur de transport combiné pourrait aussi effectuer le transport en cause totalement par fer) ; ceci accroît la proportion d'échecs de certaines transactions.

Ajoutons que l'entrée sur le marché est obérée par la multiplicité des standards et par la nécessité d'investissements spécifiques. Ces considérations justifient l'intervention publique pour aider le développement et les aides aux investissements qu'il nécessite de la part des opérateurs.

Ces difficultés à l'échelle nationale sont multipliées lorsqu'on s'intéresse aux transports internationaux : les mêmes types de structures se rencontrent à peu de choses près de chez nos voisins. S'y ajoute le fait que les marchés et les politiques ne sont pas forcément les mêmes. Ainsi, le transport routier ne pose pas en Espagne des problèmes aussi aigus que chez nous, et la volonté d'y développer le transport combiné est moins forte ; en Allemagne, les activités économiques sont diffusées sur tout le territoire, donc peu favorables au transport combiné.

Il est probable que le transport combiné se développera surtout sur certains axes, à la faveur du franchissement ferroviaire d'obstacles naturels ou de goulets particuliers comme le Tunnel sous la Manche, les traversées alpines, la Vallée du Rhône...

Le recours aux auxiliaires n'est ni obligatoire ni systématique. Lorsque le besoin de transport est simple à satisfaire, le demandeur s'adresse directement à l'opérateur. Intuitivement, on sent bien que ce n'est que pour des opérations complexes qu'on aura recours à un auxiliaire. Il est intéressant d'analyser les raisons de ce recours, autrement dit de justifier l'existence des auxiliaires, et de préciser la nature de leurs fonctions. On peut leur trouver trois origines : la gestion de l'incertitude, la mise en œuvre d'économies d'échelle dans la gestion de l'information, enfin la substitution de la coordination hiérarchique au jeu du marché dans certaines opérations d'appariement.

L'auxiliaire de transport est d'abord un gestionnaire d'incertitude. Ainsi les agences de voyage, assurant la diffusion de plusieurs compagnies aériennes, prennent, par cette diversité, une assurance contre la faillite d'une compagnie particulière, qui serait leur seule proposition, un peu à la manière d'une détaillant non exclusif, qui s'assume ainsi contre la faillite d'un fournisseur dont il dépendrait totalement si c'était son fournisseur exclusif.

D'une autre manière, les auxiliaires, lorsqu'ils réservent une partie des places dans le véhicule d'un opérateur, jouent pour ce dernier le rôle d'un assureur lui procurant une ressource à terme, et justifiant ainsi l'octroi de conditions particulières. C'est ainsi qu'opèrent les Non-Vessel-Operator-Common-Carriers, intermédiaires du transport maritime, dont le rôle essentiel est d'acheter des emplacements de cargaisons aux armateurs,

pour les placer ensuite auprès de leur clientèle de chargeurs. Les *tour-opérateurs* jouent les mêmes rôles pour les personnes. Est ainsi remplie une fonction d'éclairage du futur permettant aux opérateurs de planifier leur production, comme guidés par un marché à terme.

L'autre fonction des auxiliaires est de fournir au chargeur un savoir technique sur les procédures de transport, qui sont pour lui trop peu fréquentes pour qu'il ait intérêt à en payer le coût fixe d'acquisition. Ce savoir peut être technique et porter sur les caractéristiques des véhicules de transport et des engins de manutention ; il peut être procédural et concerner par exemple les formalités de transit douanier ; il peut aussi ressortir à la capacité d'influence sur l'opérateur (dans certains pays, les auxiliaires sont utilisés pour assurer à l'envoi par chemin de fer une priorité suffisante) ; il peut enfin être d'ordre organisationnel et résulter dans la connaissance de celui des opérateurs qui est le mieux à même de satisfaire la demande précise du client ; c'est une des fonctions importante des agences de voyage que de permettre à l'utilisateur de s'y retrouver dans le maquis des horaires et des tarifs des compagnies aériennes. Il est à cet égard significatif de voir que les grands chargeurs ont créé des filiales qui opèrent comme auxiliaires de transport : la CAT et GEFCO sont des filiales de Citroën et Renault. Réciproquement les auxiliaires effectuent des opérations de commercialisation de produits qui ne sont pas dans le cœur de l'activité des transporteurs.

Enfin, les auxiliaires de transport assurent des optimisations qui pourraient certes être assurées par le marché, mais à des coûts supérieurs. C'est le cas essentiel des situations où la coordination des actions est fondamentale.

ENCADRÉ 8-1

Coordination et optimisation

Soit une fonction de coût à minimiser qui dépend de deux facteurs sous la forme $C(x_1, x_2)$ avec une proportionnalité obligatoire entre les deux facteurs, par exemple : $x_2 = \lambda x_1$

On peut par exemple penser aux véhicules et au carburant pour les faire rouler.

Le coût minimum est atteint par exemple pour x_1 et $x_2 = \lambda x_1$. Une erreur sur la détermination de ces variables, par exemple en prenant : $x_1 + dx_1$ et $x_2 + \lambda dx_1$, est du 2^e ordre en dx_1 .

Mais si l'erreur porte sur le rapport entre les deux facteurs, par exemple : $x_1 + dx_1$ et $x_2 = \lambda x_1$, l'erreur est du 1^{er} ordre en dx_1 .

Les auxiliaires jouent ce rôle de coordination dans la mesure où grâce à leur connaissance du marché ils peuvent orienter le client vers l'opérateur le mieux placé, et assurer l'appariement de l'offre et de la demande, par exemple en groupant de façon optimale les colis dans les véhicules en

fonction des tailles des colis, des disponibilités en véhicules des différents offreurs en volume et en délais, et des délais d'arrivée.

Aussi sont-ils un enjeu important, et les relations les concernant font l'objet de manœuvres stratégiques. Les compagnies aériennes prennent appui sur l'avantage que leur donne la possession d'un Système Informatisé de Réservation (SIR), pour passer avec les agences de voyage des contrats avantageux prévoyant l'exclusivité de l'usage du SIR avec intéressement sur les ventes (Dang Nguyen, 1995). Ces actions prennent souvent un tour anticoncurrentiel : manœuvres pour donner une priorité à la compagnie promoteur et prioritaire du SIR, imposition de clauses excessives sur la durée du contrat, les rémunérations, commissions déguisées...

À l'inverse, les auxiliaires peuvent étendre leur pouvoir à la fois en direction des transporteurs, et en direction des clients. Les stratégies à l'égard des transporteurs se rencontrent en particulier dans le cas du transport routier de marchandises ; des auxiliaires ou des transporteurs de large surface contribuent à établir de petits transporteurs, par exemple en leur facilitant l'obtention d'un prêt pour acheter le véhicule, et les fournissent en commandes à bas prix ; ils rejettent aussi sur le petit transporteur nouvellement et souvent imprudemment installé la gestion du risque de l'exploitation et d'éventuelles déséconomies d'échelle.

En direction de la clientèle, les auxiliaires, et les transporteurs importants, cherchent à développer leurs prestations au-delà du transport proprement dit. Les transporteurs de marchandises étendent leur activité vers la logistique : entreposage, étiquetage, opérations commerciales. Cette extension a pour résultat recherché de fidéliser la clientèle tout en profitant d'économies d'envergure entre ces activités et le transport. Quant aux agences de voyages, elles distribuent des voyages de tourisme complets, préparés par des *tour-operators* ; ces agences profitent des économies d'envergure permises par l'existence de leur guichet et font payer la coordination qu'ils assurent entre réservation d'hôtel et de moyen de transport.

1.3. Le marché direct entre les opérateurs de transport et les clients

Même si certains marchés se présentent à première vue comme des marchés de concurrence parfaite (transporteurs routiers de marchandises) ou de monopoles (transport ferroviaire), il est rare qu'une analyse plus fine ne fasse pas apparaître un oligopole : les transports ferroviaires sont en concurrence avec les autres modes, et en matière de transport routier, les opérateurs installés sur une relation donnée sont en général en nombre limité. La règle est donc l'oligopole, et il convient de regarder lequel des différents types d'oligopoles est le mieux apte à traduire la réalité de chaque marché de transport. Pour cela on rappellera les caractéristiques des oligopoles, puis on analysera leur application au transport.

1.3.1. Rappel sur les oligopoles

Il existe deux grandes catégories d'oligopoles. Dans l'oligopole de Bertrand, la variable de décision des opérateurs est constituée par les prix, et la demande va à l'opérateur qui offre le prix le plus faible. Dans l'oligopole de Cournot les variables sont les quantités mises sur le marché ; le prix se fixe en fonction de la demande et des quantités produites par les opérateurs.

L'étude classique de ces oligopoles, qui s'effectue sur le schéma simplifié des duopoles, revient à chercher l'équilibre de Nash correspondant. On en fera une présentation rapide, des développements plus complets figurent en Annexe 7.

Soient deux entreprises $i = 1, 2$, recherchant chacun à maximiser son profit. Le profit de chaque entreprise i dépend du choix qu'elle fait de la variable stratégique x_i , qui est à sa disposition et du choix que fait l'autre de la variable stratégique x_j , qui est à sa disposition :

$$\Pi_1 = \Pi_1(x_1, x_2)$$

$$\Pi_2 = \Pi_2(x_1, x_2)$$

Un équilibre de NASH est une situation telle que si on s'y trouve, aucun agent ne regrette sa décision :

$$\Pi_1(x_1^*, x_2^*) = \max_{x_1} \Pi_1(x_1, x_2^*)$$

$$\Pi_2(x_1^*, x_2^*) = \max_{x_2} \Pi_2(x_1^*, x_2)$$

- Dans le duopole de Bertrand la variable stratégique est constituée par les prix. La présentation la plus simple est la suivante : deux entreprises ($i = 1, 2$) ont chacune un coût marginal de production constant c_i (on suppose que $c_1 \leq c_2$). Elle choisissent un prix, et les quantités qu'elles vendent sont déterminées ainsi :

$$q_1 = D(p_1) \quad \text{si} \quad p_1 < p_2$$

$$q_1 = D(p_1)/2 \quad \text{ou} \quad p_1 = p_2 \quad \text{et} \quad c_1 < c_2$$

$$q_1 = 0 \quad \text{si} \quad p_1 = p_2 \quad \text{et} \quad c_1 = c_2$$

autrement

On démontre alors que le prix se fixe au niveau de c_2 , et la quantité vendue par chaque entreprise s'en déduit par les relations précédentes. La concurrence est sévère, elle annule le profit de l'entreprise la moins performante.

- Dans le duopole de Cournot, les variables stratégiques sont les quantités, et les profits s'écrivent :

$$\Pi_1(q_1, q_2) = q_1 p(q_1 + q_2) - c_1(q_1)$$

et la relation symétrique pour l'entreprise 2.

Si on suppose que chaque entreprise fixe sa quantité indépendamment de l'autre, la solution est donnée par les conditions de maximisation des profits, dont l'écriture est simplifiée en posant : $q_1 + q_2 = Q$

$$p(Q^*) + q_1^* \frac{dp}{dQ} - \frac{dc_1}{dq_1} = 0$$

$$p(Q^*) + q_2^* \frac{dp}{dQ} - \frac{dc_2}{dq_2} = 0$$

On verrait aisément qu'alors chaque entreprise fait un profit ; la concurrence est plus douce que dans le duopole de Bertrand.

Le plus fréquemment, dans les marchés des transports, la concurrence s'exerce à travers les prix – plus précisément les prix généralisés faisant intervenir le confort, le temps passé, etc. – : c'est ainsi que les usagers choisissent leur compagnie aérienne, que les chargeurs choisissent le transporteur de leur chargement etc.

Toutefois, le duopole de Bertrand suppose que les opérateurs ne sont pas soumis à des contraintes de capacité pour fournir le marché. Lorsque les entreprises s'installent sur le marché en choisissant d'abord les capacités qu'ils choisissent, et ensuite en se livrant une compétition par les prix, un résultat de Kreps et Scheinkman (1983) établit que le résultat de cette compétition est équivalente à celui d'un duopole de Cournot. Par ailleurs les produits offerts ne sont jamais rigoureusement équivalents. Lorsqu'ils sont substitués, la concurrence à la Bertrand s'adoucit.

L'idée générale qui résulte de ces analyses théoriques c'est que, lorsque les capacités de transport sont élevées ou peuvent être augmentées par transfert en provenance d'autres lieux, et lorsque les prestations ne sont pas différenciées, alors la compétition se fera par les prix et sera à Bertrand. Ce sera par exemple le cas pour le transport routier de marchandises.

Dans les situations inverses la concurrence est moins rude. Si, par exemple, la capacité de transport est fixe ou difficile à étendre, alors la compétition se fera par les quantités et à Cournot. Ce sera le cas du transport ferroviaire ou maritime, lorsque la ligne présente des spécificités techniques, ou encore celui du transport par voie navigable lorsque le bassin en cause est de taille limitée et isolé.

1.3.2. Application aux transports

On peut fournir quelques exemples d'application de ces principes généraux.

1.3.2.a) Concurrence rail-route

Le premier concerne les modalités de la concurrence dans les transports de marchandises entre la route et le rail. La compétition que se font les transporteurs routiers entre eux ressortit à une concurrence atomique

là où un grand nombre d'entreprise sont en présence. Lorsque le nombre d'entreprises est réduit, l'oligopole est du type Bertrand, c'est-à-dire, que par un processus qui est plus celui du tâtonnement que celui d'enchères institutionnalisées, le chargeur va aller au transporteur qui offre le prix le plus bas ; on peut considérer que les coûts marginaux sont constants et que la capacité n'est pas limitée, l'entrée sur le marché d'une liaison donnée étant particulièrement facile par déplacement des véhicules opérant antérieurement dans une autre zone. Le résultat est donc que le prix routier se fixe au niveau du coût marginal routier. Alors, pour le chemin de fer, de deux choses l'une : soit son coût marginal est supérieur à celui du coût marginal routier, et il est exclu du marché ; soit son coût marginal est inférieur, et c'est lui qui opère sur le marché, éliminant la route, mais en pratiquant un prix approximativement égal à celui du coût marginal routier. Au total, le prix pratiqué est toujours égal au coût marginal de la route, quelque soit l'opérateur qui finalement prend le marché¹.

1.3.2.b) Concurrence Eurotunnel-Ferries

L'autre exemple est celui de la concurrence entre les opérateurs assurant la traversée de la Manche, c'est-à-dire les ferries avant 1995, et maintenant les ferries et Eurotunnel. Les deux modes de transport assurent la traversée pour les voitures, les camions et les piétons ; les ferries exploitent de nombreuses autres relations port à port, dont la plus importante est Calais-Douvres, la liaison assurée par Eurotunnel.

L'analyse du comportement des opérateurs en présence a été effectuée par Szymanski (1993 et 1996). Le premier texte analyse le comportement des ferries dans les années antérieures à la mise en service d'Eurotunnel. L'autre, la concurrence entre les ferries et le tunnel.

Depuis longtemps le marché des ferries est dominé par deux opérateurs, *P and O European Ferries* et *Sealink*, et peut être étudié en première analyse comme un duopole. Les deux compagnies pratiquent une différenciation des tarifs, selon des groupes de compositions similaires et la comparaison de ces tarifs fait apparaître deux faits frappants : d'abord dans chaque groupe les tarifs des deux opérateurs sont très voisins et évoluent de façon similaire ; ensuite, à partir de 1986 les tarifs des deux concurrents se sont encore plus rapprochés et ont connus une vive croissance jusqu'à la mise en service d'Eurotunnel.

La similitude des prix des compagnies pour chaque catégorie de trafic confirme ce que l'analyse du processus de concurrence fait apparaître, à savoir que la compétition se fait par les prix et non par les quantités. Toutefois, on ne peut pas analyser la compétition de l'époque entre les deux

1. Cette analyse suppose que le prix est la seule variable de choix du chargeur. En fait, les délais, la fiabilité, etc. interviennent aussi, et c'est en termes de prix et de coûts généralisés qu'il faudrait raisonner. En outre, les biens ne sont pas totalement substituables, ce qui adoucit la concurrence (voir Salini, 1995).

opérateurs comme résultant du duopole de Bertrand pur : les prix se fixent au-delà du coût marginal, couvrent les frais fixes et rapportent un profit non négligeable aux opérateurs. On ne peut pas attribuer ce fait aux limites de capacité : la capacité installée des opérateurs a toujours été de 2 à 3 fois supérieure à la demande. Cette capacité excédentaire est certainement due à la nécessité de faire face à des pointes ; mais elle a aussi un caractère stratégique : c'est un moyen de dissuader l'entrée de concurrents en leur montrant que les entreprises installées ont des coûts marginaux bas et peuvent donc supporter une guerre des prix, et en leur faisant apparaître qu'on n'est pas dans une situation où, par la contrainte de capacité, l'équilibre devient du type Cournot et laisse un profit.

L'existence de profit montre une certaine entente entre opérateurs, résultat, dans la ligne du « *Folk theorem* » (voir Annexe 7), de ce que leur relation est de longue durée. Toutefois, la collusion complète n'est pas possible, car si les profits étaient trop élevés, d'autres firmes pourraient entrer sur le marché et tirer un bénéfice suffisant de leur intervention, avant qu'une réplique des entreprises en place ne les obligent à s'en retirer. En outre, la Commission de Contrôle des Monopoles du Royaume-Uni (*Monopoly and Merger Commission*) interviendrait.

Lorsque la décision de réalisation du tunnel est prise, ces facteurs jouent moins :

- l'horizon du jeu devient fini, les probabilités de collusion se réduisent,
- la perspective prochaine du tunnel et les perturbations qu'elle entraînera sur le marché rendent l'entrée moins attrayante pour un compétiteur éventuel.

Ce dernier aspect a été modélisé par Szymanski (1993), selon un schéma qui peut être résumé de façon simplifiée ainsi : les duopoleurs font face à un régulateur qui peut soit intervenir soit ne pas intervenir. S'il intervient, cela lui coûtera en coûts de transaction diverses L , et le prix passera du prix constaté p à un prix voisin du coût marginal c ; la réglementation est valable pour n années : la fonction objectif F du régulateur dépend de l'écart entre p et c , qui traduit le gain de bien-être collectif, de n et de L , selon une formule du type :

$$F = nf(p - c) - L$$

Le régulateur n intervient que si F est positif ; il en résulte une valeur limitée p^* de p , prix pratiqué par les duopoleurs qui déclenche l'intervention du régulateur. p^* est une fonction croissante de n ; n se réduit lorsqu'on se rapproche de la mise en service du tunnel car alors une régulation naturelle sera en place et rendra inutile l'intervention de la MMC. Donc les duopoleurs peuvent augmenter leur prix sans craindre d'intervention du régulateur.

Après la mise en service d'Eurotunnel, la situation change complètement : on se trouve en présence de deux catégories d'opérateurs : l'un a un faible coût marginal, l'autre (pour simplifier en regroupant tous les ferries en un seul) un coût marginal élevé. La situation présente des similitudes avec celle de la concurrence rail-route, mais il n'est pas possible de ne pas tenir compte de la différence des qualités de service, au moins pour les passagers et les automobiles : les délais, la fiabilité, l'agrément de la traversée sont différents, et les ferries assurent plusieurs rela-

tions port à port alors que le tunnel n'en assure qu'une entre Calais et Douvres. Szymanski (1996) la modélise de la façon suivante, en prenant pour exemple le marché des passages par automobile.

Soit :

- X : le nombre total de passager supposé donné
- S_T et S_F : les parts de marché du tunnel et des ferries
- p_T et p_F : les prix
- c_T et c_F : les coûts marginaux
- F_T et F_F : les frais fixes
- Π_T et Π_F : les profits

On suppose que :

$$S_T = A + e \log(p_T/p_F)$$

On en déduit l'équilibre du jeu, en maximisant Π_T et Π_F , respectivement par rapport à p_T et p_F :

$$-e(p_T - c_T)/p_T + \log A + e \log(p_T/p_F) = 0$$

$$-e(p_F - c_F)/p_F + 1 - \log A - e \log(p_T/p_F) = 0$$

L'auteur effectue ensuite des simulations sur la base d'estimations des coûts marginaux ($c_T = 15$ £, $c_F = 75$ £) plusieurs valeurs plausibles des paramètres A et e . Les tarifs optimaux qu'il trouve sont les suivants :

	Réal (1996)	Scénario (1) A = 0,25 e = 0,8	Scénario (2) A = 0,25 e = 1,0	Scénario (3) A = 0,75 e = 0,8	Scénario (4) A = 0,75 e = 1,0
Prix tunnel	120	81	60	126	85
Prix ferry	100	133	100	119	91

On voit que le modèle fournit des valeurs qui, pour un balayage assez large des paramètres, englobent bien les prix réels. Les paramètres qui donnent un meilleur accord avec les prix réels correspondent à une plus grande attractivité du tunnel (A plus élevé), ce qui est conforme à l'instruction. Les résultats suggèrent que les prix d'Eurotunnel, comme ceux du ferry, sont peut être un peu élevés ; la position prise par les opérateurs est compréhensible, car psychologiquement il est plus facile vis-à-vis de la clientèle de baisser un tarif que de l'augmenter.

Après la mise en service du tunnel, on est donc devant une forme adoucie de compétition à la Bertrand. Dans cette situation, les deux types d'opérateur subsistent. On peut s'interroger sur le point de savoir si Eurotunnel n'a pas intérêt à éliminer ses concurrents. Deux arguments s'y opposent : d'abord le fait que l'entrée est facile et que les concurrents éliminés pourraient facilement être remplacés ; ensuite en cas de succès, le danger de voir un régulateur intervenir pour contrer les excès éventuels du monopole. La collusion est un risque plus sérieux. Les plus grands gagnants y seraient les ferries, dont la marge de profit par rapport au coût marginal est la plus faible, et risque de couvrir à peine les coûts fixes. Mais précisément pour cette raison, on ne voit pas clairement l'intérêt qu'y trouverait Eurotunnel.

1.3.2.c) Concurrence entre opérateurs sur de mêmes liaisons

Le troisième exemple est celui de la concurrence entre compagnies aériennes sur des liaisons communes : Brander et Zhang (1993) ont analysé ainsi la concurrence entre American Airlines et United Airlines. Ils calculent statistiquement le coefficient de réaction de chaque firme, c'est-à-dire le rapport entre une variation des quantités produites par le concurrent, et la variation qui en résulte pour les quantités produites par la firme elle-même. Ce coefficient est de -1 pour le duopole de Bertrand, 0 pour celui de Cournot et q_1/q_2 pour la collusion entre les deux opérateurs (cf. Annexe 7). Les auteurs testent également la stratégie du jeu dynamique consistant pour chaque compagnie à coopérer (attitude de collusion) tant que l'autre en fait autant, puis à punir si l'autre abandonne la coopération. Le jeu dynamique aboutit à des coefficients de réaction qui sont égaux à -1 en phase de punition, et sinon compris entre -1 et ceux de la collusion (car, en raison de l'incertitude, la punition ne se fait que lorsqu'un seul de baisse des profits est atteint). Dans le cas du duopole de Cournot les valeurs de -1 sont remplacées par des valeurs de 0 .

Les résultats de l'étude montre que, si l'on prend comme hypothèse que les jeux successifs sont indépendants, les coefficients de réaction moyens indiquent une concurrence plus proche de Cournot que de Bertrand : ils sont de l'ordre de $0,40$. Leur variation au cours de la période donne à penser que des phases de collusion alternent avec des phases de guerres de prix.

1.3.2.d) La concurrence en réseau

Les exemples qui viennent d'être présentés sont simplistes en ce qu'ils ne prennent pas en compte les spécificités résultant de la structure en réseaux des opérateurs. Cette structure en réseaux, provenant de ce que les opérateurs exploitent simultanément plusieurs liaisons géographique, entraîne des interdépendances à la fois au niveau de la demande et au niveau de l'offre. Ces interdépendances se manifestent à des degrés divers dans tous les modes de transports ; mais là où elles jouent de la façon la plus complète, c'est dans les transports aériens. Aussi est-ce en prenant cet exemple qu'elles seront présentées.

• Effet de la taille du réseau

Les effets de la taille du réseau sur la demande peuvent s'exprimer en termes du nombre et de la structure des liaisons desservies¹. Ils peuvent aussi s'exprimer en termes de compatibilité des horaires : deux liaisons AB et BC peuvent être desservies par la même compagnie ; mais si les horaires correspondent pas, la présence simultanée de ces deux dessertes n'apporte pas grand chose à l'utilisateur. La taille du réseau constitue un avantage

1. Curien et Dupuy (1997) et Dupuy (1985) analysent ces structures et définissent les indicateurs permettant d'en mesurer les propriétés : connectivité, connectivité...

évident pour l'utilisateur : lorsque la compagnie dessert un réseau étendu, par besoin de changer de compagnie pour atteindre une destination lointaine, plus grande facilité d'information, compatibilité des horaires mieux assurée, plus de facilité pour modifier le voyage si on change de destination... La stratégie des compagnies en matière d'extension de leur réseau ou de mise en compatibilité de ces réseaux peut revêtir plusieurs aspects, allant de la coordination des horaires jusqu'à la fusion, en passant par des accords sur l'acceptation réciproque des billets.

L'analyse dont les éléments sont résumés dans l'encadré joint montre que la compatibilité est en général globalement favorable, mais qu'elle bénéficiera surtout à la plus petite compagnie.

ENCADRÉ 8-2

La compatibilité des réseaux

Supposons que deux entreprises soient en compétition sur une ligne donnée, et possèdent par ailleurs un réseau de taille x_A pour l'entreprise A, x_B pour l'entreprise B, et que leur coût marginal soit nul. Supposons aussi que la taille du réseau x présente pour l'utilisateur de la ligne une utilité $u(x)$. Alors chaque compagnie qui tarifie la ligne en cause au prix p offre à l'utilisateur de cette ligne un coût généralisé :

$$P = p - u(x)$$

Si les entreprises se livrent une compétition par les quantités ou capacités q_i , avec égalité du prix généralisé, et si la demande sur la ligne est :

$$P_A = P_B = 1 - q_A - q_B$$

on en déduit :

$$\Pi_i = q_i(1 - q_A - q_B + u(x_i))$$

les quantités d'équilibre sont :

$$q_B = (1 + 2u_B - u_A)/3$$

$$q_A = (1 + 2u_A - u_B)/3$$

Si le réseau de B est plus grand que celui de A :

$$q_B > q_A$$

Si maintenant les entreprises rendent leur réseau compatible, la taille commune est :

$$x_A + x_B$$

$$u_{A+B} > u_A + u_B$$

les nouveaux trafics d'équilibre sont :

$$q_A^* = q_B^* = (1 + u_{A+B})/3$$

On voit que :

$$q_B^* - q_B < q_A^* - q_A, \quad \Pi_B > \Pi_A$$

$$\Pi_A^* = \Pi_B^*, \quad \Pi_A^* > \Pi_A$$

mais on ne peut rien dire sur Π_B et Π_B^* .

La petite compagnie gagne à la compatibilité ainsi que les usagers ; mais on ne peut rien dire pour la grande entreprise.

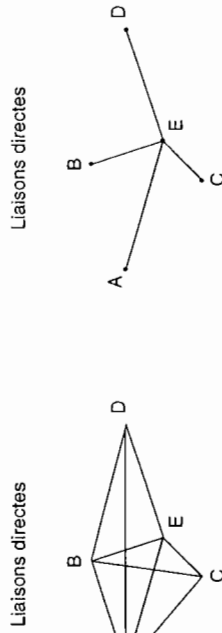
• Liaisons directes ou étoilées

La structure réticulaire de l'activité de transport, combiné avec les économies d'envergure dans les coûts de production des différentes relations, et les économies d'échelle par rapport à l'intensité du trafic sur chaque relation, entraîne des particularités de l'offre.

Dans le cas des exploitations en réseaux, les économies d'envergure se traduisent par le fait que l'exploitation de relations différentes est réalisée de façon plus économique par une seule entreprise que par plusieurs. La mise en commun de plusieurs lignes permet, avec un parc de matériel plus grand, une meilleure gestion des incidents, une certaine compensation des pointes. De même les économies d'échelle impliquent que sur une relation donnée le coût de transport ne croît pas aussi vite que le trafic ; ces dernières se rencontrent dans tous les modes ; la tonne x km offerte coûte moins cher sur un PL de 40 tonnes que sur un PL de 20 tonnes, le siège x kilomètre offert (*skol*) est moins coûteux sur un Boeing de 500 places que sur un Fokker de 40 places, etc.

Une première conséquence de ces traits porte sur la structure des réseaux, et notamment sur le partage entre liaisons directes et réseau étoilé (cf. figure 8-2).

FIGURE 8-2
Liaisons directes et étoilées



En l'absence d'économies d'échelles et d'envergures, les liaisons directes sont le système de coût minimum, puisque c'est lui qui minimise le nombre total de passagers x kilomètres. Il en va différemment avec les économies d'échelle et d'envergure : la liaison entre A et B peut être assurée à un coût plus faible en passant par E si les économies d'échelle sont suffisamment élevées et induisent suffisamment les coûts unitaires sur AE et EB qui concentrent un trafic supérieur à celui de AB seul. La structure la plus avantageuse est alors celle d'un réseau étoilé (*hub and spokes*) On peut concrétiser cet arbitrage sur des réseaux simples.

Toutefois lorsque le passage d'un réseau point à point à un réseau étoilé est avantageux en termes de coût pour l'opérateur à trafic donné, il peut devenir désavantageux si l'on tient compte de la réponse de la clientèle qui sera en général désavantagée à la suite des pertes de temps et du coût de changement. De la même manière le bilan collectif peut être négatif.

ENCADRÉ 8-3

Économies d'échelles et réseau étoilé d'après Encaoua et Perrot (1992)

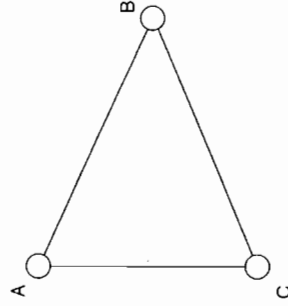
Soient 3 villes équidistantes entre lesquelles s'échangent des trafics égaux q ; ces trafics sont assurés par un transporteur qui, pour chaque couple O, D a une fonction de coût dépendant du trafic x de la forme :

$$C(x)$$

La fonction de coût étant croissante et concave (existence de rendements d'échelle). À quelle condition la liaison AB est-elle assurée plus économiquement en passant par C que directement ? Pour le voir, il faut comparer :

$3C(q)$, coût de transport correspondant au réseau en liaisons directes à :

$$2C(2q), \text{ coût de transport correspondant au réseau étoilé}$$



La condition d'apparition d'un réseau étoilé en termes de coût pour l'entreprise est donc :

$$3C(q) > 2C(2q)$$

Le réseau étoilé sera rentable si les rendements d'échelle sont suffisamment élevés : notons toutefois que cette analyse est valable pour l'entreprise, mais que le coût pour l'utilisateur doit comprendre aussi le coût du changement, ainsi que le coût du temps supplémentaire pour les usagers dont le trajet est allongé.

• L'interdépendance entre les marchés

Une conséquence des réseaux étoilés et de la non constance des rendements est l'interdépendance entre les marchés, un phénomène qu'on peut exposer sur un modèle simple : soient 3 villes ABC desservies par un réseau en étoile ne comportant que AB et BC, comme indiqué dans l'encadré. Le trafic AC effectue un changement en B. Appelons q_{ij} le trafic entre deux villes ij ; le trafic sur l'arc AB est :

$$Q_{AB} = q_{AB} + q_{AC}$$

FIGURE 8-3

Interdépendance des marchés



le coût pour l'opérateur sera :

$$C(Q_{AB}, Q_{BC})$$

et même chose pour l'arc BC. Les marchés seront interdépendants même si ce coût est séparable entre les deux liaisons, c'est-à-dire même si :

$$C(Q_{AB}, Q_{BC}) = C_1(Q_{AB}) + C_2(Q_{BC})$$

En effet une augmentation de la demande q_{AB} va entraîner une diminution du coût unitaire sur AB, donc (dans la plupart des situations de concurrence) une diminution du prix sur AB, donc une augmentation de la demande et en particulier de q_{AC} et un effet sur le marché de la liaison BC, qui sera probablement une réduction du prix et une augmentation du prix et une augmentation des demandes q_{BC} et q_{AC} : à travers la réduction du coût C_2 .

On peut étudier les conséquences de telles situations d'un point de vue soit théorique soit appliqué. Sur le plan théorique, Encoua et Perrot (1992) ont analysé les conditions d'équilibre d'un marché constitué par un réseau d'arcs sur lesquels opèrent des compagnies munies de fonctions de coût présentant les caractéristiques d'économies d'échelle et d'envergure ; la demande sur chaque $O . D$ est fixée. Les usagers choisissent la compagnie leur fournissant le plus faible coût généralisé de transport, qui comprend le prix du billet, le temps de trajet et les temps d'attente. Les opérateurs, en nombre fixé à l'avance, se livrent une concurrence en deux temps : d'abord définition du réseau sur lequel chacun s'installe, ensuite compétition par les prix sur ce réseau. L'optimum collectif est une configuration d'équilibre, mais il peut exister un ou plusieurs autres équilibres de Nash du jeu en cause, qui ne seront pas optimaux. Avec des hypothèses un peu différentes (en particulier en introduisant la congestion aéroportuaire), Hong et Harker (1992) aboutissent à l'existence d'un équilibre unique et explorent les conséquences de différentes modalités de répartition des slots sur l'optimum collectif.

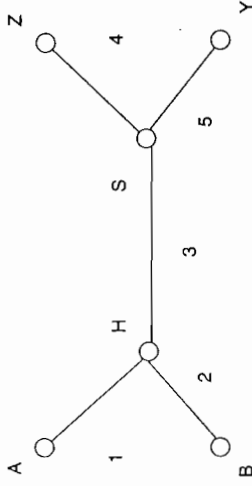
D'un point de vue plus appliqué, Nero (1996) modélise les effets de la dérégulation européenne à travers un modèle de duopole représenté par la figure jointe, et schématisant les liaisons entre deux pays : HA et HB représentent les liaisons intérieures au pays 1, SZ et SY celles du pays 2. HS représente la liaison internationale. Le modèle est complété par la donnée des fonctions de demande :

$$P(Q_{ij}) = a - Q_{ij}/2$$

où i et j représente les $O . D$ possibles pour chaque compagnie (AH, AS, AB, BH, BS, HS pour la compagnie 1), et où la fonction de coût est :

$$C = \sum_i C_i(Q_i) = \sum_i (Q_i - \theta Q_i^2 / 2)$$

où Q_i est, pour chaque compagnie, le trafic sur chaque arc (par exemple 1, 2 et 3 pour la compagnie 1).



Plusieurs situations sont analysées :

- monopole sur les liaisons intérieures et collusion sur la liaison internationale HS : c'est le schéma de la situation européenne avant déréglementation.
- monopole sur les liaisons intérieures et duopole de Cournot sur HS¹. C'est le schéma représentatif de la situation européenne après déréglementation.
- fusion des deux compagnies.

Les résultats, qui dépendent bien sur des paramètres a et q , sont parfois contre intuitifs, en raison des interdépendances des coûts sur les différents marchés origine-destination. Pour des valeurs raisonnables des paramètres (notamment existence d'économies d'échelle pas trop fortes), on aboutit à ce que :

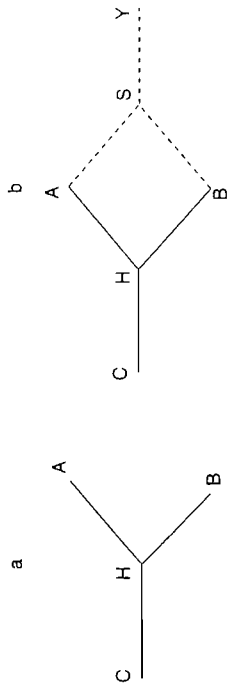
- la collusion entraîne des prix sur HS supérieurs aux prix intérieurs, et des trafics intérieurs (ce résultat provient de l'existence d'économies de densité qui ne sont pas totalement exploitées sur HS, car le trafic est partagé entre deux compagnies, contrairement à ce qui se passe sur les liaisons intérieures).
- la compétition sur HS produit une augmentation des trafics et une réduction des prix non seulement sur HS mais aussi sur les liaisons intérieures. Car la réduction des coûts sur HS entraîne sur cette liaison une réduction des prix, donc une réduction des prix sur AS par exemple, donc une augmentation des trafics intérieurs donc une réduction des coûts intérieurs. Elle s'accompagne d'une augmentation du surplus des consommateurs et du surplus collectif.
- quant à la fusion, elle est meilleure que la collusion pour toutes les lignes, car elle permet de pleinement bénéficier des économies d'échelle sur HS, lesquelles économies d'échelle ont des effets positifs sur les trafics intérieurs. Elle est également meilleure que la compétition pour les lignes intérieures, le résultat dépendant des paramètres en ce qui concerne la ligne internationale HS : dans ce dernier cas, la fusion est meilleure si les rendements d'échelle, symbolisés par θ , sont suffisamment élevés.

1. Même si la concurrence se fait en dernier ressort par les prix, la limitation des capacités entraîne un adoucissement de la concurrence et un rapprochement de la situation de Cournot (théorème de Kreps-Schenkman), orientation renforcée par le fait que les biens considérés ne sont pas substitués parfaits.

Deux conclusions de politique peuvent en être tirées : la compétition est meilleure que la collusion, même pour les lignes intérieures, et la fusion n'est pas trop à craindre si les économies d'échelle sont suffisamment élevées.

D'autres modèles simples peuvent être analysés pour explorer les situations de concurrence. Considérons par exemple la situation inspirée de Bruchner et Spiller (1989) et reprise par Encaoua et Perrot (1992) : dans un premier temps, le réseau est constitué comme sur la figure 8-4a.

FIGURE 8-4



Puis, la compatibilité avec une autre compagnie est créée, et le schéma devient celui indiqué sur la figure 8-4b. Quels effets sur le marché ? Les indications qualitatives qui suivent, et qui se réfèrent au duopole de Cournot, permettent de faire apparaître les mécanismes en jeu : l'arrivée de la compagnie n° 2 aura pour effet de réduire le trafic entre A et B. De ce fait, les coûts unitaires moyens sur les maillons HA et HB vont croître, et le coût des liaisons HA et HB, CA et CB va augmenter d'autant. Les prix sur ces liaisons vont augmenter et le trafic se réduire. Ces conséquences peuvent être contrecarrées par la baisse des prix consécutive à la concurrence, en fonction de l'élasticité du trafic aux prix. Ce qui apparaît clairement, c'est un effet de réseau, à travers lequel un changement des conditions de concurrence sur un arc peut avoir des répercussions sur les autres arcs.

Les complémentarités entre lignes d'un réseau peuvent également être analysées sur un modèle simple représentatif d'une liaison d'apport : une compagnie exploite une liaison AB de fort trafic, où elle se trouve en compétition avec d'autres compagnies ; le prix du marché y est P, et son coût marginal, supposé constant, est Γ ; le profit par unité de trafic est donc : $P - \Gamma$, strictement positif. Si la compagnie exploite une ligne CA, feeder de la précédente où le coût unitaire est γ , elle peut admettre des pertes sur cette ligne, et se contenter d'un prix p inférieur à γ et dépendant de la part π du trafic du feeder dans le trafic total sur CA. La condition est :

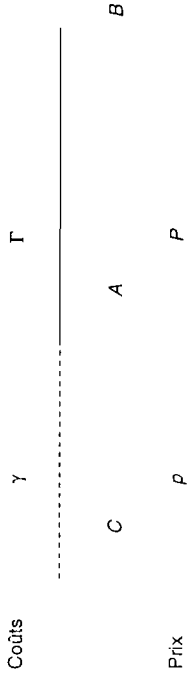
$$(1 - \pi)(p - \gamma) + \pi(p - \gamma + P - \Gamma) > 0$$

ou :

$$p > \gamma - \pi(P - \Gamma)$$

Cet effet dépend de la répartition des propriétés. Il ne se produirait pas si les deux liaisons étaient exploitées par des compagnies différentes qui ne bénéficieraient pas des externalités positives qu'elles se créent l'une l'autre.

FIGURE 8-5
La problématique des liaisons d'apport



D'une manière générale la combinaison des substitutions et complémentarités entre maillons et des différences d'élasticités entre $O \cdot D$ entraîne que les tarifs ne sont pas additifs.

La complémentarité et la substitution peuvent aussi s'exercer sur le choix des aéroports-hub par les compagnies. Pels, Nijkamp et Rietveld (1997) analysent les politiques de choix du hub par les compagnies en fonction des systèmes de tarification (coût marginal ou coût complet) des aéroports. Ce système de tarification influe ou non sur le choix du hub selon divers paramètres, en particulier le volume de la demande et l'importance des économies d'échelle.

1.3.2.e) La discrimination des produits et des services

Une forme de concurrence très fréquente dans les transports est la discrimination des produits. Cette discrimination des produits s'analyse par des modèles formellement analogues à ceux présentés au chapitre II pour étudier la concurrence spatiale. On y interprète la distance x comme un paramètre de différenciation des utilités des agents, les produits présentés par les deux entreprises en concurrence correspondent à des valeurs particulières différentes de ce paramètre ; et pour chaque consommateur et chaque produit, on peut définir un prix généralisé p' défini par :

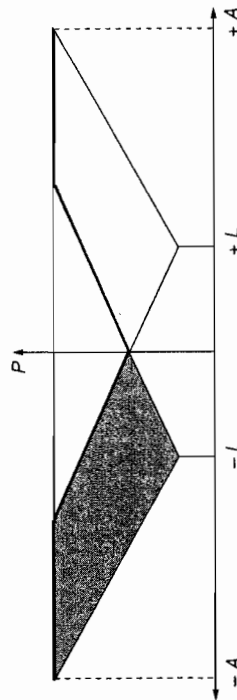
$$p' = p + f(x - x_0)$$

où p est le prix du produit, x_0 la valeur du paramètre pour le produit, x la valeur du paramètre pour le consommateur ; la stratégie des entreprises consiste à choisir la valeur du paramètre x_0 qui leur sera le plus favorable. Si on se place dans le cadre d'une concurrence par les prix, et si les utilités des usagers varient comme le carré de l'écart entre leur paramètre et celui qui caractérise le produit, on retrouve le résultat que les entreprises ont intérêt à se différencier au maximum pour exploiter au mieux l'avantage monopolistique que leur procure la différenciation ; si en revanche les utilités varient comme la valeur absolue de l'écart, alors les entreprises ont intérêt à rapprocher davantage leur produit pour ne pas abandonner le marché à leur concurrent.

C'est une forme de discrimination des produits que constitue la concurrence sur une même liaison de transports de performance, et plus particulièrement de vitesses, différentes.

Le schéma correspondant peut aussi être appliqué à un cas particulier intéressant, celui de la concurrence entre compagnies aériennes pour le transport de fret (Bensimon, 1993). La situation est représentée sur la figure 8-6¹. Les compagnies offrent à leur client une prestation de bout en bout qui comprend l'enlèvement du colis et son acheminement terrestre par camion jusqu'à l'aéroport, puis son transport aérien, et enfin son acheminement terrestre à destination. Chaque compagnie offre un prix auquel correspond un coût de transport terrestre plus un coût de transports aérien. Supposons que deux compagnies basées chacune sur des aéroports distants de $2L$ soient face à des clients uniformément répartis sur leur axe, et ayant chacun à faire un envoi avec un prix de réservation de P , que le coût de transport aérien soit le même pour les deux, et que le coût de transport terrestre soit proportionnel à la distance. Il est facile de voir qu'en chaque point, les coûts des deux aéroports sont représentés par les deux entonnoirs en traits fins, que le prix est représenté par la partie supérieure de ces deux entonnoirs (traits épais), et que le profit de chaque compagnie est représenté par les aires grisées. Les entreprises se partagent le marché, et le prix réalisé dépend à la fois de la concurrence et des coûts. Il est minimum au milieu des deux localisations, là où la concurrence est la plus vive, mais où le coût n'est pas minimum.

FIGURE 8-6
Compétition entre aéroports



Un aspect de la différenciation des produits est celui de la qualité du service offert. Lorsque l'opérateur joue sur un paramètre de qualité, lui donne-t-il le niveau optimal d'un point de vue collectif ? On peut examiner cette question en utilisant un modèle simple adapté de Tirole (1988) ; considérons un monopole dont le coût de production dépend à la fois des quantités q et de la qualité fournie θ et dont la demande est de la forme :

$$p = p(q, \theta)$$

le profit du monopole est :

$$\Pi = qp - C(q, \theta)$$

La qualité est optimale pour l'entreprise si :

$$\frac{\partial \Pi}{\partial \theta} = q \frac{\partial p}{\partial \theta} - \frac{\partial C}{\partial \theta} = 0$$

Quant au surplus collectif, il vaut :

$$SC = \int_0^q p(q, \theta) dq - C(q, \theta)$$

La qualité est optimale pour la collectivité si :

$$\frac{\partial C}{\partial \theta} = \int_0^q \frac{\partial p}{\partial \theta} dq$$

Ces conditions sont équivalentes si $\partial p / \partial \theta$ est une constante. Cette dernière condition implique :

$$p = f(q) - a\theta$$

ou encore :

$$q = f(p + a\theta)$$

Or, on a vu au chapitre IV que la plupart des modèles de trafic satisfaisaient à cette condition pour ce qui est de principaux paramètres : temps, sécurité, confort. Ces résultats tendent à considérer que les opérateurs ont eux-mêmes intérêt à fournir la bonne qualité de service ; ceci sous la réserve que ces paramètres n'engendrent pas d'effets externes et que l'information des utilisateurs à leur sujet soit bonne.

Une dimension particulière de la compétition dans les transports est celle de la concurrence sur les horaires. On a vu plus haut les circonstances de cette concurrence sur les horaires pour des entreprises en situation de réseau. L'effet de cette compétition lorsqu'elle porte sur les horaires de services réguliers de transports collectifs a été analysée par plusieurs auteurs. Quinet (1991) se place dans l'hypothèse où les heures de départ souhaitées par les usagers sont uniformément répartis, et où le prix généré par le transport varie linéairement en fonction du temps d'attente ; on voit que la compétition sur les horaires est instable ; cela correspond aux constatations faites dans les situations où on a pu constater cette sorte de concurrence (courses entre bus de compagnies différentes). Il apparaît aussi que le monopole, l'une des formes qui peut résulter, offre un nombre de services inférieur d'environ deux fois à l'optimum collectif, et cela va bien dans le sens de ce que l'on constate lorsque l'on compare des situations de gestion monopolistique privée avec des situations de gestion publique menées avec le souci du service public.

1.3.2.f) La discrimination des prix

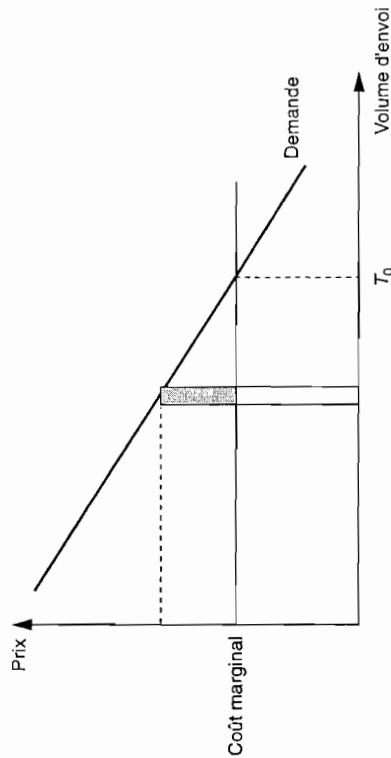
Les transports sont aussi le siège des discriminations par les prix usuelles, que l'on classe traditionnellement en trois catégories en se plaçant dans la situation d'un monopole.

1. Analogie à la figure 2-1 du chapitre II, et soutenant un raisonnement similaire.

• **Discrimination de 1^{re} espèce**

La discrimination de 1^{re} espèce est celle où l'opérateur peut connaître les consentements à payer de chacun de ses clients et peut demander à chacun un prix différent. La situation est approximativement réalisée en matière ferroviaire pour les envois par trams complet. L'opérateur se trouve en face d'un petit nombre de clients, dont il finit par bien connaître les besoins et ce qu'ils sont prêts à payer pour l'envoi (c'est approximativement ce que coûterait le transport par le mode concurrent, en général la route). Il peut donc, dans le contrat qu'il établit, fixer son prix de façon à s'approprier la totalité du surplus du consommateur. À noter que dans ces conditions le volume de trafic T_0 est optimal, la différence avec l'optimum concurrentiel réside dans l'appropriation du surplus.

FIGURE 8-7
Discrimination de 1^{re} espèce



• **Discrimination de 2^e espèce**

La discrimination de 2^e espèce est celle qui peut être opérée lorsque l'entrepreneur, qui vend son produit à plusieurs catégories de clients, est capable de différencier selon un critère observable, par exemple les usagers dépassant un âge donné et les autres, ou les usagers de semaine et ceux du week-end. Alors il peut fixer un prix différent pour chacune des catégories. Analysons cela sur un modèle simple. Soit un monopole fabricant son bien selon une fonction de coût $C(q)$. Il s'adresse à plusieurs clientèles définies par leur fonction inverses de demande, supposées indépendantes :

$$P_i = P_i(q_i)$$

Le profit de l'entrepreneur est :

$$\Pi = \sum_i q_i P_i(q_i) - C\left(\sum_i q_i\right)$$

il est maximum pour :

$$q_i \frac{dp_i}{dq_i} + p_i(q_i) - \frac{dC}{dq_i} = 0$$

On en tire facilement que le prix doit vérifier :

$$\frac{p_i - \frac{dC}{dq_i}}{p_i} = - \frac{1}{e_i}$$

e étant l'élasticité de la demande.

Cette tarification de Ramsey-Boiteux conduit à faire payer davantage aux usagers dont l'élasticité est faible, et moins à ceux dont l'élasticité est élevée¹. Elle explique les pratiques de tarification différenciée très utilisées dans les transports : tarifs vermeil, cartes jeunes, tarifs week-ends... Cette discrimination aboutit-elle à une augmentation du bien être ? Il n'y a pas de résultat général, mais on démontre que si la discrimination conduit à réduire les quantités, alors le surplus collectif diminue, la réciproque n'étant pas vraie (Mougeot et Naegelien, 1994).

• **Discrimination de 3^e espèce**

Le troisième type de discrimination se produit lorsque la population comporte deux ou plus généralement plusieurs catégories d'usagers, mais qu'il n'est pas possible de les distinguer par un critère objectif. Par exemple les consommateurs sont caractérisés par un paramètre θ variant entre $\underline{\theta}$ et $\bar{\theta}$ et obéissent à la loi de répartition :

$$F(\theta) = \Pr(\theta \leq \theta_0)$$

leur utilité indirecte pour le bien considéré étant :

$$S(\vartheta, q) = \theta u(q) - T(q)$$

Enfin $u(q)$ étant telle que $u' > 0$ et $u'' < 0$, et $T(q)$ étant le tarif proposé par le monopole pour l'achat de q quantités. Le problème du monopole est de fabriquer le tarif $T(q)$ qui lui rapporte le profit maximal.

ENCADRÉ 8-4

Discrimination de 3^e espèce et menus de tarifs

Les consommateurs sont caractérisés par un paramètre θ variant entre $\underline{\theta}$ et $\bar{\theta}$

$$F(\theta) = \Pr(\theta \leq \theta_0)$$

leur utilité indirecte pour le bien considéré étant :

$$S(\vartheta, q) = \theta u(q) - T(q)$$

1. En cas d'interdépendance des demandes, la règle s'exprime de façon moins simple (cf. l'encadré 7-5 du chapitre 7).

Enfin $u(q)$ étant telle que $u' > 0$ et $u'' < 0$ et $T(q)$ étant le tarif proposé par le monopole pour l'achat de q quantités. Le problème du monopole est de fabriquer le tarif $T(q)$ qui lui rapporte le profit maximal.

L'effet de ce tarif est que la quantité q sera achetée par l'utilisateur caractéristique θ . Soit $q = q(\theta)$ cette fonction. Elle vérifie :

$$S(\theta, q(\theta)) > S(\theta, q(\theta')) \quad \theta \neq \theta'$$

soit

$$S(\theta, q(\theta)) = \partial u(q(\theta)) - T(q(\theta)) > \partial u(q(\theta')) - T(q(\theta')) \quad \forall \theta, \theta'$$

En tenant compte de ce que :

$$T(q(\theta')) = -S(\theta', q(\theta')) + \theta' u(q(\theta'))$$

on en déduit

$$S(\theta, q(\theta)) \geq S(\theta', q(\theta')) + u(q(\theta')) (\theta - \theta')$$

et par intervention :

$$S(\theta', q(\theta')) \geq S(\theta, q(\theta)) + u(q(\theta)) (\theta' - \theta)$$

d'où

$$u(q(\theta)) (\theta - \theta') \geq u(q(\theta')) (\theta - \theta')$$

c'est-à-dire que si :

$$\theta > \theta'$$

alors :

$$u(q(\theta)) > u(q(\theta'))$$

et :

$$q(\theta) > q(\theta')$$

Il faut aussi bien sûr que chaque agent ait intérêt à acheter, c'est-à-dire que :

$$\partial u(q(\theta)) - T(q(\theta)) \geq 0$$

Mais cette dernière contrainte est atteinte pour tout θ dès qu'elle l'est pour l'utilisateur $\underline{\theta}$. En effet :

$$\partial u(q(\theta)) - T(q(\theta)) > \partial u(q(\underline{\theta})) - T(q(\underline{\theta})) \geq \partial u(q(\underline{\theta})) - T(q(\underline{\theta}))$$

Par ailleurs :

$$\frac{dS}{d\theta} = \frac{\partial S}{\partial \theta} + \frac{\partial S}{\partial q} \frac{dq}{d\theta}$$

Mais, comme S est maximum en q ,

$$\frac{dS}{d\theta} = \frac{\partial S}{\partial \theta} = u(q(\theta)) \quad (\text{c'est le théorème de l'enveloppe})$$

et :

$$S(\theta) = \int_{\underline{\theta}}^{\theta} u(q(y)) dy + S_{\underline{\theta}}$$

Donc :

$$T(q(\theta)) = \partial u(q(\theta)) - \int_{\underline{\theta}}^{\theta} u(q(y)) dy - S_{\underline{\theta}}$$

l'espérance de profit de l'entreprise est :

$$\begin{aligned} E(\Pi) &= \int_{\underline{\theta}}^{\bar{\theta}} (T(q(y)) - cq(y)) f(y) dy \\ &= \int_{\underline{\theta}}^{\bar{\theta}} [(y u(q(y)) - cq(y)) f(y) - u(q(y)) (1 - F(y))] dy - S_{\underline{\theta}} \end{aligned}$$

La maximisation de $E(\Pi)$ implique :

$$S_{\underline{\theta}} = 0$$

et :

$$\partial u(q(\theta)) = c + \frac{1 - F(\theta)}{f(\theta)} u'(q(\theta))$$

ou enfin :

$$\frac{u' - c}{u'} = \frac{1 - F}{\theta f(\theta)}$$

On voit, en posant $\theta = \theta + d\theta$ que :

$$\partial u'(q(\theta)) = T'_q(q(\theta)) = p(q)$$

en appelant $p(q)$ le coût de l'unité supplémentaire par rapport à la consommation q .

D'où :

$$\frac{p(q) - c}{p(q)} = \frac{1 - F(\theta)}{\theta f(\theta)}$$

On en déduit plusieurs conclusions :

Si $(1 - F)/\theta f$ est non croissante¹, il en résulte que $p(\theta)$ l'est aussi :

$$dp/d\theta < 0$$

et puisque comme on l'a vu

$$dq/d\theta < 0$$

alors :

$$T''_q = dp/dq = (dp/d\theta) : (dq/d\theta)$$

est négatif.

Résumons les caractéristiques de la solution, dont la démonstration est esquissée dans l'encadré 8-4 :

- la satisfaction de l'utilisateur qui a la plus faible valeur du paramètre (qui ressent le moins d'attrait) est nulle,

1. Cette condition est remplie pour la plupart des distributions de probabilité.

- la quantité consommée est d'autant plus grande que θ est élevé (que le consommateur est attiré),
- le prix payé augmente avec la quantité mais moins vite,
- le prix marginal est supérieur au coût marginal et décroît avec θ (avec l'attrait du consommateur pour le bien) ; il devient égale au coût marginal pour $\bar{\theta}$: seul l'usager correspondant (celui qui exprime le plus d'attrait) est servi de façon optimale.

Ce schéma de discrimination explique l'existence de tickets en carnet moins cher qu'à l'unité, de cartes forfaitaires à la durée dans les transports collectifs urbains qui reviennent moins cher que les tickets à l'unité pour les gros utilisateurs, le fait que les cartes au mois reviennent moins cher qu'à l'unité. Il explique aussi la pratique des abonnements, permettant de bénéficier d'un prix réduit moyennant le paiement initial d'un forfait. Ce dernier mécanisme ressortit aussi, plus simplement, au tarif binôme dont le principe est présenté dans l'encadré 8-5.

ENCADRÉ 8-5 Le tarif binôme

Supposons que la clientèle d'un monopole, dont le coût marginal de production est nul, soit composée de deux catégories d'usagers, tous identiques à l'intérieur d'une catégorie, mais indiscernables par le monopole. Dans la première catégorie, la courbe de demande inverse est :

$$p(q_1) = (a - q_1) \quad \text{d'où : } q_1 = a - p$$

Dans la deuxième catégorie elle est :

$$p(q_2) = k(a - q_2) \quad \text{d'où : } q_2 = ka - kp$$

Avec ces caractéristiques, la deuxième catégorie est plus « demanderesse » que la première. Si le monopole fixe un prix unique, son bénéfice sera, aux dépenses fixes près :

$$B = p(a - p) (1 + k)/2$$

et vaudra :

$$B = a^2(1 + k)/8$$

le prix étant :

$$p = a/2$$

Le monopole peut augmenter son bénéfice avec un tarif binôme de la forme :

$$p(q) = A + bq$$

Si b est fixé, l'abonnement maximum permettant de ne pas faire sortir du marché la clientèle à faible demande est :

$$A = (a-b)^2/2$$

Alors tout le surplus de cette clientèle est accaparé par le monopole. Le bénéfice devient :

$$B = (a-b)^2/2 + b(a-b) (1 + k)/2$$

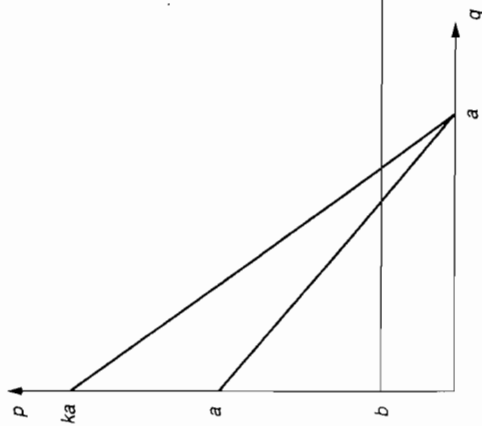
Il est maximum pour :

$$b = a(k - 1)/(2k)$$

et vaut alors :

$$B = a^2(k + 1)/4$$

On verrait facilement que le surplus collectif est augmenté.



Si le paramètre θ est interprété comme un paramètre de qualité, le mécanisme décrit explique par exemple l'existence de plusieurs classes dans les transports collectifs, avec ce paradoxe que les usagers des classes économiques « en ont juste pour leur argent » alors que ceux des classes de luxe bénéficient d'un surplus.

- Le *yield management*

Le *yield management* combine la différenciation des prix et la gestion d'une demande aléatoire pour optimiser les recettes dans les systèmes de vente par réservation tels qu'ils se présentent dans le transport aérien et dans certains transports ferroviaires.

De la différenciation des prix il retient les discriminations de 2^e espèce (tarifs réduits pour les personnes âgées, suppléments pour période de pointe pour certaines catégories d'usagers). Il définit non seulement des prix différents mais aussi des quotas par catégorie. Ces quotas, fixés initialement d'après l'expérience, varient au cours de la période de réservation, pour tenir compte de la manière dont la demande aléatoire se développe au cours de cette période. Ainsi, si elle est plus forte que prévu, on réduira les quotas de tarifs réduits.

Le système comporte aussi des réductions pour achats à l'avance fermes et non remboursables. Les justifications de ces réductions, sont les suivantes (Caillaud et Quinet, 1992) :

- un allègement de la trésorerie de la compagnie,
- une assurance contre les aléas de la demande liées à l'aversion au risque de la compagnie.

Ces réductions pour achat à l'avance sont aussi soumises à quotas ; ces quotas varient aussi au cours de la période de réservation en fonction de l'évolution constatée de la demande.

Pour toutes ces catégories, la compagnie cherche aussi à faire varier les prix. Ceci est rarement fait ouvertement en raison des réactions négatives de la clientèle¹, mais par le biais de la création de nouvelles catégories différenciées visuellement par rapport aux anciennes, dont on ferme le quota. Ainsi dans le cas du transport aérien on voit quelque fois les quotas de tarifs réduits fondre, lorsque la demande à plein tarif est suffisante, ou au contraire, quelques jours avant le départ, apparaitre de nouvelles offres bradées.

Tout ceci est géré par des programmes d'optimisation en avenir aléatoire dont le but est de maximiser la recette et le remplissage en jouant sur le moment où ce remplissage maximal est atteint le plus efficacement (voir Sauvart, 1996).

1.3.2.g) La concurrence destructrice

L'analyse économique permet d'éclairer une question souvent évoquée dans le monde des transports, celui de la concurrence destructrice. On peut l'appréhender cette question à travers la théorie des jeux coopératifs et l'existence de cœur. Une présentation formalisée peut en être trouvée dans Sharkey (1982) et est résumée ci-dessous :

Un jeu coopératif est défini par N joueurs, qui recueillent des utilités u_i . Une coalition S de joueurs peut recueillir un ensemble d'utilités définies comme un vecteur \bar{u} . Le cœur du jeu, s'il existe (les jeux n'ont pas tous un cœur), est l'ensemble des utilités atteignables v_i (i allant de 0 à N) telles que aucune coalition ne puisse procurer à ses participants j des utilités \bar{u} vérifiant : $u_j > v_j$.

Si les utilités sont transférables d'un individu à l'autre, les vecteurs \bar{u} et \bar{v} peuvent être remplacés par des nombres :

$$v = \sum v_i \quad \text{et} \quad u = \sum u_i$$

S'il y a une infinité de joueurs, on peut les représenter par des points d'un segment $(0, \tau)$. Une coalition sera un ensemble de points, et le jeu sera défini par une fonction caractéristique, utilité que la coalition des joueurs compris entre 0 et s peut se garantir, soit : $V(s)$.

On démontre qu'un tel jeu a un cœur si et seulement si :

1. Réactions dont la SNCF a fait l'amer expérience lors de l'introduction du *yield management* à l'occasion de la mise en service du TGV Nord.

$$V(s) \leq V(t) \quad \forall s \leq t$$

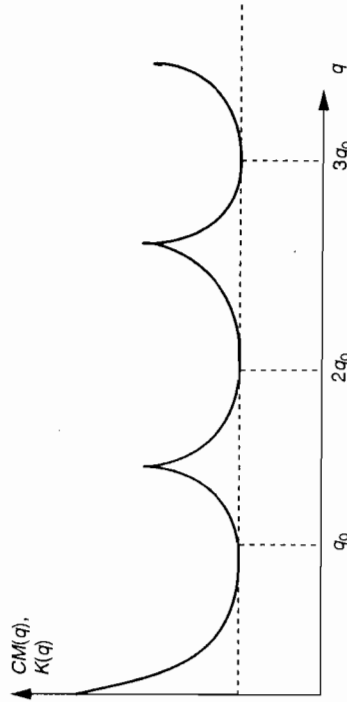
Considérons alors un marché composé d'un nombre fini d'entreprises ayant toutes la même fonction de coût moyen $CM(q)$ ($CM(q)$ a une forme en U) et d'une infinité d'acheteurs ayant chacun la même utilité u pour le bien en cause.

Le nombre d'entreprises assurant une production donnée q à un coût moyen minimum ainsi que le coût moyen minimum $K(q)$ correspondent apparaissent sur la figure jointe ; ce coût $K(q)$ est défini par :

$$K(q) = \min_n (CM(q/n))$$

Si, comme on en fait souvent l'hypothèse, la courbe $CM(q)$ a une forme en U et est plate au voisinage de q_0 , $K(q)$ est une succession de festons de plus en plus plats ; à la limite, lorsque q est très grand, on retrouve la situation de concurrence parfaite où chaque entreprise, très petite par rapport au marché, opère à son coût moyen minimum. L'utilité d'un groupe de consommateurs est alors l'inverse du coût minimum. Dans ces conditions, on peut représenter le marché comme un jeu entre les consommateurs, jeu dont la fonction caractéristique est :

$$V(s) = sK(s)$$



La condition de stabilité précédente n'est remplie que si : $q = nq_0$. Ceci se produit lorsque n est très grand, ou, lorsque n est petit, pour des valeurs ad hoc de la production ($q = nq_0$).

On peut expliquer ainsi ce mécanisme : lorsque la capacité de l'industrie (définie comme le produit associé à un coût moyen minimum) n'est pas égale à la demande à ce prix, l'équilibre du marché nécessiterait que le prix dépasse ce coût minimum ; mais une ou plusieurs firmes auraient intérêt à s'entendre pour produire au coût minimum, en ne servant qu'une partie de marché. L'argument est voisin de celui présenté dans l'encadré 6-10.

Les courbes de coût considérées ici sont les courbes de coût moyen ; elles constituent un guide pour l'entreprise dans une vue d'équilibre de long terme. Le raisonnement fait avec leur intermédiaire est donc valable lorsque les conditions du marché en offre et en demande sont jugées durables.

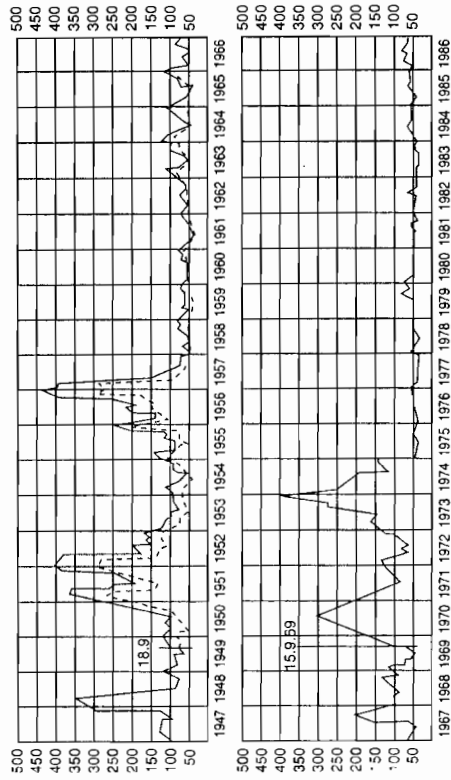
Lorsqu'il n'est pas ainsi, lorsque les conditions du marché sont transitoires, les courbes précédentes doivent être vues comme représentant le coût moyen

variable hors dépenses fixes ; si par exemple le niveau de la demande baisse (c'est à court terme la source de variabilité la plus forte), l'équilibre du marché nécessiterait qu'une ou plusieurs entreprises sortent du marché pour restaurer l'égalité entre le prix et le coût moyen minimum. Mais les entreprises qui auraient à sortir peuvent avoir intérêt à attendre des jours meilleurs et à « casser le prix » par rapport au coût moyen tant que ces prix dépassent le coût moyen variable, pour assurer une couverture, même partielle, des frais fixes.

Ces situations d'instabilité en concurrence par les prix rejoignent les analyses de duopole de Bertrand (Annexe 7) qui montrent de nombreuses situations sans équilibre.

L'idée générale est que lorsque les capacités de production ne sont pas adaptées au niveau de la demande, des instabilités vont apparaître et le marché concurrentiel ne fonctionnera plus. Or, dans les transports les situations sont fréquentes où ces causes d'instabilité peuvent apparaître ; elles résultent de la nature de la demande et de la nature de l'offre¹. La demande est variable ; elle subit des variations saisonnières et des variations conjoncturelles avec l'activité économique. Quant à l'offre, elle est aussi parfois aléatoire. C'est ce qui résulte du caractère non régulier des flux dans les transports de marchandises, entraînant des retours à vide. Dans l'ensemble les excédents de capacité sont la situation la plus courante, encouragée par les chargeurs et par les constructeurs de matériels. En outre, les équipements en capital prennent une part importante dans le coût des opérateurs, leur durée de vie est longue, ce qui fait qu'une capacité excédentaire aura des effets durables. La conséquence la plus nette est l'instabilité des marchés de transports non réguliers, où les opérateurs se livrent en permanence une guerre des prix, et où les prix fluctuent largement. On peut en prendre une vue en considérant les taux de fret du transport maritime de vrac, l'un des rares marchés pour lesquels on dispose de statistiques de prix (voir figure 8-8).

FIGURE 8-8
Volatilité des cours du fret maritime



Une manière de lutter contre cette instabilité serait la collusion entre les entreprises. Cette collusion peut également apparaître en l'absence de raisons d'instabilités du marché, simplement parce que c'est l'intérêt des concurrents dans le cadre d'une confrontation durable (voir Annexe 7 et le « *Folk theorem* »).

Button (1991) a analysé les conditions de ces deux types de collusion et a résumé en un tableau les conditions de collusion pour recherche de rente et pour lutter contre l'instabilité :

Probabilité de collusion

Condition	Recherche de rente	Lutte contre l'instabilité
Offre hétérogène	Faible	Faible
Demande inélastique	Faible	Élevé
peu d'agents	Élevé	Élevé
Récession	Pas clair	Élevé
Offre ou demande variable	Faible	Élevé
Restriction légales	Élevé	Faible

L'analyse du tableau suggère qu'en général, dans le secteur des transports, les risques de collusions pour recherche de rente sont faibles ; ceux de collusion pour lutter contre l'instabilité des marchés semblent un peu plus élevés.

1. Button (1996) a spécialement analysé en effet dans le cas du transport aérien de passagers.

2. Les effets de la concurrence

On vient de voir que, dans les transports, la concurrence peut utiliser des outils très diversifiés, qu'elle est souvent ruineuse et instable. Autant de raisons pour que, plus qu'ailleurs, les entreprises cherchent à s'en protéger, et les pouvoirs publics à la surveiller.

2.1. L'intervention des pouvoirs publics dans les marchés concurrentiels

Dans les transports, les pouvoirs publics ne sont jamais totalement absents des marchés de toute nature. On se concentrera ici sur les marchés concurrentiels, laissant le cas des marchés de monopole pour le chapitre suivant. Mais la distinction est parfois difficile à faire : un même marché peut selon les époques, être monopolistique ou concurrentiel, ou passer insensiblement de l'un à l'autre ; il peut y avoir un monopole sur le bien, mais avec des substituts proches ; enfin les pouvoirs publics peuvent agir sur la structure du marché en poussant au développement de la concurrence ou en légalisant le monopole. Cela explique les renvois au chapitre suivant qu'on rencontrera parfois dans la suite du texte.

L'intervention des pouvoirs publics est plus ou moins forte selon les périodes. Elle était très prégnante dans l'après-guerre. Puis, il y a environ quinze ans a commencé une période de libéralisation qui s'est développée sous des formes et avec un calendrier différents selon les pays. L'analyse des causes de l'interventionnisme antérieur, ainsi que celle des effets de la libéralisation récente, fournissent des renseignements sur le fonctionnement des marchés et l'efficacité de l'intervention publique.

L'intervention publique a d'abord un rôle général de surveillance d'une activité économique : vérifier le respect des lois et règlements, s'assurer du respect de la qualité des produits lorsque celle-ci n'est pas immédiatement perceptible par l'utilisateur.

Elle doit également assurer l'exécution des éventuelles obligations de service public, qui se traduit alors souvent par l'instauration de monopole. Les deux outils en sont la protection du marché par l'attribution de droits exclusifs et la mise en œuvre des subventions croisées. Ainsi, dans son attribution des lignes aériennes, le Civil Aeronautic Board des USA s'arrangeait pour équilibrer, dans la dotation d'une même compagnie, les lignes rentables et les lignes non rentables, que des raisons de service public et d'aménagement local conduisaient à maintenir malgré leur pertes. De même les transports publics urbains et régionaux en France sont exercés sous forme de droit exclusifs, dans l'idée que la concurrence ne permettrait pas d'assurer la régularité et les conditions de tarifs qui exigeait le service public. Ces aspects seront développés au chapitre suivant.

Elle doit aussi remplir une fonction d'efficacité en corrigeant les conséquences négatives des effets externes. On en a vu les modalités essentielles au chapitre précédent : réglementation et taxation.

Elle doit enfin remplir une fonction de régulation en veillant à la stabilité et au fonctionnement harmonieux des marchés. Une bonne partie de la réglementation qui frappait, et qui frappe encore, les transports était inspirée de l'idée que le libre jeu des marchés conduirait à des désordres qu'il convenait d'éviter.

Ainsi, aux USA la réglementation des chemins de fer avait été mise sur pied au 19^e siècle pour éviter que ces derniers n'usent de leur pouvoir de monopole à l'égard de leurs clients, en particulier les agriculteurs. Ultérieurement la réglementation du transport routier a rempli un rôle analogue : il fallait éviter une concurrence désordonnée avec les chemins de fer, qui aurait pu entraîner des faillites et disparitions d'activités. La réglementation était très contraignante ; elle concernait à la fois les tarifs et l'entrée sur le marché : il fallait disposer d'une licence, accordée limitativement, définissant la nature du produit, la quantité autorisée, la liaison concernée.

Le transport aérien aux USA a été également encadré à sa naissance jusqu'à la fin des années 70, pour éviter une concurrence néfaste entre entreprises à rendements croissants, et pour empêcher les monopoles qui en auraient résulté d'abuser de leur pouvoir.

En France, l'intervention de l'État a eu deux motifs principaux : la protection d'activités menacées, et la crainte qu'un marché totalement libre soit l'objet d'instabilités dommageables. Pour protéger le chemin de fer, une réglementation du transport routier de marchandises, comportant licences et tarification à fourchettes, a été établie en 1947 ; elle ne fut guère efficace, n'empêchant pas les pertes de marchés pour la SNCF. De même la batellerie a été protégée par la mise en œuvre de licences, par une tarification obligatoire et par le système du tour de rôle, d'après lequel, dans chaque marché local, les offres de fret étaient, non pas mises aux enchères, mais présentées au premier inscrit sur la liste d'arrivée des péniches.

Ceux qui prônaient la déréglementation dénonçaient les effets négatifs de l'intervention publique : obstacles au progrès technique, à la mise en œuvre des gains de productivité, création de rentes susceptibles d'être dissipées en gaspillage, mauvaise satisfaction des besoins des consommateurs. Ils contestaient aussi l'intérêt de l'attribution de monopoles à travers des arguments qui seront développés au chapitre suivant.

2.2. Les expériences de libéralisation et leurs conséquences

À partir de la fin des années 80, un vent de libéralisme a soufflé aux États-Unis, où il a touché le transport aérien, le transport routier inter-États de marchandises et de voyageurs, et le transport ferroviaire. Dans toutes ces activités le contrôle de l'entrée sur le marché et la réglementation des prix ont été supprimés.

Ce vent de libéralisme a traversé l'Atlantique et a atteint l'Europe. En Angleterre il a pris un tour extrême, puisqu'il a également touché les transports publics urbains et qu'il s'est accompagné d'une privatisation des entreprises publiques. Il a également été fortement marqué dans les pays

nordiques et en Hollande. L'Allemagne a suivi le mouvement avec un certain retard ; la France fut la plus lente. Seul le transport routier de marchandises est complètement libre, et les mesures qui ont été prises dans le transport aérien ou dans le transport routier de voyageurs le furent sous la pression des directives de l'Union Européenne. Car cette dernière a eu un rôle non négligeable en termes de libéralisation des marchés de transports.

ENCADRÉ 8-7

La politique européenne des transports

Les transports sont un des trois secteurs qui, selon le traité de Rome de 1958, devaient faire l'objet d'une politique commune.

Jusqu'en 1968, les mesures prises en faveur de cette politique commune furent très limitées. Elles se cantonnèrent aux transports terrestres, à la suite d'une interprétation très restrictive des termes du traité. Les premières mesures effectives ne furent prises qu'après 1972, à partir de l'adoption par le Conseil des Ministres d'une ligne politique visant à réaliser l'uniformisation et l'harmonisation entre modes et pays avant de libéraliser le marché. Mais ces mesures n'arrivèrent pas à atteindre l'harmonisation de la tarification des infrastructures et des choix d'investissement, premier objectif avoué ; quelques textes significatifs sont pris concernant les conditions de travail et les normes techniques dans le transport routier ainsi que l'harmonisation des comptes des chemins de fer.

C'est à partir de 1985 que va véritablement s'instaurer une politique des transports, et l'impulsion est venue de la Cour de Justice qui a condamné la Commission pour carence, plus précisément pour ne pas assurer la liberté des prestations. En outre, portant sur une affaire de transport aérien¹, la décision impliquait l'inclusion de ce mode dans la politique européenne. L'Acte Unique de 1986 confirme cette orientation en prévoyant l'instauration d'un espace intérieur où serait assuré la libre circulation des personnes et des biens.

La Commission a alors pris une série de mesures mettant en priorité la libéralisation par rapport à l'harmonisation des conditions de concurrence, en démolissant progressivement les systèmes de licences du transport routier de marchandises et d'autorisations bilatérales dans le transport aérien, et en définissant parallèlement les conditions d'accès à la profession. Ainsi, le transport routier de marchandises sera complètement libéralisé en 1998 avec l'autorisation de cabotage ; le transport aérien l'a été en 1997.

Dans le même temps, l'harmonisation accomplissant de grands progrès sur les plans des normes techniques et des conditions de travail ; la Commission prendrait aussi un certain nombre de directives importantes dans les domaines de l'harmonisation fiscale : taxes sur les carburants ; adoption du principe de territorialité pour la tarification des infrastructures, aboutissant à la mise en place d'une vignette pour les poids lourds dans les pays où il n'y a pas de péage autoroutier ; la Commission a aussi pris des directives réglementant

1. L'arrêt « Nouvelles frontières ».

l'attribution des créneaux horaires (une certaine priorité aux entrants dans l'attribution des créneaux inutilisés) et l'usage des Systèmes Informatisés de Réserve dans le transport aérien.

L'action de la Commission fut plus limitée concernant les autres modes, à savoir les voies navigables et les chemins de fer. Dans les voies navigables les décisions ont surtout concerné la limitation des capacités nécessaires pour faire face à la chute de la demande (actions en faveur du déchargement, restrictions à la création de capacités nouvelles).

Dans le chemin de fer (voir chapitre IX) l'action de la Commission pour libéraliser le marché a démarré en 1991 avec une directive prévoyant la séparation, au moins comptable, de la gestion de l'infrastructure et de celle des activités, et la libéralisation de certains marchés : ceux du transport combiné international et ceux du transit international. Parallèlement ont été prises des directives définissant les règles d'accès au marché, celles relatives à l'attribution et à la tarification des sillons horaires.

En dehors de ces actions essentiellement réglementaires visant à libéraliser les marchés, l'Union Européenne a promu une politique d'infrastructure, à travers l'adoption de programmes pour un Réseau Trans Européen comprenant des autoroutes, des lignes ferroviaires et des voies navigables, et à travers des subventions accordées par divers fonds européens.

En matière de transport maritime, l'action de la Commission s'est surtout exercée à travers une lutte contre les pratiques tarifaires déloyales, une réglementation assez libérale des Conférences maritimes, et la libéralisation du cabotage.

Ces expériences de déréglementation ont été marquées par un certain nombre de caractères communs qu'on va exposer rapidement avant de détailler trois cas particuliers : la déréglementation des transports aériens aux États-Unis ; celle du transport routier de marchandises en Europe et plus particulièrement en France ; enfin celle des autocars de Grande-Bretagne.

2.2.1. Les traits communs des déréglementations

La caractéristique la plus frappante est peut-être la transformation de la structure d'exploitation des réseaux : les pouvoirs publics imposaient des réseaux constitués de relations directes, alors que les entreprises, laissées libres du choix de leur structure, sont passées à une exploitation étoilée, les fameux « *hub and spokes* ». Elles pouvaient ainsi bénéficier des économies d'échelle permises par de plus grands flux de trafics.

Une conséquence en fut la baisse des coûts. Cette dernière résulta aussi des effets de la concurrence accrue entre firmes. Les gains de productivité se combinèrent avec des réductions de coûts unitaires, notamment de salaires. Ainsi, Glaister (1997) estime que les réductions de coûts intervenues dans les transports collectifs de voyageurs au Royaume-Uni ont porté sur le poste « salaires » et sont intervenues pour 1/3 par réduction des effectifs et pour 2/3 par compression des salaires unitaires.

Les restructurations ont aussi porté sur les produits offerts, dont la diversité s'accrut : les systèmes de navettes dans les transports réguliers, une augmentation des fréquences, accompagnée d'une adaptation de la taille des véhicules (minibus, avions, moyen porteur) ; des services particuliers adaptés à chaque clientèle. Les tarifs se sont diversifiés ; leur nombre a explosé dans le cas du transport aérien. Cette diversification des tarifs s'est accompagnée de la mise en œuvre du « *yield-management* », qui vise à réviser ces tarifs en permanence en fonction de l'évolution de la demande et des prévisions de son évolution, pour leur faire produire la recette maximale.

De ce fait certaines dessertes ont vu leur trafic accru. D'autres, celles qui n'étaient pas rentables et dont la puissance publique imposait l'exploitation, ont été fermées. Ce fut le cas par exemple de petites lignes à faible trafic, qui auparavant étaient financées par les subventions croisées en provenance des lignes rentables. Ces lignes à faible rentabilité n'ont pu être maintenues lorsque la puissance publique leur attachait une subvention, comme ce fut le cas par exemple en Grande-Bretagne pour certaines lignes d'autobus.

Le processus d'adaptation aux nouvelles conditions de concurrence prit en général plusieurs années, marquées par des faillites, des fusions, des créations d'entreprises dont certaines disparaissaient très vite. Il est toutefois tout à fait frappant de constater que dans tous les domaines les opérateurs ont cherché à se reconstituer un pouvoir de monopole par l'accaparement des ressources rares. Dans la plupart des expériences, le nombre d'opérateurs, qui avait augmenté après la libéralisation, s'est finalement retrouvé souvent plus bas qu'initialement ; corrélativement, les prix qui avaient substantiellement baissé immédiatement après la déréglementation, se sont ensuite stabilisés.

On a pu, dans certains cas, mesurer la variation de surplus liée à ces déréglementations. Winston (1993) s'est attaché à le faire pour les États-Unis.

TABLEAU 8-1
Les gains de la déréglementation aux USA
(en millions de dollar 1990 par an)

Section	Surplus des consommateurs	Surplus des producteurs	Surplus total
Compagnies aériennes	8,8 à 14,8	4,9	13,7 à 19,7
Chemin de fer	7,2 à 9,7	3,2	10,4 à 12,9
Transport routier de marchandises	15,4	- 4,8	10,6

2.2.2. La déréglementation du transport aérien aux USA

Au milieu des années 70, le transport aérien aux USA était étroitement réglementé. En dehors des réglementations technique et sociale, très tatillonnes, la mainmise de l'administration sur l'attribution des lignes, fréquences, et des tarifs était totale. C'est alors que la pratique du *Civil Aeronautic Board* devint beaucoup plus libérale, jusqu'à ce que l'*Airline Deregulation Act* de 1978 vint acter ces pratiques et permettre de les accélérer. Cela fut fait à telle enseigne que le CAB fut dissous. Et maintenant l'entrée et la sortie ainsi que les tarifs sont entièrement libres.

Les conséquences de ces mesures ont été nombreuses. Elles se sont manifestées par une différenciation des tarifs et des services, ainsi que par une transformation des structures de production : réduction du nombre des dessertes directes, généralisation des réseaux de desserte en étoile, avec des « *hubs* » sur lesquels convergent les vols de la compagnie en cause, permettant à la fois d'augmenter la taille des avions et de bénéficier des économies d'échelle qui en découlent, et de mieux faire face aux aléas de l'exploitation. De ce fait les escales faites au cours d'un même voyage ont augmenté, mais pas forcément les changements d'avion, et en tout cas pas les changements de compagnies : 11,2 % des voyages en 1978, 1,2 % en 1990. Par ailleurs les fréquences ont été accrues. Les compagnies ont aussi abandonné des liaisons non-rentables que l'administration forçait à exploiter pour des raisons de service public. N'ont alors subsisté de ces liaisons que celles pour lesquelles une subvention publique a été versée à l'exploitant.

Structuration plus efficace de l'offre, baisse des prix et coûts, meilleure adaptation à la demande : l'ensemble a été hautement bénéfique, comme l'ont prouvé les analyses citées plus haut. Enfin il ne semble pas que la sécurité aérienne ait été dégradée, contrairement aux craintes de certains.

Mais la dynamique ainsi créée s'est arrêtée, en grande partie sous l'action des compagnies et de leurs efforts pour échapper à la concurrence, efforts couronnés de succès puisque, après une période de foisonnement d'entreprises avec entrées, fusions, faillites, on a assisté à une reconcentration et maintenant le marché est plus concentré qu'avant : les 8 plus grandes compagnies détiennent plus de 60 % du marché, contre 56 % antérieurement. En outre, les fusions de compagnies aériennes qui se partageaient un hub n'ont pas été prohibées ; on a d'ailleurs logiquement constaté qu'à l'issue de ces fusions les profits augmentaient ; mais le résultat en termes de surplus du consommateur n'est pas net car les coûts diminuent.

En dehors des actions classiques et peu pernicieuses pour différencier les tarifs et les services, ou pour fidéliser la clientèle, d'autres stratégies ont contribué plus profondément à réduire la concurrence, et notamment celle des Systèmes Informatisés de Réservation (*Computerized Reservation System*). Ces systèmes, très lourds, sont d'élaboration longue et coûteuse. Seules les grandes compagnies peuvent en créer, et celles qui en disposent

peuvent alors concevoir le système en leur faveur, et peuvent également faire payer très cher aux autres la possibilité d'y accéder. Il ne faut pas s'étonner que ces systèmes aient donné lieu à de nombreuses actions en justice pour pratique anticoncurrentielle.

Un autre facteur de restriction de la concurrence résulte directement de l'organisation en « *hubs* ». Sur l'aéroport correspondant à son hub la compagnie se construit un pouvoir de monopole, à partir de la possession des aéroports et des portes d'embarquement, de la saturation de l'espace aérien combinée avec la règle d'attribution des slots à travers les « droits du grand-père » qui empêche la remise périodique sur le marché des créneaux horaires d'atterrissage ou de décollage. Ainsi, se sont constitués des monopoles locaux. Les systèmes de régulation de la rareté par les tarifs n'ont eu que très peu d'application. Seuls quelques aéroports appliquent pour l'attribution des slots un système d'enchères et encore ce système est-il fortement handicapé notamment par la séparation entre le marché des liaisons courtes et celui des liaisons longues (Starkie, 1994).

Une attente des économistes fut déçue, c'est celle qui concerne les effets de la concurrence potentielle. La théorie des marchés contestables suppose que, lorsque l'entrée et la sortie du marché sont sans coût, la concurrence potentielle qui en résulte est aussi efficace que la concurrence réelle ; elle avait constitué un des arguments importants en faveur de la déréglementation ; il n'était pas nécessaire de réglementer le pouvoir de monopole du transport aérien, la concurrence potentielle s'en chargera puisque l'entrée et la sortie sont sans coût. Il s'est avéré qu'il n'en était rien. L'étude des tarifs et des coûts a montré que les profits dépendaient étroitement du nombre de concurrents réels et que la concurrence potentielle n'avait qu'un effet limité, perceptible uniquement si l'entrant potentielle était déjà installé aux deux extrémités de la ligne.

Au total, si la déréglementation du transport aérien a eu dans l'ensemble des effets positifs, elle s'est accompagnée de conséquences non prévues initialement, et le marché, qui était auparavant rigidifié par l'action des pouvoirs publics, l'est maintenant par la capacité que les grands opérateurs ont montré à se constituer des pouvoirs de monopole, qui ne pourraient être rompus que par des politiques actives énergiques de la puissance publique.

Le diagnostic est clair : c'est par défaut d'action que la puissance publique a pêché. Ce à quoi elle doit s'appliquer, c'est à organiser et discipliner la gestion des ressources rares que sont les aéroports lorsque la saturation est atteinte, ou les Systèmes Informatisés de Réserve (SIR), dont les coûts fixes élevés empêchent la multiplication. Il faut pour cela par exemple élaborer des codes de conduite pour l'usage des SIR ; mettre au point des procédures d'attribution des terminaux d'aéroport et des créneaux horaires d'atterrissage-décollage par des systèmes d'enchères avec attribution de franchise pour une durée limitée, et qui n'assurent pas la main mise définitive d'un des opérateurs privés ; également mener une politique judicieuse d'extension des infrastructures.

Certaines de ces actions ressortissent au rôle traditionnel de la puissance publique ; c'est le cas de la réglementation ou de la construction d'infrastructure. D'autres sont plus inhabituelles, au moins pour nous : c'est d'organiser une répartition efficace des ressources rares, par exemple sous forme d'enchères. Mais toutes exigent de cette puissance publique une vigilance attentive et active, susceptible d'interventions anticipatrices éventuelles, et ce, dans un domaine qui n'est pas de son champ usuel : le fonctionnement d'un marché privé. Elles nécessitent aussi de sa part autorité, puisqu'il lui faudrait lutter contre des droits acquis tels que les « droits du grand-père », et exproprier certains acteurs de droits de propriété exclusifs sinon abusifs.

2.2.3 La déréglementation du transport routier de marchandises en Europe et en France

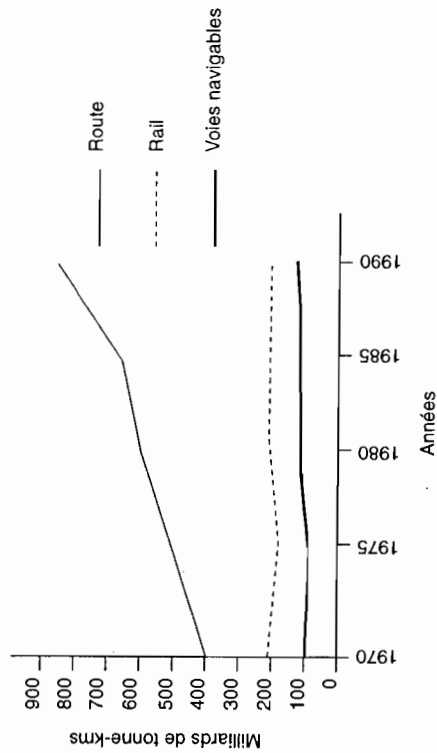
Jusqu'au début des années 80, la politique commune des transports ne s'était penchée que sur des sujets relativement mineurs en ce qui concerne les transports routiers de marchandises. Elle avait établi des directives concernant l'harmonisation des conditions de travail, des caractéristiques techniques des véhicules, et s'était efforcée, sans grand succès d'harmoniser la fiscalité des carburants et les redevances d'infrastructures. Mais l'accès à la profession était strictement réglementé dans la plupart des pays, selon un système complexe de licences, non harmonisé entre les pays. Ainsi, en France la réglementation mise en place en 1949 avait pour but de protéger la SNCF et d'éviter une concurrence néfaste en période de basse conjoncture. L'accès au marché international était réglé par des accords bilatéraux définissant des contingents de licences, et par des maigres contingents multilatéraux de licences européennes. Le cabotage (possibilité pour un transporteur à la fois de prendre du fret et de le livrer dans le pays étranger où il effectue un transport international) était interdit.

Les choses changèrent en 1985, à partir de la décision de la Cour de Justice enjoignant à la Commission de mettre en œuvre des dispositions du Traité pour assurer la libre circulation des biens et personnes et pour réaliser une politique commune des transports. Suivirent alors une série de dispositions qui renforcèrent l'harmonisation des conditions techniques, sociales et fiscales, qui éliminèrent les contrôles frontaliers internes, qui accrurent progressivement les contingents de licences internationales finalement abolies en 1993, et qui créèrent des licences de cabotage, progressivement accrues ; le cabotage lui-même deviendra libre en 1998.

L'Union a également défini les règles d'accès à la profession, valables pour tous les pays, les autorisations d'installation étant automatiquement mutuellement acceptées.

Les trafics routiers ont augmenté. Le graphique joint montre l'évolution des trafics des différents modes pour l'ensemble de l'Europe.

FIGURE 8-9
Évolution des trafics de transport de marchandises



Source : Bayliss et Millington, 1995.

Cette croissance a surtout profité au transport routier pour compte d'autrui, comme le prouvent les chiffres suivants pour la France, tiré de OEST 94 :

En G tonne × km	1985	1991
Trafic compte propre	12,3	13,9
Trafic compte d'autrui		
- routier	43,9	71,1
- SNCF	55,8	51,5
- transport fluvial	7,6	6,8

La libéralisation du transport international intra-européen a considérablement accéléré la libéralisation des marchés intérieurs. Le Royaume-Uni avait complètement libéralisé son système de transport routier de marchandises dans les années 80. C'est l'Allemagne qui a marqué le plus de réticences avec un système de licences. La France n'était pas en retard. La réglementation du transport routier pour compte d'autrui y avait été établie en 1949 et comportait des mesures similaires à celles qui existaient dans les autres pays occidentaux ; elle consistait d'abord en un système de licences : chaque camion roulant en zone longue (c'est-à-dire en gros, faisant un transport de plus de 150 km) devait posséder une licence, et les licences étaient délivrées limitativement par l'État. La « tarification routière obligatoire » (TRO), tarification à fourchette fixée par les pouvoirs publics, jouait surtout comme plancher, pour limiter à la baisse des tarifs.

Au milieu des années 80, la libéralisation a comporté la suppression de la tarification obligatoire et une politique très large de distribution des licences, qui ne constituent plus qu'un moyen très lâche de vérifier la capacité de l'entrant, et dont le prix est tombé à zéro. Cette libéralisation est entrée en synergie avec les transformations structurelles que connaissait notre appareil productif : recentrage sur les activités de base ; recours à la sous-traitance ; recherche de meilleurs fournisseurs, même éloignés ; généralisation de la production en flux tendus et en juste à temps. L'organisation de la production reposait aussi de plus en plus sur des transports bon marché et de bonne qualité. Ces transformations se sont répercutées sur les entreprises de transports. Les exigences de flexibilité, le besoin de développer des structures bénéficiant d'effets de réseaux, ont entraîné une augmentation de leur taille : le nombre total d'entreprise augmentait de plus de 10 % en raison des possibilités nouvelles ouvertes à l'entrée sur le marché ; mais le nombre d'entreprises de plus de 100 salariés est passé de 111 en 1986 à 170 en 1991. Parallèlement, le système de sous-traitance s'est développé dans le secteur : la part du chiffre d'affaire sous traité est passée de 16 % à 19 % au cours de la période.

Tous ces éléments ont eu des conséquences favorables : des coûts plus bas, une meilleure adaptation du système productif, un développement de l'activité et de l'emploi. L'OEST (1994) a évalué ces gains à 9 milliards de F 1991 dont il retire les coûts externes liés à la circulation routière supplémentaire, non compensés par les recettes fiscales supplémentaires soit 3,6 milliards de G. La liberté des prix et la suppression des licences ont ainsi contribué à développer la concurrence et à libérer les initiatives privées. Mais ces effets bénéfiques et faciles à constater se sont accompagnés d'autres conséquences, dommageables elles.

D'abord, la croissance du trafic routier, qui a quasiment doublé en dix ans alors que les trafics ferroviaires et fluviaux régressaient ; c'est le résultat direct d'une compétitivité accrue, et un facteur de congestion et de pollution. Mais aussi les conditions d'exercice de la concurrence qui se manifestent par la pression sur la main-d'œuvre : des durées de travail de 50 ou 60 heures par semaines ne sont pas rares, les salaires du personnel – et les profits des artisans, nombreux dans la profession – sont faibles, voisins du SMIC. Elle se manifeste aussi par la surcharge de nombreux véhicules : le dépassement des poids autorisés est tout bénéfice pour le transporteur, mais il conduit à une dégradation accrue des routes, à des risques d'accident augmentés.

Il en résulte une concurrence déloyale à l'égard de la SNCF et coûteuse pour les budgets publics. Il en résulte aussi périodiquement des crises se manifestant par les grèves des routiers et le blocage des routes. Alors, quelques mesures d'apaisement sont prises, et tout rentre dans l'ordre... jusqu'à la prochaine fois. C'est qu'une solution durable nécessiterait une analyse de la cause profonde de la situation, et une action publique énergique.

La cause profonde c'est que la concurrence en ce secteur y est souvent destructrice, et cela pour deux raisons. La première est la facilité d'entrée sur le marché : le coût d'achat d'un véhicule est faible, l'emprunt est aisément accordé puisque gagé sur un bien réel, les compétences techniques nécessaires modestes ; du coup l'offre est pléthorique. L'autre raison est le déséquilibre permanent entre l'offre et la demande sur un grand nombre de liaisons, déséquilibre dû aux fluctuations de la demande, à la mobilité de l'offre, et aux fréquentes disparités structurelles des flux aller et retour (zones exportatrice ou importatrice net en tonnage). D'où la pression à la baisse sur les prix du fret, les faillites d'entreprises, d'où aussi les tentations à contourner les limites de chargement et les conditions de travail. Pour lutter contre ces excès il faudrait dresser au bon niveau les barrières à l'entrée sur le marché, afin d'éviter l'installation d'entrepreneurs mal préparés sur les plans financier et organisationnel. Ceci a été fait, mais un peu tardivement peut-être. Il faudrait surtout que la puissance publique sanctionne de façon plus efficace les infractions au code de la route et à la réglementation sociale, de façon que la fraude ne soit plus une décision rationnelle de la part des transporteurs.

La libéralisation n'est pas la démission de l'État. La puissance publique doit être forte dans le contrôle des lois et règlements, elle doit intervenir pour corriger les effets externes - ici pour assurer un partage plus favorable au rail, moins polluant que la route - et aussi, c'est plus rare mais cela arrive, pour stabiliser les marchés dans les situations de fluctuations domageables. Là aussi, comme dans l'exemple précédent, on voit la nécessité d'une action forte et énergique des pouvoirs publics.

2.2.4 La déréglementation des transports par autocar en Grande-Bretagne¹

Le transport régulier par autocars en Grande-Bretagne était caractérisé avant 1980 par une apparente ouverture à la concurrence cachant en fait une forte main mise des pouvoirs publics. En théorie il suffisait, pour être opérateur d'autocar, de disposer d'une licence obtenue en démontrant que l'entrée sur le marché allait dans le sens de l'intérêt général. Les arguments négatifs des entreprises en place ou des chemins de fer (« *British Rail* » à l'époque) entraînaient qu'en fait très peu de nouvelles licences étaient accordées chaque année. Le marché était dominé par deux entreprises publiques, la *National Bus Company* qui, par l'intermédiaire d'opérateurs regroupés sous une même marque *National Express*, détenait une très forte majorité du marché et *Scottish Bus Group* qui opérerait en Écosse.

La libéralisation s'est effectuée en deux temps. En 1980, le marché a été déréglementé, c'est-à-dire que toute entreprise satisfaisant à des critères simples pouvaient s'installer comme opérateur et assurer les services

qu'elle voulait à des prix librement fixés. Comme pour le transport aérien aux États-Unis, les autorités publiques locales qui souhaitaient maintenir une ligne rentable passaient un contact avec un opérateur, en lui versant une compensation financière pour l'exécution du service public correspondant. En 1987 le groupe NBC a été privatisé.

On a d'abord assisté à une série de tentatives d'opérateurs nouveaux pour s'installer sur le marché. Ceux-ci ont essayé différentes stratégies. Certains se sont installés sur les lignes les plus fréquentées et les plus rentables en pratiquant une concurrence par les prix. D'autres se sont lancés dans une politique de qualité, proposant des services de haut de gamme.

La réponse de *National Express* a consisté d'abord à restructurer son réseau en éliminant les relations non rentables et en adoptant une structure en étoile, avec plusieurs *hubs* dont un central à Birmingham. L'entreprise a également créé des services haut de gamme et abaissé ses tarifs en manifestant une capacité de réponse rapide aux changements de tarifs, ce qui a réduit les possibilités de gains des concurrents. *National Express* a bénéficié de la taille de son réseau, permettant une publicité plus large, une diffusion commerciale dans les agences de voyages et un réseau de vente plus important. À ces atouts de fait se sont ajoutées plusieurs actions stratégiques, notamment le blocage de l'accès aux gares routières, la mise en œuvre d'un système électronique de réservation sur tout le territoire, ainsi qu'un certain nombre de pratiques visant à se faire donner la priorité de l'information dans les gares routières et dans les agences de voyages.

Enfin *National Express* a résisté aux attaques des entrants et a regagné les parts de marché contestées initialement. Ne subsistent que quelques autres petits opérateurs dans des niches de faible taille. À noter que *Scottish Bus Group* n'a pas connu la même évolution favorable et a finalement été absorbée par cette dernière, qui maintenant privatisée, représente à peu près 80 % du marché total, le reste étant assuré par des entreprises de petite taille sur des lignes particulière isolées ou des relations d'appoint.

La *Monopolies and Merger Commission* dans un rapport de 1994 juge que compte tenu des effets de réseau, aucune entrée importante sur le marché n'est possible. C'est reconnaître que la concurrence ne peut s'exercer qu'à la marge. Elle estime toutefois que cette concurrence à la marge, ainsi que celle provenant des autres modes et notamment du fer, sont des aiguillons suffisants pour *National Express*, rendant inutile une réglementation par les prix.

L'ensemble de l'action de libéralisation a eu des effets bénéfiques pour la clientèle : diversification des services, abaissement des tarifs, amélioration de la qualité de service, accroissement des trafics¹, ceci accompagné d'un abaissement des coûts dans le cadre d'une augmentation de la production du travail, et se traduisant essentiellement par un transfert dans

1. Sources : Porte et Rémy (1995), Glaister (1997), Mackie (1997).

1. Le nombre de déplacements par autocars, qui avait chuté de 23 à 15 millions entre 1975 et 1980, était remonté à 17 millions en 1985.

l'affectation du surplus de l'entreprise, qui profitait auparavant aux employés, et va maintenant aux usagers. Mais ces effets se sont surtout fait sentir au début, durant la phase de concurrence, et les progrès se sont ensuite ralentis.

Ces enseignements de l'expérience anglaise peuvent être complétés par les résultats d'études portant sur l'activités des autocars en Europe. Banister, Berechman et de Rus (1992) ont comparé la situation en Grande-Bretagne, marquée par la réglementation et la privatisation, avec celles, plus variées, qu'on rencontre dans le reste de l'Europe ; ils mettent l'accent sur les manœuvres des entreprises en place pour dissuader les concurrents d'autres sur le marché, en agissant sur la qualité des produits (accès aux terminaux, information) ; ils rappellent aussi les actions prédatrices susceptibles d'être menées une fois que l'entrée a eu lieu et tiennent la position que ces actions ont été limitées en Grande-Bretagne, comme le prouve le fait que les tarifs ont augmenté après la déréglementation ; ils en voient la raison dans les fortes économies de réseau qui marquent l'activité. Ils notent enfin les effets négatifs sur la clientèle de l'instabilité des services et de la mauvaise information des usagers qui en est résulté.

Cet exemple illustre un certain nombre de points qu'on retrouve constamment dans les expériences de libéralisation : les vertus de la concurrence même si celle-ci s'exerce de façon atténuée (ici, dans la situation finale, la concurrence se fait à la marge ou provient des autres modes) ; l'intérêt d'y recourir même si ce faisant on perd quelques économies d'échelle par la duplication des coûts fixes ; la reconcentration des entreprises après une période de perturbations, et les nécessaires vigilance et continuité dans l'action de pouvoirs publics.

3. Conclusion

Les exemples qui viennent d'être présentés permettent de comparer les effets attendus de la déréglementation aux effets réellement survenus et de prendre une vue des conditions de la concurrence dans les transports.

Par rapport aux attentes mises en elles, la déréglementation a effectivement été bénéfique pour les usagers. Les prix ont été réduits, au moins dans un premier temps, les services diversifiés et mieux adaptés à la clientèle, les clients mieux satisfaits se sont accrus, et les effets parfois craints, par exemple sur le plan de la sécurité, ne se sont pas produits. Quant à l'emploi, son évolution a été favorable là où l'augmentation de la clientèle a permis de compenser les réductions de personnel liées à l'abaissement des coûts. La déréglementation a aussi permis une meilleure adaptation des moyens techniques (réseaux étoilés, taille des véhicules). Elle a conduit à une réduction sensible des coûts ; mais cette réduction est pour une part due à des réductions de salaires, et pour le reste à des réductions d'effectifs. En ce sens la déréglementation a conduit à un transfert de surplus des salariés vers les usagers.

Mais, sauf dans les transports routiers de marchandises, la concurrence est difficile à maintenir : les cas de collisions semblent fréquents, la compétition entre duopoles n'a pas l'âpreté qu'on pourrait attendre d'une concurrence par les prix ; les cas sont nombreux où cette compétition s'effectue plutôt à travers la qualité de service ou les horaires. La tendance à la concentration est un trait général, preuve de l'existence d'économies d'échelle et d'effets de réseaux. Cette tendance n'est d'ailleurs qu'une des façons par lesquelles les entreprises essaient d'échapper à la concurrence en se reconstituant des situations de monopoles grâce à la mainmise sur les ressources rares : aéroports, gares routières, SIR...

De ce point de vue on peut considérer que le désengagement des pouvoirs publics a atteint ses objectifs en libérant les activités de transport des carcans qui empêchaient de répondre aux besoins des utilisateurs ; mais qu'en laissant s'opérer une appropriation des ressources rares, il a été trop absolu. Et si, dans le transport comme dans beaucoup d'autres secteurs, la libéralisation est une nécessité pour rendre l'activité plus efficace, les voies de sa réalisation sont parsemées d'embûches.

Il faut parfois, paradoxalement, se garder des excès de la compétition comme on l'a vu dans le transport routier de marchandises, et intervenir pour en modérer les effets. La concurrence destructrice n'est pas une curiosité intellectuelle, on la rencontre dans la réalité avec ses conséquences néfastes qui s'exercent d'abord à l'intérieur de l'industrie qui en est l'objet, et s'étendent, lors de crises, à l'ensemble de l'appareil productif. Mais la plupart des autres compartiments du secteur présentent des économies d'échelle et donnent un grand avantage aux entreprises en place. L'état doit alors intervenir de manière active et dynamique, notamment pour assurer la gestion des ressources rares, comme on l'a vu dans le cas du transport aérien aux États-Unis. Il doit lutter pour vaincre les nombreuses résistances au changement qui s'exercent à la fois en son propre sein et dans les organismes qui lui sont proches, notamment les entreprises publiques. Il doit faire preuve de persévérance dans l'action et de sens dans la durée, comme le montre l'histoire de la libéralisation des transports interurbains de voyageurs au Royaume-Uni et de celle du transport aérien aux USA.

La mise en place d'une concurrence efficace est donc tout l'inverse du laisser-faire. L'appareil public doit être fort et puissant, et doté de qualités très différentes de celles que lui confèrent son rôle traditionnel. Il doit en effet non seulement disposer d'un bon système d'information sur les secteurs concernés, mais aussi d'une bonne connaissance du fonctionnement des activités productives et commerciales. Il doit être non seulement juge mais aussi acteur du processus. Tout autant que le fonctionnement des entreprises, la libéralisation remet en question celui des pouvoirs publics.

BIBLIOGRAPHIE

- B. Bayliss et A. Millington (1995), « Deregulation and logistic systems in a single European Market », *Journal of Transport Economics and Policy*, Septembre.
- S. Borenstein (1992), « The evolution of US Airline competition », *Journal of Economic perspectives*, vol. 6, Spring.
- J. Brander et A. Zhang (1993), « Dynamic oligopoly behaviour in the Airline industry », *International Journal of Industrial Organization*, 11.
- J. Bruckner et P. Spiller (1989), « Competition and mergers in Airline networks », *International Journal of Industrial organization*.
- K. Button (1996), « Liberalising European Aviation : is there an empty core problem », *Journal of Transport Economics and Policy*, Septembre.
- N. Curien et G. Dupuy (1997), *Réseaux de communication. Marchés et territoires*, Presses des Ponts et Chaussées, Paris.
- G. Dang NGuyen (1995), *Économie industrielle appliquée*, Vuibert.
- G. Dupuy (1985), *Systèmes, Réseaux, Territoires*, Presses des Ponts et Chaussées, Paris.
- D. Encaoua et A. Perrot (1992), *Concurrence et coopération dans le transport aérien en Europe*, Rapport pour la CEE. Publications officielles des Communautés Européennes.
- S. Glaister (1998), « L'expérience de libéralisation des transports de voyageurs en Grande-Bretagne », *Transport*, mars.
- E. Golob (1995), « Impact of deregulation on investment and production strategies in the commercial aircraft industry », *Transportation Research Record*, n° 1480.
- B. Haudeville (1994), « Les stratégies concernant l'entrée dans un oligopole étroit : l'exemple de la construction aéronautique civile », *Revue d'Économie Industrielle*, 3^e trimestre.
- S. Hong and P. Harker (1992), « Air traffic network equilibrium : toward frequency, price and slot priority analysis », *Transportation Research*, part A, n° 2.
- D. Kreps et J. Scheinkman (1983), « Quantity precommitment and Bertrand competition Yield Cournot outcomes », *Bell Journal of Economics*, 14.
- P. Mackie (1997), « Réglementation ou concurrence », Communication au 14^e Symposium de la CEMT, Innsbruck.
- M. Mougeot et F. Naegelen (1994), *La discrimination par les prix*, Economica.
- G. Nero (1996), « A structural model of intra-european union duopoly airline competition », *Journal of Transport Economics and Policy*, May.
- OEST (1994), « La déréglementation du transport routier de marchandises. Essai de bilan », document de travail, Ministère de l'Équipement.
- E. Pels, P. Nijkamp et P. Rietveld (1997), *Substitution and complementarity in aviation : Airports vs Airlines*, ronéoté, Free University, Amsterdam.
- A. Perrot (1993), « Compatibility, networks and competition : a review of recent advances », *Transportation Science*, February.
- P. Ponsolle (1996), « Le financement privé des grands projets d'infrastructure : l'expérience Eurotunnel », *Revue Transport*, sept-oct.

- O. Porte et A. Rémy (1995), « La déréglementation du transport interurbain par autocars en Grande-Bretagne : bilan 1950-1995 », *Synthèses OEST*, décembre.
- E. Quinet (1991), « Organizational structure of public transport and assessment of schedules », *Transport planning and technology*, vol. 16.
- E. Quinet (1997), « Issues in building the European high speed train network », *Journal of Infrastructure Systems*, January.
- Salini (1995), *Économie politique du transport routier de marchandises*, Celse.
- M. Savy (1997), *Le rail et la route : le commerce et la dispute*, Presse de l'ENPC.
- W. Sharkey (1982), *The theory of natural monopoly*, Cambridge University Press.
- D. Sauvant (1996), *Cours d'économie des transports*, Ronéoté Université de Paris II.
- D. Starkie (1994), « Air slot markets », *Journal of Transport Economics and Policy*, Septembre.
- S. Szymanski (1993), « Making hay when the sun shines », working paper, Imperial College Management School.
- S. Szymanski (1996), « Rational pricing strategies in the cross channel tunnel », in Quinet and Vickerman editors, *The econometrics of major transport infrastructures*, Mac Millan.
- J. Tirole (1988), *The theory of industrial organization*, the MIT Press.
- P. Whiter (1995), « Deregulation of local bus services in Great Britain : an introductory review », *Transport reviews*, vol. 15.
- C. Winston (1993), « Economic deregulation : days of reckoning for microeconomists », *Journal of Economic Literature*, September.

Monopole et service public dans les transports

Les monopoles font depuis toujours l'objet de l'attention des économistes et des décideurs politiques. Ce sont eux qui ont été les premiers objets des politiques de libéralisation il y a une vingtaine d'années, et l'expérience accumulée sur les mérites et inconvénients des différentes formes de leur contrôle ne fait qu'enrichir les débats. Ceux-ci prennent un tour particulier dans les transports, où les monopoles naturels sont nombreux. En dehors peut-être du transport routier de marchandises, on y rencontre fréquemment des rendements d'échelle, faibles et dépendant des circonstances locales pour le transport aérien et le transport routier interurbain de voyageurs, plus élevés et plus systématique pour les transports collectifs urbains et pour le transport ferroviaire.

Le monopole naturel s'accompagne souvent d'obligations de service public qui lui donnent ses caractéristiques. On les présentera en partant d'un exposé des nombreuses réformes qui, dans presque tous les pays ont, depuis quelques années, touché les chemins de fer, le secteur le plus exemplaire. Leur analyse permettra de faire apparaître les différents enjeux de la gestion des monopoles : choix entre fragmentation ou intégration, problème de l'accès aux réseaux, succédanés de concurrence, modalités de contrôle de l'État. Les sujétions particulières qu'impose l'exercice d'obligations de service public seront ensuite examinées, à partir des modalités d'organisation auxquelles elles donnent lieu : propriété publique ou privée, concessions, concurrence réglementée.

1. Les réformes des chemins de fer

Il y a deux décennies, la quasi totalité des compagnies de chemins de fer européennes étaient organisées sur une base nationale, sous forme d'entreprises publiques intégrées. Il était ainsi possible de profiter des économies d'échelle généralement reconnues comme fortes dans ce mode de transport, tout en évitant les abus d'un monopole privé et en assurant

commodément les obligations de service public qui lui était demandées dans le plupart des pays.

Toutefois, au fil des années, les conséquences néfastes de ce système se sont développées, et d'abord les incohérences et défauts de gestion résultant des interventions publiques dans la gestion, interventions souvent contradictoires, dont le nombre est accru par la dérive bureaucratique, et dont les thèmes sont subordonnés au calendrier politique plus qu'aux besoins du marché ; ensuite l'insuffisance des contrôles et des incitations entraînant l'apparition de gaspillages et de déficits.

Des efforts ont été faits depuis longtemps pour parer à ces inconvénients. La procédure des contrats de plans a permis des progrès pendant un certain temps, mais n'a pas empêché les dérives financières ; et dans la plupart des pays des réformes plus radicales ont été opérées récemment ou sont en cours. Ces réformes ont eu un certain nombre de traits communs quant à leurs objectifs et leurs moyens : introduire la pression de la concurrence pour obtenir une réduction des coûts et une meilleure adaptation à la demande ; couper ou réduire les liens avec l'État ; diminuer la taille des organismes pour éviter les effets néfastes de la gestion bureaucratique.

1.1. En dehors d'Europe

Au Japon (Mituzani et Nakamura, 1997), la réforme a abouti à scinder les J.N.R. en sept entreprises qui se partagent le territoire nippon, assurent l'exploitation des services voyageurs et la gestion des infrastructures, qu'elles louent à une 8^e entreprise chargée du fret. Les entreprises de voyageurs ne sont donc pas en concurrence, sauf peut-être aux frontières de leurs domaines respectifs, ou sous la forme de concurrence par comparaison (« *yardstick competition* ») de leurs performances respectives. L'effet de cette réforme a été une réduction très sensible des coûts, accompagnée d'une réduction des effectifs de l'ordre de 30 %, d'une reprise de la clientèle, et d'un retour aux profits. Enfin, les entreprises issues de la fragmentation suivent un processus de privatisation. Le ressort de la réforme ne semble pas être la recherche de la concurrence, qui comme on l'a vu reste faible, mais la destruction du pouvoir syndical pour aboutir à transférer le surplus de l'entreprise des salariés vers les usagers, et pour casser des structures jugées trop rigides et incapables d'adaptation.

En Amérique du Sud, l'Argentine, le Brésil, le Chili ont ou sont en train de reformer leur chemin de fer selon des modalités similaires, mais en appuyant davantage leurs réformes sur l'introduction de la concurrence.

Ainsi en Argentine, le réseau, surtout utilisé par les marchandises, était en mauvais état technique, déficitaire, et aux performances médiocres. La réforme a consisté à le découper en quatre entités géographiques attribuées chacune à un concessionnaire par appel d'offre. Pour la durée de 30 ans de son contrat, le concessionnaire gère l'infrastructure, exploite les transports de fret, et loue son infrastructure aux compagnies de voyageurs

en général des opérateurs suburbains de petite taille, qui ont aussi fait l'objet de concessions). La concurrence est une concurrence ex ante, pour le marché, les réponses des concessionnaires étant jugées selon un système de notes prenant en compte les subventions demandées, le business-plan, les tarifs projetés et les services assurés, les améliorations techniques envisagées et notamment le programme d'investissement.

Là aussi les effets ont été significatifs, marqués par une amélioration des conditions techniques et une réduction des coûts accompagnée d'une réduction des effectifs de l'ordre de 4 à 1¹. Le nombre total d'employés est passé de 92 000 à environ 20 000. Les pertes des chemins de fer, qui s'élevaient à 800 millions de dollars en 1989, n'étaient plus que de 200 millions de dollars au milieu des années 90 (Kogan et Thompson, 1994).

ENCADRÉ 9-1

La réforme des chemins de fer japonais

Avant la réforme de 1989, l'entreprise publique *Japan National Railways*, fondée en 1949, assurait la quasi totalité des transports ferroviaires à longue distance. Mais l'entreprise voyait ses parts de marché décliner dans l'ensemble des déplacements interurbains (de 45 % en 1965 à 23 % en 1985 en termes de passager × kilomètre). Par ailleurs la rentabilité baissa à partir de 1964, les pertes apparurent et crurent régulièrement, jusqu'à 1 000 milliards de yen (environ 8 milliards de dollars) en 1980, et l'entreprise publique coûtait de plus en plus cher au budget de l'État. L'analyse fut faite que cette situation découlait en partie des modalités de gestion de l'entreprise : trop forte intervention de l'État dans cette gestion, péreuxation des tarifs qui ne reflétaient pas les différences de coûts des différentes lignes, interventions politiques pour maintenir des services non rentables, antagonismes forts entre la direction et les salariés.

La réforme entreprise en 1987 consista à fragmenter les JNR en 7 entreprises se partageant le territoire japonais et assurant chacune sur leur territoire la gestion des infrastructures et l'exploitation des services voyageurs, et louant leur infrastructure à la société de fret, le 8^e démembrement de JNR. Par ailleurs, les entreprises étaient débarrassées de la tutelle publique, libres de fixer leurs tarifs, leurs services, de gérer à leur guise et de se diversifier ; leur statut privé ne se traduisit par immédiatement par une diffusion de leurs actions dans le public ; celles-ci furent d'abord la propriété d'une holding publique, et n'ont été que récemment cotées en bourse.

Le résultat fut une meilleure adaptation aux besoins des consommateurs : accroissement des services offerts, amélioration des temps et des fréquences des conditions de correspondance, faible augmentation des tarifs, entraînant une augmentation du trafic de près de 20 % entre 1987 et 1991. La productivité s'est améliorée considérablement, surtout dans les premières années qui ont suivi la privatisation : la productivité du travail par wagon × km a augmenté de 15,5 % par an entre 1985 et 1989, et ensuite de 7,8 % contre 4,5 % par an auparavant.

1. Les concessionnaires devaient reprendre en priorité le personnel de la compagnie japonaise, le personnel excédentaire faisant l'objet de plans de reclassement et de mises à la retraite anticipées.

Cette amélioration provint pour une très large part de la réduction des effectifs qui passèrent de 280 000 à 160 000 en 1996. Maintenant la productivité des JRs se compare à peu près à celle des compagnies privées, même si elle leur est encore inférieure, d'environ 20 %.

Par ailleurs, les JRs ont développé les activités de diversification : urbanisme commercial autour des gares, tourisme, vente de terrains. Enfin, le budget de l'État a été soulagé : les déficits se sont transformés en excédents. Les moteurs de cette transformation sont la libération de la tutelle publique et l'autonomie de gestion, l'aiguillon de la concurrence : concurrence avec les autres modes, et surtout l'avion ; concurrence par comparaison entre compagnies et notamment avec les compagnies privées anciennes ; la concurrence réelle ne se produit qu'aux franges des réseaux, sur quelques relations desservies par des compagnies différentes.

Les points noirs résident dans la situation des petites dessertes, qui s'est renchérie et dont la qualité de service s'est réduite ; dans les conséquences pour le personnel – le personnel et les syndicats furent les grands perdants –. Par ailleurs, la dette accumulée au cours de l'ancienne gestion se résorbe mal : on comptait sur les ventes de terrains et sur les ressources de vente d'action, mais la mauvaise conjoncture économique n'a pas permis de remplir tous les espoirs en ce domaine.

1.2. En Europe

La démarche européenne est partie, comme dans les autres pays, de la constatation du déclin des chemins de fer et des charges croissantes qu'ils imposaient aux budgets publics. Elle exprime la volonté de donner un nouveau souffle à ce mode de transport, en raison notamment des possibilités qu'il offre en termes de substitut à la route pour lutter contre la congestion et les atteintes à l'environnement.

TABLEAU 9-1
Charges pour les budgets publics, en pourcentage
des recettes d'exploitation (année 1994)

RENFE (Espagne)	61,6 %
SNCF (Belgique)	46,0 %
SNCF (France)	31,8 %
CFF (Suisse)	36,0 %
NSB (Norvège)	61,8 %
DB (Allemagne)	36,3 %
FS (Italie)	36,7 %

Source : UIC. Annuaire statistique.

1.2.1. La directive 91-440 et ses suites

Cette volonté s'est traduite par la directive 91-440. Celle-ci établit un objectif général, qui est de faire des chemins de fer des compagnies indépendantes des États, commercialement viables, et tournées vers la satisfaction des besoins de marché. Pour cela un certain nombre d'orientations sont définies : la plus importante est la nécessité d'une séparation au moins comptable entre les activités de gestion de l'infrastructure et celles d'exploitation. Ensuite, une certaine dose de concurrence est introduite dans l'exploitation : des entreprises ferroviaires internationales peuvent utiliser les infrastructures pour acheminer du trafic de transit, ou pour réaliser des opérations de transport combiné international. Ce premier pas vers une éventuelle concurrence dans l'exploitation nécessite néanmoins la fixation de règles pour l'organisation du marché de l'exploitation ferroviaire ; ces règles ont été définies dans leur grande ligne dans les directives 95/18 et 95/19. La première fixe les conditions d'entrée sur le marché, celles qui doivent être satisfaites pour qu'un opérateur puisse prétendre à faire circuler des trains : sa capacité professionnelle et financière pour assurer des services internationaux sera vérifiée par une instance dont chaque État définira la nature et la composition, et les candidats devront satisfaire aux normes et réglementations en vigueur dans chaque pays. La directive 95-19 définit les procédures par lesquelles les candidats sont attribués et tarifés : une instance définie par chaque État sera chargée de ces tâches, avec un souci de non discrimination et de recherche de l'usage optimal de l'infrastructure. Dans l'allocation des lignes, une priorité pourra être donnée aux services d'intérêt public (transports suburbains), ou à ceux pour lesquels les infrastructures ont été spécialement prévues (cas des lignes à grandes vitesses). Les charges d'infrastructure devront être fixées, en accord avec le gestionnaire d'infrastructure, en fonction de la nature du service, de la période, des dépenses imposées en termes d'infrastructure, et de la situation du marché.

Enfin, un livre blanc « une stratégie pour revitaliser les chemins de fer communautaires », paru en 1996, met en débat un certain nombre de mesures qui viendraient compléter et renforcer celles qui résultent de la directive 91-440, en abordant successivement les aspects suivants :

- dispositions financières : les États-membres devraient libérer les chemins de fer de leur dette et assainir leurs finances, en se conformant aux règles communautaires sur les aides de l'État ;
- organisation du marché : séparation de la gestion des infrastructures et de l'exploitation en unités distinctes. Mise en place de « freeways ferroviaires », grands corridors dont l'accès serait ouvert et les conditions d'utilisation simplifiées ;
- obligations de services publics : elles devraient être exécutées à travers des contrats entre l'État et les opérateurs ;

En 1994 est intervenue une réforme beaucoup plus radicale, impliquant à la fois fragmentation en une centaine d'organismes et privatisation.

Les grandes lignes de cette réforme sont décrites dans la figure 9-1. Les différentes unités figurant sur ce graphique ont des statuts différents :

- Certaines sont privatisées, comme les compagnies de fret, les compagnies de matériel roulant, les compagnies de gestion de l'infrastructure (entretien et renouvellement), et de nombreuses autres compagnies de services.

- D'autres sont ouvertes à concessions à travers des appels d'offre. C'est le cas des 25 compagnies de voyageurs.

- *Railtrack*, le propriétaire des infrastructures, d'abord propriété de l'État, a été introduite en bourse en 1996. Ses ressources de 2 275 millions de L en 1995 proviennent des charges d'accès payées par les exploitants (voyageurs : 1 955 millions de L ; marchandises : 191 millions de L ; l'évolution des charges d'accès est bornée à 2 % de moins que l'inflation). L'évaluation des actifs avait été fixée à 5 600 millions de L, mais la vente des actions n'a rapporté que 1 950 millions, probablement en raison des incertitudes sur l'organisation du système ferroviaire et sur son évolution (Brasdhaw, 1996).

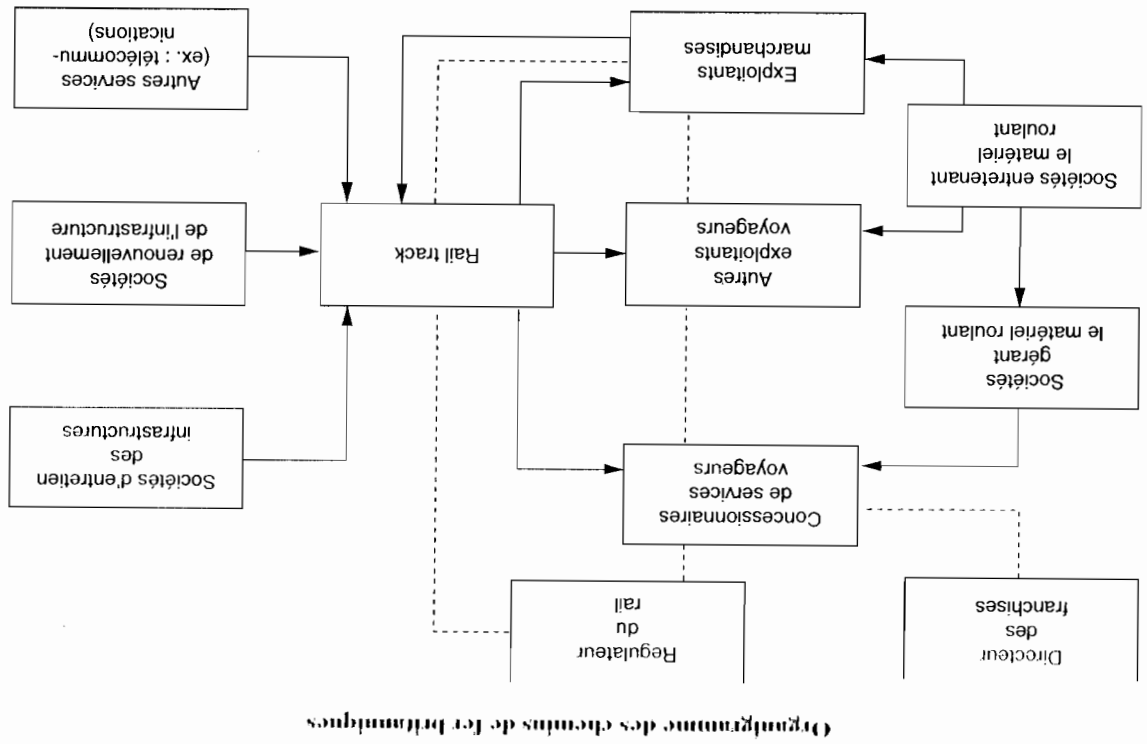
- Enfin apparaissent deux organismes de régulation, le *Rail Regulator* et l'*OFRRAP*.

1.2.2.b) La situation en France

En France, les réformes ont été beaucoup plus limitées. La SNCF procède à une réforme interne visant à séparer infrastructure, voyageurs et marchandises ; cette réforme touche progressivement l'ensemble des échelons hiérarchiques à partir du sommet. Par ailleurs, un nouvel établissement public, Réseau Ferré de France, a été créé en 1997 pour prendre en charge une partie de la dette antérieure de la SNCF, et surtout, pour l'avenir, pour assurer la gestion de l'infrastructure. Mais cet organisme léger s'apparente plus à un régulateur qu'à un gestionnaire. Il définira la politique d'investissement et d'entretien des infrastructures en tant que maître d'ouvrage, mais le maître d'œuvre est la SNCF, avec laquelle RFF passe des contrats pour l'exécution des travaux. Les ressources de RFF proviennent d'emprunts, de subventions de l'État et des charges d'accès que la SNCF et les opérateurs ferroviaires lui versent.

2. Analyse des expériences.

Débats et enjeux autour de la gestion des monopoles



- interopérabilité des réseaux des différents pays (harmonisation des normes techniques) ; cette interopérabilité sera favorisée ;
- mesures sociales, pour les reclassements et reconversions de personnels.

1.2.2. Les mises en œuvre selon les pays

Les directives et recommandations de la commission sont mises en œuvre de façon très différente selon les pays.

Le Royaume-Uni a élaboré une réforme allant bien au-delà des exigences communautaires, et d'ailleurs mise en œuvre bien avant ces directives. La Suède a également anticipé, mais sans procéder à des réformes aussi profondes ; l'Allemagne et les Pays-Bas suivent à peu près le rythme communautaire. La France, en revanche, est sensiblement en retrait par rapport aux pays qui viennent d'être cités.

ENCADRÉ 9-2

Les réformes des chemins de fer en Suède et en Allemagne

La réforme suédoise fut motivée par les faibles performances de l'entreprise qui assurait le transport ferroviaire : déficit élevé, lourdeur du processus décisionnel, imbrication et conflit entre les fonctions commerciales et de service public des S.J. En 1988, une loi fut votée qui stipulait la séparation entre l'infrastructure et l'exploitation (Hysten, 1997).

L'infrastructure est confiée à une entreprise BV (*Barverker*). Celle-ci plannifie les investissements d'infrastructure selon des critères de rentabilité collective prenant en compte les externalités. Leur financement est assuré par l'État, les utilisateurs acquittant des redevances d'usage calculés sur la base des coûts marginaux sociaux harmonisés entre les modes, calculés en tenant compte de l'imparfaite tarification routière.

L'exploitation des trains est essentiellement assurée par les S.J. (*Statens Järnvägar*), entreprise publique qui opère sur une base purement commerciale, sans subvention publique, et qui fixe librement les tarifs et les services assurés. Les S.J. ont un quasi monopole sur les grandes lignes ; les lignes régionales sont sous la maîtrise des autorités organisatrices de transports qui définissent les services à assurer et peuvent contracter avec d'autres opérateurs que les S.J.

Les flux financiers nets entre l'État et les chemins de fer sont passés de 2 277 millions de SEK en 1987 à 7 281 en 1993 ; en contrepartie les investissements d'infrastructure se sont considérablement accrus de 882 à 7 213 entre les deux mêmes années. Par ailleurs, les subventions publiques à l'exploitation se sont sensiblement réduites (de 1 400 à 700 M SEK), et ne concernent plus que les lignes régionales et certaines lignes nationales non rentables identifiées par des contrats spécifiques. Les redevances d'accès, nulles en 1987, se montaient à 659 M SEK en 1993. Enfin, la réforme s'est accompagnée d'une

restructuration des S.J., avec une baisse des effectifs, qui sont passés de 29 000 à 14 000, une organisation en « *business-unit* » tournée vers la satisfaction du consommateur.

Comme on le voit, la réforme n'a pas eu pour effet de réduire les appels au budget public, mais d'améliorer la qualité et l'efficacité du service, et en conséquence la fréquentation : depuis 1988 le trafic ferroviaire a augmenté de près de 30 % et le trafic aérien a baissé de 15 % environ. On peut toutefois s'interroger sur l'avenir du système, qui actuellement repose sur une coopération entre B.V. et S.J. : quelle est la précision des critères d'investissement des B.V., compte tenu de la prise en compte des externalités, et des prévisions d'utilisation des S.J., inoins fiables que si l'entreprise était intégrée ? Comment articuler les investissements de matériel roulant de S.J. et ceux d'infrastructure de B.V. notamment pour le développement des grandes vitesses (train pendulaire ou ligne nouvelle ?). En tout cas, il semble que la séparation infrastructure-exploitation n'ait pas posé de problèmes d'exploitation opérationnelle ; mais il n'y a qu'un exploitant. La situation pourra être différente au fur et à mesure que la concurrence va se développer. Ainsi, depuis 1996, les régions peuvent faire circuler des trains sur les lignes principales, et le trafic de marchandises est déréglé.

En Allemagne, la réforme a été plus récente (Hafner, 1996) et a suivi les principes de la commission. Les relations entre l'État et *Deutsche Bahn AG* (résultat de la fusion entre les chemins de fer des deux Allemagnes) ont été normalisées dans le sens d'une réduction des possibilités d'intervention dans la gestion. Par ailleurs, une restructuration interne de l'entreprise a été entreprise : séparation en 4 Divisions (régional voyageurs, national voyageurs, fret, infrastructure). En 1998 ces divisions deviendront des filiales de DB AG et ultérieurement elles pourront être privatisées. Ce changement de l'organigramme s'est accompagné d'une réduction des effectifs, passés de 350 000 en 1993 à environ 250 000 en 1996 ; par ailleurs, l'État prend en charge les investissements d'infrastructure sous la forme de prêts sans intérêt. Le bilan de l'opération devrait néanmoins être favorable au budget public : les promoteurs de la réforme ont calculé qu'elle ferait gagner en 10 ans 100 milliards de DM au contribuable. Enfin, la concurrence est introduite dans l'exploitation ; les länder peuvent concéder leurs services régionaux à d'autres opérateurs que la DB et l'entrée sur le marché du transport de marchandises est libre. Les tiers utilisateurs paient une redevance d'infrastructure, leur nombre se développe rapidement, mais la part de trafic qu'ils assurent reste limitée.

1.2.2.a) Le cas de la Grande-Bretagne

C'est en Grande-Bretagne que la réforme des chemins de fer a été la plus complète. Pendant une période de plus de 10 ans, les *British Rails* avaient déjà été réorganisés et restructurés. Ces efforts avaient porté certains fruits, puisque les subventions publiques étaient passées de 1 430 millions de livre en 1983 à 1 035 en 1992, pour un trafic un peu accru, et une productivité bien améliorée (Nash, 1997).

En 1994 est intervenue une réforme beaucoup plus radicale, impliquant à la fois fragmentation en une centaine d'organismes et privatisation.

Les grandes lignes de cette réforme sont décrites dans la figure 9-1. Les différentes unités figurant sur ce graphique ont des statuts différents :

- Certaines sont privatisées, comme les compagnies de fret, les compagnies de matériel roulant, les compagnies de gestion de l'infrastructure (entretien et renouvellement), et de nombreuses autres compagnies de services.
- D'autres sont ouvertes à concessions à travers des appels d'offre. C'est le cas des 25 compagnies de voyageurs.
- *Railtrack*, le propriétaire des infrastructures, d'abord propriété de l'État, a été introduite en bourse en 1996. Ses ressources de 2 275 millions de L en 1995 proviennent des charges d'accès payées par les exploitants (voyageurs : 1 955 millions de L ; marchandises : 191 millions de L ; l'évolution des charges d'accès est bornée à 2 % de moins que l'inflation). L'évaluation des actifs avait été fixée à 5 600 millions de L, mais la vente des actions n'a rapporté que 1 950 millions, probablement en raison des incertitudes sur l'organisation du système ferroviaire et sur son évolution (Bradshaw, 1996).

- Enfin apparaissent deux organismes de régulation, le *Rail Regulator* et l'OFRAP.

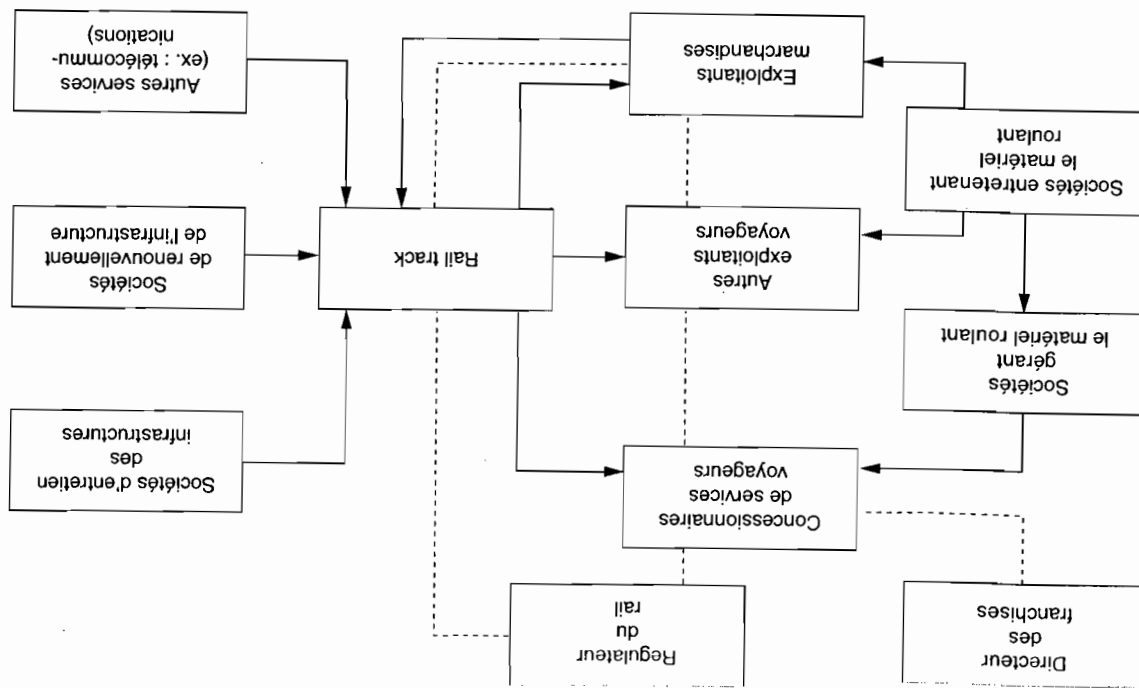
1.2.2.b) La situation en France

En France, les réformes ont été beaucoup plus limitées. La SNCF procède à une réforme interne visant à séparer infrastructure, voyageurs et marchandises ; cette réforme touche progressivement l'ensemble des échelons hiérarchiques à partir du sommet. Par ailleurs, un nouvel établissement public, Réseau Ferré de France, a été créé en 1997 pour prendre en charge une partie de la dette antérieure de la SNCF, et surtout, pour l'avenir, pour assurer la gestion de l'infrastructure. Mais cet organisme léger s'apparente plus à un régulateur qu'à un gestionnaire. Il définira la politique d'investissement et d'entretien des infrastructures en tant que maître d'ouvrage, mais le maître d'œuvre est la SNCF, avec laquelle RFF passe des contrats pour l'exécution des travaux. Les ressources de RFF proviennent d'emprunts, de subventions de l'État et des charges d'accès que la SNCF et les opérateurs ferroviaires lui versent.

2. Analyse des expériences. Débats et enjeux autour de la gestion des monopoles

Les solutions choisies par les différents pays sont diverses ; mais elles présentent un aspect commun, c'est qu'elles ont eu pour objectif et pour résultat d'ensemble une dynamisation des compagnies de chemins de fer,

Figure 9-1
Programme des chemins de fer britanniques



Source : Bradshaw, 1996.

une amélioration de leur productivité et de la satisfaction de la clientèle, une gestion du personnel plus souple et moins coûteuse. Elles ont pour cela mis en œuvre des mécanismes économiques variés dont les avantages et inconvénients vont être maintenant examinés : fragmentation des entreprises, modalités d'accès au réseau, instauration de la concurrence, nature du contrôle publique.

2.1. Modalités de fragmentation

La fragmentation est une constante de la plupart des réformes ferroviaires. Elle répond à deux objectifs : dans presque tous les cas, en réduisant la taille d'un organisme lourd à gérer, on cherche à le rendre plus maniable et plus réactif en réduisant les coûts de motivation (chap. VI, encadré 6-3). Souvent aussi la fragmentation est une condition préalable à l'introduction de la concurrence dans une partie du système. Elle peut prendre différentes formes.

2.1.1. Fragmentation géographique

La séparation géographique peut être justifiée par les limites des économies d'envergure. Des études économétriques (Jara-Diaz et Muniza, 1992 ; Preston, 1994) tendent à montrer que dans les grands réseaux européens tels que l'Allemagne, la France, l'Angleterre, la fonction de coût présente des déséconomies d'envergure, c'est-à-dire qu'un découpage en quelques réseaux géographiquement distincts conduirait à des coûts plus faibles. On peut toutefois objecter à ce raisonnement que, si les réseaux étaient effectivement séparés, des coûts supplémentaires pourraient apparaître dans la fourniture du matériel (commandes de taille plus réduite et non coordonnées), et que des effets de réseau sur la demande risqueraient de disparaître. Il faudrait, pour pallier ce dernier inconvénient, qu'un système d'information et d'uniformisation des systèmes billettiques soit mis en place, éventuellement par intervention publique.

Enfin la fragmentation géographique peut avoir des effets négatifs sur l'efficacité lorsque les deux organismes séparés fournissent des services complémentaires dont les tarifications et la fourniture en quantité et qualité doivent être coordonnées, comme ce peut être le cas pour une ligne principale et son feeder ou pour un aéroport « hub » et ses « spokes », comme l'ont montré Oum, Zhang et Zhang (1996).

2.1.2. Séparation horizontale

La séparation horizontale est celle qui consiste par exemple à séparer l'exploitation des voyageurs et celle des marchandises. L'intégration des deux activités permet d'absorber les pointes d'activité, à la fois en hommes et en matériel. Mais elle a un coût, celui de la polyvalence des hommes et du matériel (Henry et Quinet, 1996) : une locomotive adaptée à la fois aux marchandises et aux voyageurs est plus coûteuse qu'une locomotive spécialisée.

2.1.3. Séparation verticale

La séparation verticale est, dans les chemins de fer, celle qui serait introduite entre l'infrastructure et les activités. Elle conduit à abandonner la coordination hiérarchique à l'intérieur d'un même organisme, et à lui substituer une coordination par marché ou par marchandage entre des entités séparées, avec l'intervention plus ou moins directe d'un régulateur.

Il est d'abord clair que la séparation entraîne une augmentation des coûts de transaction. L'établissement des horaires et des priorités entre les services ne nécessite qu'une conférence interservice au sein d'une entreprise intégrée ; lorsqu'elle est le fait de deux entités juridiques différentes, elle entraîne des contrats volumineux et complexes, source de contentieux : qu'on pense par exemple à la gestion opérationnelle au jour le jour et à la manière de faire face aux incidents.

Certes, on peut supposer que la décision qui en résultera sera meilleure, par ce que fondée sur une confrontation plus claire des intérêts en jeu et débarrassés des considérations annexes (tradition en faveur d'un service, prestige des personnes, origine des décideurs) qui peuvent vicier la politique d'un monopole intégré, notamment lorsque ses objectifs de gestion ne sont pas clairement définis (cas des entreprises publiques). On pense généralement en France que l'allocation des sillons entre voyageurs et marchandises donne une trop forte priorité aux premiers.

Mais cet argument tombe lorsqu'il s'agit de décisions de long terme, notamment d'investissement. Les investissements d'infrastructure et de matériel roulant sont très liés, par les spécifications techniques (vitesse de base, poids des wagons...) et par les interactions des systèmes de signalisations et de sécurité. Chacun des partenaires sera incertain sur les choix de l'autre, sauf contrat de long terme, donc complexe et incertain, avec un risque de sous-investissement, sauf intervention d'une autorité publique. Encore faut-il qu'on puisse identifier les partenaires, ce qui peut s'avérer impossible si l'accès des tiers au réseau se développe.

2.2. La concurrence par le marché comme contrôle des monopoles

Un des objectifs avoués de nombreuses réformes et une des possibilités offertes par la fragmentation est la mise en place d'une concurrence entre opérateurs. Cette concurrence, on en attend une réduction des inefficiences statiques (réduction de la production) et dynamiques (gaspillages). Mais l'optimum collectif y perd la duplication des coûts fixes, dont d'ailleurs la présence constitue un obstacle à l'instauration de cette concurrence. Aussi n'obtient-on en général que des formes détournées de concurrence.

2.2.1. La concurrence à la frange et la concurrence par comparaison

La concurrence au sens plein et usuel n'existe nulle part dans les chemins de fer. La forme qui s'en rapproche le plus est la concurrence à la

frange où, à côté d'un opérateur principal détenant la majorité du marché, on rencontre un ou plusieurs autres petits opérateurs. C'est cette concurrence qui existe sur le marché du fret ferroviaire en Allemagne et en Suède.

Cette concurrence à la frange s'exerce aussi dans le cas des chemins de fer japonais. Les 6 compagnies de voyageurs disposent en fait de monopoles locaux et ne sont soumises à la concurrence qu'aux limites de leur réseau, dans les situations assez rares où les mêmes relations sont assurées par deux compagnies voisines.

La situation japonaise fournit aussi des éléments de concurrence par comparaison, une comparaison résultant de l'analyse des comptes et des performances des 6 compagnies qui opèrent parallèlement et indépendamment sur le territoire nippon. C'est également une concurrence de ce type qui peut être attendue de la fragmentation géographique des services ferroviaires de voyageurs en Grande-Bretagne ou que l'on peut obtenir de la comparaison de coûts et de statistiques concernant différents pays. Ainsi Rémy (1996) analyse les performances des principaux opérateurs ferroviaires de l'OCDE à partir des indicateurs tels que les indicateurs de productivité ou les indicateurs financiers usuels ; Oum et Yu (1994) utilisent des méthodes économétriques plus raffinées fondées sur la détermination d'une frontière de production et sur le calcul de la distance de chaque opérateur à cette frontière, distance mesurant l'inefficacité de l'opérateur. La Grande-Bretagne fournit aussi, par les concessions, l'exemple d'une concurrence pour le marché, qui sera analysée plus bas.

2.2.2. La concurrence potentielle et la résistance des entreprises en place

La concurrence peut être effective, par la présence de plusieurs entreprises sur le marché. Elle peut être aussi potentielle. La théorie des marchés contestables analyse les effets de la concurrence potentielle.

Développée par Baumol, Panzar et Willig (1982), elle repose sur l'idée qu'un monopole ne peut pas exercer pleinement son pouvoir et effectuer des profits en raison de la menace que fait peser sur lui la concurrence potentielle, sous la condition que l'entrée et la sortie du marché soient sans coût et que la réaction du monopole à une entrée ne soit pas immédiate. Alors, en cas de profit du monopole, un concurrent peut s'installer sur le marché, l'écarter tant que le monopole ne réagit pas, et ensuite se retirer.

Tableau de coûts d'entrée et de sortie d'un monopole non l'absence de

concurrence potentielle contrôlerait l'activité avec plus d'efficacité que la réglementation. En fait, il apparaît que l'entrée et la sortie sont rarement sans coût. La contestabilité n'est pas une notion « tout ou rien », elle fonctionne avec des limites (Morrison et Winston, 1997).

Par ailleurs, l'entreprise en place, qui cherche à réduire les possibilités de concurrence, peut avoir intérêt à empêcher l'entrée sur le marché de concurrents¹. L'analyse économique a recensé les moyens de blocage de l'entrée par l'entreprise en place :

- une surcapacité qui permet de produire à ses coûts marginal et/ou d'accroître les quantités mises sur le marché, provoquant une baisse des prix et donc de la rentabilité de l'entrée,
- un développement des dépenses de recherche-développement, signalant des baisses de coût,
- un haut niveau de production, fournissant à l'entreprise en place un niveau d'expérience élevé,
- l'acquisition de licences et brevets,
- la diversification des produits, comblant les niches éventuellement disponibles pour l'entrant,
- l'utilisation des prix comme un signal adressé à l'entrant potentiel sur les conditions de l'offre ou de la demande. Ainsi, un prix bas fera croire à l'entrant éventuel que les coûts de l'entreprise en place sont bas²,
- la réputation. Ainsi une compagnie luttera contre un entrant éventuel sur une ligne, même si ce n'est pas son intérêt immédiat, pour se faire une réputation d'agressivité et dissuader les tentatives d'entrées sur d'autres lignes.

Parmi ces moyens, il semble que ce soient surtout les premiers, les surcapacités et les dépenses de recherche-développement, ainsi que la maîtrise des ressources rares, qui constituent les barrières à l'entrée les plus fréquemment utilisées dans les transports, notamment dans les chemins de fer et dans les transports routiers de voyageurs, comme on l'a vu au chapitre précédent.

2.3. L'accès au réseau

La mise en place d'une concurrence sur le marché, même sous les formes réduites qu'on vient d'évoquer, pose le problème des conditions et de la tarification de l'accès au réseau pour les exploitants.

2.3.1. Les conditions d'accès

Les conditions d'accès sont définies de manière très générale par les directives communautaires.

Elles concernent d'abord la certification des entreprises, qui comporte en général des clauses financières et techniques : capacité d'assurer une sécurité et une régularité suffisante pour ne pas perturber la marche des autres trains.

Elles concernent aussi les critères d'attribution des sillons, sur lesquelles on se penchera plus particulièrement. Actuellement cette attribution est essentiellement fondée sur des règles de priorité ; en général la priorité va aux services publics suburbains, puis aux grandes lignes, puis au fret. En France, c'est la SNCF qui attribue les sillons, sous l'autorité et avec le recours possible du régulateur, en l'occurrence RFF, à partir de grilles horaires qui fixent des priorités entre les catégories de trains selon les heures et les types de voies. On peut craindre que cette attribution faite par un des concurrents ne soit biaisée en sa faveur, et le rôle du régulateur sera à cet égard primordial.

Une manière efficace d'attribuer les sillons serait de recourir à un système d'enchères, par exemple le système qui consisterait à faire payer au plus offrant le prix de la plus élevée des offres non retenues. C'est le système d'enchères de Vickrey, qui présente la propriété que chacun des participants est incité à déclarer son vrai prix. Nilsson (1995) a mené des expériences d'enchères fictives tendant à montrer que ce système est praticable et aboutit à dégager dans la plupart des cas les allocations les plus efficaces. Le principe est le suivant : chaque enchérisseur définit les sillons qu'il désire et la valeur qu'il accorde à chacun d'eux, ainsi que la diminution de valeur résultant d'un éventuel décalage dans le temps par rapport à l'horaire désiré. Ensuite, le planificateur optimise l'ensemble, c'est-à-dire attribue les sillons, éventuellement décalés, de façon à maximiser la valeur totale de l'ensemble du graphique¹. Le résultat de cette maximisation est le même que celui d'un système d'enchères ou les individus se comporteraient en déclarant la vérité. La procédure est encore expérimentale en raison de sa complexité due aux conflits entre services de vitesses différentes, et aux nécessaires complémentarités à assurer dans l'attribution de sillons voisins. Une autre procédure possible d'enchères est décrite dans Brewer et Plott (1996) ; elle est fondée sur le principe d'enchères ascendantes et simultanées pour tous les sillons. Mais ces systèmes en sont encore à l'état de recherche et ne sont pas encore opérationnels.

En particulier ils ne sont pas mis en œuvre dans les entreprises intégrées, alors que qu'ils permettrait à l'entreprise de choisir ses services de la façon la plus rationnelle. Ils régleraient à la fois la question de l'allocation des sillons et celle de leur tarification. La Grande-Bretagne met en

œuvre une procédure qui s'en rapproche : tous les six mois, des sillons supplémentaires sont ouverts par Railtrack et attribués par négociation, ce qui constitue une sorte d'enchère informelle. Mais la qualité du résultat est sous la dépendance des asymétries d'information dont souffre le régulateur. En dehors de ce cas, dans la pratique actuelle, les deux étapes sont séparées. L'attribution des sillons dépend peu d'un système de filtrage par les tarifs d'infrastructure et les grilles de tarification sont beaucoup plus fondées sur les conditions de coût que sur les conditions de demande (valeur des sillons pour l'utilisateur).

2.3.2. La tarification de l'accès

En l'absence d'un système d'enchères, l'attribution d'un sillon entraîne le paiement d'un tarif. Dans une optique de recherche d'un optimum premier, la tarification optimale devrait imputer à chaque allocataire d'un sillon le coût marginal d'usage du sillon (dépenses marginales d'entretien et d'exploitation pour le gérant d'infrastructure) ; il faudrait y ajouter le coût d'opportunité : c'est le coût supplémentaire que la présence de l'allocataire impose aux autres exploitants en raison des déplacements ou suppressions de sillons, ou par pertes de recettes infligées aux concurrents non déplacés mais atteints dans leur clientèle. On a vu au chapitre VII la difficulté que présente sa détermination en raison du caractère particulier de la congestion ferroviaire, à la suite des deux dimensions de cette congestion : la localisation et le temps. Il faudrait y ajouter, comme on l'a vu, des pénalités lorsque l'horaire prévu n'est pas respecté et que les retards d'un train occasionnent des pertes aux trains suivants. Cette procédure de tarification serait le pendant pour le cas ferroviaire de la tarification au coût marginal social (on peut ici difficilement parler de coût marginal, s'agissant de variations discrètes de la clientèle).

La tarification au coût marginal – ou plutôt ici au coût d'opportunité – n'est valable que dans une situation d'optimum premier. Lorsque les fonds publics ont un coût, ou lorsque le mode concurrent n'est pas tarifé de façon optimale (c'est le cas du transport routier de marchandise par gros porteur, tarifé en dessous du coût marginal, et concurrent direct du chemin de fer), ou lorsque pour des raisons d'équité ou d'incitation à un bon niveau d'investissement, on veut assurer un certain niveau de recouvrement des dépenses fixes, des adaptations sont à mettre en œuvre. Pour la couverture des dépenses deux voies sont possibles.

La première est celle d'une tarification binôme, (voir encadré 8-5 chapitre V III). Ce point est développé par exemple dans Caillaud et Quinet (1992b) qui analysent les autres justifications d'une tarification dégressive pour les sillons ferroviaires, par exemple la dégressivité des coûts équilibrés avec le volume des services et avec la durée du contrat. Elle peut aussi constituer un moyen de correction des imperfections du marché aval et celui-ci est oligopolistique : pour corriger l'effet de double monopolisation en chaîne il convient que le monopole d'amont (ici le gérant d'infrastructure) tarifie en dessous du coût marginal ; un moyen de combler le

1. Dans certains cas le décalage ne sera pas suffisant, la suppression d'un sillon sera nécessaire.

déficit qui en résulte est d'imposer un droit fixe, ce qui aboutit à un tarif binôme. Enfin, le tarif dégressif peut aussi rejoindre les raisons de la pratique commerciale courante, pour les contrats de long terme, et constitue une sorte d'assurance du gérant d'infrastructure vis-à-vis du risque. Ce système est celui pratiqué par *Railtrack* dans ses négociations avec les concessionnaires de services voyageurs et avec l'entreprise exploitant le fret.

La tarification binôme présente toutefois l'inconvénient majeur de désavantager l'entreprise qui cherche à entrer sur le marché. Un menu de tarifs, avec des parts fixes d'autant plus faibles que les parts variables sont élevées, atténuerait cet inconvénient sans le supprimer. Ces considérations conduisent à préférer une tarification du type Ramsey-Boiteux (Nera, 1998), telle qu'elle est présentée au chapitre VII.

2.3.3. Le cas de l'accès des tiers au réseau

Un problème particulier se pose si, comme c'est souvent le cas, le gestionnaire d'infrastructure est également un des exploitants, et se trouve ainsi être le fournisseur de son concurrent. Laffont et Tirole (1993, 1994) ont étudié le problème, et les grandes lignes présentées ci-dessous s'inspirent de leur analyse, qui aboutit à une forme de tarification de Ramsey-Boiteux.

ENCADRÉ 9-3

Problèmes de tarification de l'accès des tiers au réseau

Soit un opérateur ferroviaire dont les coûts fixes sont k et donc le coût marginal de gestion de l'infrastructure est c_0 . Il est également opérateur, assure deux sortes de trafics q_1 et q_2 et son coût marginal pour l'exploitation est c . D'autres opérateurs lui font concurrence sur le réseau ; ils lui achètent des sillons à un prix a et ont un coût d'exploitation de γ pour un trafic q_3 , substitut de q_2 .

Le trafic total est donc :

$$Q = q_1 + q_2 + q_3$$

L'optimum premier conduirait à fixer un coût d'accès égal à c_0 , mais entraîne un déficit k . Si les fonds publics ont un coût λ , son financement peut être récupéré sur les produits finaux, par une formule du type Ramsey-Boiteux¹ :

$$\begin{aligned} \frac{P_1 - c - c_0}{P_1} &= \frac{\lambda}{1 + \lambda} \frac{1}{\varepsilon} \\ \frac{P_2 - c - c_0}{P_2} &= \frac{\lambda}{1 + \lambda} \frac{1}{\eta_2} \\ \frac{P_3 - \gamma - c_0}{P_3} &= \frac{\lambda}{1 + \lambda} \frac{1}{\eta_3} \end{aligned}$$

1. Il peut aussi l'être par un tarif binôme, comme on l'a vu dans l'encadré 8-5.

Les η sont les superélasticités définies au Chapitre VII (section 2.2.1, encadré 7-5).

Si l'on se rappelle que, pour des biens substitués, les superélasticités sont plus faibles que les élasticités directes on voit que les 2 biens en concurrence seront plus frappés que le bien 1 qui n'a pas de substitut, à l'intérieur du système considéré.

Le recouvrement des charges fixes peut donc se faire par une taxe à la production, différenciée par rapport à la taxe Ramsey-Boiteux simple, et frappant proportionnellement plus fortement le plus gros opérateur, et en outre d'autant plus forte que les substitutions entre produits le sont.

Si l'État peut imposer les prix p_1 et p_2 au monopole, mais pas la taxe à l'opérateur concurrent, l'optimum pourra être réalisé en imposant des prix d'accès différenciés :

$$c_0 + \frac{\lambda p_3}{(1 + \lambda) \eta_3} \quad \text{à l'opérateur concurrent}$$

$$c_0 + \frac{\lambda p_2}{(1 + \lambda) \eta_2} \quad \text{au monopole pour le trafic } q_2$$

$$c_0 + \frac{\lambda p_1}{(1 + \lambda) \varepsilon_1} \quad \text{au monopole pour le trafic } q_1$$

On peut comparer ce type de résultat avec celui que donne une règle fréquemment utilisée pour la tarification de l'accès, celle de « l'efficient component pricing » (Laffont et Tirole, 1994), qui revient à imputer au concurrent le manque à gagner que sa présence impose au monopole, soit ici :

ou encore :

$$a = p_2 - c$$

$$a = c_0 + \lambda(1 + \lambda) p_2 / \eta_2$$

On voit que les règles coïncident si les deux biens substitués ont les mêmes caractéristiques tant en termes de demande (élasticités identiques) que d'offre (prix équivalents, ce qui implique des coûts marginaux équivalents). Elle peut, dans les autres situations, constituer une première approximation et une référence.

2.3.4. Considérations pratiques

Les considérations qui précèdent peuvent être complétées dans trois directions. La première concerne l'extension du domaine susceptible de payer une contribution aux dépenses fixes. Rien, dans ce qui précède, ne conduit à limiter cette contribution aux opérateurs qui bénéficient de l'accès à l'infrastructure. Le déroulement du raisonnement qui conduit à la tarification de Ramsey-Boiteux montre au contraire qu'il est bénéfique de l'étendre aux produits substitués, et cela, d'autant plus que la substitution est forte. C'est ainsi qu'on pourrait atteindre une plus grande efficacité en faisant participer le trafic routier aux dépenses fixes du chemin de fer, et réciproquement d'ailleurs.

La deuxième direction peut conduire à infléchir les résultats. Ceux-ci ont été établis avec quelques hypothèses implicites, à savoir que le gérant d'infrastructure est un agent docile dans les mains d'une puissance publique dont il épouse les objectifs et les contraintes ; que cette puissance publique connaît parfaitement toutes les données d'offre et de demande du marché ; enfin, que le monopole a un comportement passif à l'égard de ses concurrents, et qu'il ne tente aucune manœuvre pour contrecarrer leur action. Ces hypothèses sont évidemment simplistes : le monopole a ses objectifs propres, différents de ceux de l'autorité publique, et il va essayer de les réaliser en utilisant sa supériorité d'information par rapport à la puissance publique, pour accroître ses avantages ou ses profits, réduire les tensions qui s'exercent sur lui, et en particulier s'efforcer d'éliminer la concurrence. Ceci peut conduire l'autorité de tutelle à réduire les charges d'accès, de façon à faciliter l'entrée sur le marché. Des considérations d'incitation ou d'asymétrie d'information peuvent militer en sens inverse (Laffont et Tirole, 1996), et conduire à proposer des menus de tarifs judicieusement calculés qui conduiraient les utilisateurs performants à choisir une part fixe élevée et une part variable plus faible, et permettrait de profiter de leur bonne productivité¹.

La troisième direction résulte de la structure de réseau et souligne la complexité de la couverture des coûts fixes. Ainsi, dans le cas de la tarification de Ramsey-Boiteux, les tarifs devraient dépendre, non seulement des coûts, mais aussi des conditions de la demande, et donc être définis par origine-destination.

Face à ces considérations complexes, les pratiques actuellement enregistrees sont frustes. Les tarifications mises en place sont en général fondées sur des formules faisant intervenir un terme fixe et des termes proportionnels au kilométrage un terme qui dépend du sillon horaire et qui rend compte de la congestion, un terme qui dépend du type de train et qui rend compte de la capacité contributive de chaque catégorie de train, et enfin un terme qui traduit les coûts d'entretien et d'exploitation. On en est actuellement à une période de tâtonnements, au cours de laquelle les tarifs vont s'ajuster progressivement, et on ne peut déceler que certains grands traits selon les pays et son activité commerciale.

Aux Pays-Bas les infrastructures sont financées et gérées par l'État, et leur tarification est très faible. On voit bien la raison de cette politique qui avantage les ports du pays.

La Suède a établi un système où les charges d'infrastructure sont également faibles ; elles sont fondées sur une tarification au coût marginal amendée et réduite pour tenir compte de la sous-tarification de la route en termes d'externalités.

1. À l'inverse, les entreprises peu performantes seront induites à choisir dans le menu proposé une rémunération à part fixe plus faible. Ceci ne les incitera pas à une productivité élevée, mais l'incitation coûterait très cher. Les mécanismes régissant la fabrication de ces menus sont similaires à ceux qui président à la différenciation des prix (chap. VIII, encadré 8-4).

En France la tarification provisoire actuelle, valable jusqu'en 1999, est établie à partir de principes similaires, mais en incluant des coûts de congestion évalués de façon probablement fort modeste par rapport aux coûts réels.

En Allemagne les charges d'infrastructures sont fixées à un niveau élevé visant à récupérer la totalité des coûts d'infrastructure supportés par l'entreprise ; mais cette entreprise est fortement subventionnée pour la réalisation des investissements.

Dans tous ces pays les charges sont essentiellement variables. Ainsi en France le montant total des charges s'élève à 6 milliards de Francs qui ne comporte que 300 millions de Francs de charges fixes.

En Grande-Bretagne, il y a couverture totale des coûts, mais cette couverture est essentiellement assurée par des charges fixes, celles qui sont négociées entre Railtrack et les concessionnaires d'opérateurs voyageurs ; en outre ces charges sont payées par l'État au travers du prix de la concession. On conçoit qu'il faille alors un contrôle actif et très bien informé du régulateur.

Dans quelques pays existe un système de pénalités pour irrégularités et retards : c'est le cas en Grande-Bretagne et en Suède ; en France le décret qui fixe les principes de la tarification en ouvre la possibilité, mais aucune application n'en a été faite dans le cadre de la tarification provisoire mise en place jusqu'en 1999.

Il ne faut pas s'étonner dans ces conditions que les tarifs réels soient très disparates, comme le montre par exemple la comparaison des tarifs suédois et anglais, tirée de Nash (1996) :

	Suède	Grande-Bretagne
Prix d'un train × km de marchandise en Livres	0,69	4,72

On imagine dans ces conditions la difficulté qu'il aura à harmoniser les charges d'infrastructures européennes. C'est une tâche que se fixe la Commission de l'Union Européenne, dans le souci de revigorer le chemin de fer et d'harmoniser les conditions de concurrence entre les pays.

2.4. Le contrôle public des monopoles

L'alternative à la discipline qu'impose la concurrence est le contrôle public des monopoles. Ce contrôle est particulièrement nécessaire lorsque ceux-ci sont investis d'une mission de service public, car c'est là que l'efficacité externe de leur action est le plus en contradiction avec l'objectif de profit.

2.4.1. L'organisation du contrôle

Ce contrôle peut prendre de multiples formes, qui diffèrent selon la propriété du monopole. Ainsi, dans l'URSS de la planification impérative,

le contrôle était direct et hiérarchique : les monopoles (comme d'ailleurs les autres opérateurs) étaient des services de l'administration.

Dans la structure européenne ancienne, à base d'entreprises publiques, il s'exerçait par le Ministère chargé de la tutelle de l'opérateur de service public, très proche de ce dernier ; et d'ailleurs nombre de tâches de réglementation étaient, en droit ou en fait, exercées par l'opérateur ; c'est la SNCF qui elle-même définissait les normes techniques de sécurité et les règles de police dans les lieux publics. Il y avait en outre une communauté d'état d'esprit et d'objectif entre les deux, et on a pu qualifier la situation de capitalisme de confiance. En France, bien que la SNCF reste entreprise publique, Réseau Ferré de France, établissement public, est maintenant dépositaire de certaines fonctions de régulation, notamment concernant les charges et les conditions d'accès, et le contrôle s'est en quelque sorte éloigné de l'État.

Dans les pays anglo-saxons, le régulateur et les opérateurs sont plus éloignés encore, la divergence de leurs objectifs est reconnue, et le statut des régulateurs est diversifié. La régulation exercée par l'administration est marginale à côté de celle exercée par des agences indépendantes (qui apparaissent aussi en France, voir l'Agence de Régulation des Télécommunications). La réglementation repose plus ou moins fortement sur l'action des tribunaux ; ainsi aux États-Unis les abus de position dominante, les infractions pour prix prédateurs sont combattus surtout par le pouvoir judiciaire, alors que ces fonctions sont davantage remplies par des agences indépendantes en Angleterre.

On voit les motifs de telles organisations : soustraire la régulation aux vicissitudes, aux pressions et à l'agenda du pouvoir politique, lui donner une stabilité et une continuité suffisante pour rendre ses actions efficaces et ses orientations crédibles. Le régulateur est, par les textes qui le définissent, investi d'un certain nombre de missions telles que : promouvoir une concurrence efficace et équitable, assurer le développement de l'industrie contrôlée, la satisfaction du consommateur et l'exercice du service public. C'est ainsi que le régulateur des chemins de fer britannique a un droit de regard (et de refus) sur les contrats liant Railtrack et les opérateurs, sur les prix et conditions d'accès à l'infrastructure, et sur les concessions de service voyageurs quant à leurs conséquences pour la concurrence et le développement de l'activité ferroviaire. Il dispose d'un certain nombre de moyens qui sont essentiellement les entrées sur le marché, les prix et l'approbation des contrats qui interviennent entre opérateurs. En revanche ce n'est pas lui qui choisit les opérateurs de transport de voyageurs. Cette fonction est remplie par l'*Office of Passenger Rail Franchising* (OFRAP) qui a donc par là un rôle de surveillance des tarifs ferroviaires.

Les difficultés sont multiples : c'est d'abord celle du choix du régulateur, qui doit présenter des qualités d'honnêteté, d'indépendance et de dévouement à l'intérêt général. C'est ensuite celle de sa tâche, puisqu'il doit guider des acteurs dans une situation d'asymétrie d'information quant aux conditions d'exercice de leur activité, concernant notamment

leurs coûts et leur marché, et éviter d'être capturé par le régulé, de finir par épouser le point de vue et devenir l'avocat de celui qu'il contrôle. C'est enfin celle du contrôle de son action ; le régulateur ne peut pas être un despoite tout-puissant, il convient de doser les modalités d'appel pour qu'elles ne nuisent pas à son indépendance tout en assurant les assujettis contre l'arbitraire de celui qui doit les contrôler.

L'efficacité du contrôle peut être amoindrie par des phénomènes de collusion ou de capture. L'importance de ces phénomènes dépend d'éléments tels que le mode de nomination des régulateurs, la durée de leur mandat, la précision avec laquelle sont définis leurs objectifs. Elle dépend aussi des actions de lobbying des parties prenantes en la régulation, parmi lesquelles les plus actifs sont en général les producteurs car ils sont plus concentrés que les utilisateurs, et leur action est moins sujette à externalités.

2.4.2. La régulation des prix

Elle a pour but d'éviter que le monopole rançonne le consommateur et produise une quantité sous-optimale. Elle a fait l'objet de nombreuses études théoriques qui ont porté sur les différents systèmes, à savoir :

- la régulation « *cost plus* » : le régulateur étudie les coûts de production des services et fixe les prix à un niveau légèrement supérieur pour attribuer un profit à l'opérateur,
- la régulation « *price-cap* » ou « *RPI-X* » : le régulateur fixe une évolution des prix forfaitaires qui tiennent compte de l'inflation (*Retail Price Index*), moins un certain pourcentage correspondant aux progrès de productivité attendus,
- la régulation « *rate of return* » dans laquelle c'est le taux de rendement du capital qui est plafonné.

L'analyse théorique et l'expérience ont fait apparaître les vertus et défauts de ces systèmes :

- le système « *cost plus* » n'est pas incitatif à l'effort puisque toutes les dépenses sont remboursées.
- le système « *price-cap* » à l'inverse, est très incitatif à l'effort puisque tous les gains de productivité sont empochés par l'entreprise. Mais ce système a un coût pour la puissance publique, dans le cas très général où il y a asymétrie d'information ou risque ; alors, pour attirer l'opérateur, il faut lui laisser une prime qui l'assure contre le risque et dépasse son aversion pour par exemple l'encadré 7-13).
- le système de contrôle « *rate of return* » par le rendement du capital peut avoir un effet sur la politique d'investissement du monopole. C'est l'effet Averch-Johnson, qui se produit lorsque le contrôle porte sur le rendement du capital (voir encadré 9-5). Le monopole a intérêt à surinvestir pour augmenter son profit.

Une adaptation du système *price-cap* en situation de multiproduction

Le cas d'entreprises multiproduits est particulier. Dans ce cas, la condition de maximisation du bien être social, pris sous la forme du surplus collectif traditionnel, auquel s'ajoute un coût des fonds publics, aboutit à une tarification du type Ramsey-Boiteux. Or, il se trouve qu'un système de contrôle des prix converge vers cette tarification : c'est le système proposé par Vogelsang et Finsinger (1979), dans lequel la contrainte porte sur une somme pondérée des prix des différents produits :

$$S = \sum_i w_i P_i$$

les poids étant égaux aux quantités de l'année précédente, et la valeur de S est égale aux dépenses totales de la période précédente (s'il n'y avait qu'un produit, la contrainte aboutirait à ce que le prix de la période t ne doit pas dépasser le coût moyen de $t-1$). Le mécanisme suppose toutefois que l'entreprise est myope ; dans le cas contraire elle pourra se préparer des profits futurs en augmentant ses dépenses immédiates, au démarrage du processus (R et D , investissements...).

Des stratégies similaires peuvent apparaître dans le cadre dynamique des régulations *cost plus* ou *price-cap* ; l'entreprise soumise au *price-cap* peut modérer ses efforts de productivité de crainte que le régulateur, acquérant avec le temps des informations sur les coûts et l'efficacité de l'effort, n'en tire argument pour réduire la hausse des prix autorisée. Le résultat dépend alors de l'annonce de sa stratégie à long terme et de la crédibilité des engagements ainsi pris par le régulateur.

L'effet Averch-Johnson

Soit un monopole produisant un bien q au prix p , avec du capital K , dont le coût unitaire annuel est i et du travail L (w étant le salaire) la fonction de production est :

$$q = f(K, L)$$

Si le régulateur impose un taux de profit maximum r , le monopole aura à résoudre le programme :

$$\max (pq - (iK + wL))$$

$$q - f(K, L) \leq 0$$

$$pq - (iK + wL) - rK \leq 0$$

Sa solution donne, avec les multiplicateurs de Lagrange usuels :

$$-i + \lambda f'_K + \mu(r+i) = 0$$

$$-w + \lambda f'_L + \mu w = 0$$

$$p - \lambda - \mu p = 0$$

D'où :

$$\lambda = p(1 - \mu)$$

et, comme :

$$\lambda \geq 0, \quad 1 - \mu \geq 0$$

et :

$$\frac{f'_K}{f'_L} = \frac{i - \mu(r+i)}{w - \mu w} = \frac{i}{w} - \frac{\mu r}{w(1 - \mu)}$$

Si $\mu = 0$, la contrainte ne joue pas, on aboutit à l'égalité classique :

$$\frac{f'_K}{f'_L} = \frac{i}{w}$$

Si $\mu = 0$, la contrainte joue et :

$$\frac{f'_K}{f'_L} < \frac{i}{w}$$

La productivité marginale du capital est inférieure à l'optimum, il y a surinvestissement.

Une analyse de la dynamique entre monopole et régulateur peut au contraire conduire le monopole à sous-investir. Cela se produira lorsqu'on se trouve en présence d'investissements spécifiques non réutilisables, et que, une fois ces investissements faits, le monopole craint que le régulateur ne fixe des prix trop bas, mettant en cause leur rentabilité. Le nœud de la question réside là aussi dans les engagements pris par le régulateur et dans leur crédibilité. Helm et Thompson (1991) ont analysé l'effet du système de contrôle sur l'incitation à investir de l'opérateur.

La régulation par les prix est, dans les transports, essentiellement une régulation *price-cap* ; c'est ainsi que fonctionne la régulation anglaise, celle qui est définie par le *Rail Regulator* et concerne la tarification des sillons, comme celle définie par l'OFRAP pour les tarifs voyageurs. Pour les concessions d'autoroutes en France, le niveau des péages est plafonné par les autorités régulatrices (en l'occurrence le Ministère des Finances), par référence au taux d'inflation ; mais l'écart avec le taux d'inflation n'est pas limité statutairement, il résulte d'accords non contractuels à caractère non juridique. Enfin les Sociétés concessionnaires sont souvent implicitement régularisées par une forme non officielle de régulation *rate of return* : si leur profit était trop élevé, elles risqueraient une forme de *hold-up*. Alors elles s'efforcent de les réinvestir, un moyen d'échapper à cette menace. La régulation *cost plus* est en vigueur dans certaines organisations de transport public : c'est en particulier celle qui fonctionne de fait pour la RATP, qui est remboursée de ses augmentations de coût non compensées par des augmentations de tarifs : la puissance publique lui rembourse le manque à gagner de façon que l'entreprise couvre ses coûts ; c'est le système du verbatim compensatoire.

3. L'exercice du service public

Entendu au sens large, le caractère de service public peut être attaché à nombre d'activités de transport. En France, la Loi d'Orientation des Transports Intérieurs (LOTI) du 31 décembre 1982 prend une position extrême et l'attribue à l'ensemble du secteur. On peut avoir des conceptions plus réduites, comme c'est le cas dans les pays anglo-saxons ; mais même alors il serait difficile d'en dénier le caractère aux taxis ou aux ambulances, ou aux transports collectifs assurant les déplacements de base tels que les déplacements urbains.

L'exercice du service public n'implique donc pas nécessairement la gestion publique, ni même le monopole. Les taxis et ambulances déjà cités sont dans tous les pays assurés par des opérateurs privés nombreux, en situation de concurrence les uns par rapport aux autres, et très réglementés par les pouvoirs publics. L'intervention publique est bien sûr d'autant plus forte que le service public comporte des contraintes par rapport à ce qu'entraînerait la gestion privée : exécution de services non rentables, tarification différente de ce qu'implique la recherche du profit.

Dans ces situations, l'intervention publique est nécessaire pour imposer à l'opérateur la contrainte souhaitée, pour l'en dédommager éventuellement si cette contrainte implique un déficit ; pour le protéger si l'on veut lui faire financer les services déficitaires par subventions croisées en provenance de services bénéficiaires (dans ce cas les services bénéficiaires sont fragilisés à l'égard d'une concurrence éventuelle qui pourrait se contenter de profits réduits, cf. chapitre VI section 3.2 et encadré 6-10).

Cette intervention peut prendre différentes formes.

3.1. La gestion publique

Un moyen apparemment simple pour être sûr que l'opérateur concoure aux objectifs de la puissance publique est de l'intégrer à cette puissance publique, en l'érigant en service administratif. C'est ainsi qu'étaient organisés les services ferroviaires et aériens dans les pays de l'Est européen après la guerre ; c'est aussi sous la forme d'entreprises publiques qu'ils étaient organisés dans la plupart des pays de l'Europe de l'Ouest ; ces entreprises avaient des statuts juridiques divers (service de l'administration, établissement public, entreprise à statut particulier...). Mais leur caractéristique commune était la toute puissance de l'État dans leur gestion. Ce système a bien fonctionné pendant de nombreuses années, puis sont survenus des vices de fonctionnement.

Il est apparu que la tutelle de l'État s'exerçait dans le désordre. L'État tenait plusieurs rôles de façon contradictoire : prescripteur de services publics, il imposait des obligations (tarifs sociaux, dessertes non rentables...) dont, en tant qu'actionnaire, il ne voulait pas assumer les conséquences financières ; pour des raisons de régulation macro-économique, il bloquait à certaines périodes les tarifs des entreprises, mais sans accepter

ENCADRÉ 9-6

La gestion bureaucratique

Ce mode de gestion, caractéristique des administrations publiques, s'oppose sur la plupart des points à la gestion privée et du marché :

- par l'origine des ressources utilisées : l'attribution d'un budget de la part d'une autorité supérieure plutôt que le produit d'une activité marchande,
- des emplois stables et un avancement à l'ancienneté, par opposition à une insécurité d'emploi, et à une progression en fonction des mérites et des résultats,
- le respect des normes et des règles prime sur la recherche et l'efficacité,
- le contrôle et la tutelle portent sur les moyens plus que sur les fins et sur les inefficaces,
- les communications et la circulation de l'information vont de haut en bas et circulent mal entre niveaux hiérarchiques, alors que dans la gestion privée la communication est ascendante, et est plus active dans le sens vertical que dans le sens horizontal.
- l'objectif est plus tourné vers la maximisation du budget de l'organisation que vers la recherche du projet de l'entreprise.

de financer les déficits qui en résulteraient. En outre, l'État était multi-fonctionnel et engendrait des interventions venaient de nombreuses administrations non coordonnées, parfois contradictoires et toujours tatillonnes. Les administrateurs de l'entreprise publique poursuivaient leurs objectifs propres dans le cadre d'un fonctionnement bureaucratique opposé au fonctionnement managérial (voir encadré 9-6) : l'avancement à l'ancienneté et non à l'efficacité, un éventail des salaires resserré, la recherche du volume d'activité et d'avantages discrétionnaires, sans compter d'autres caractéristiques de la gestion interne de l'organisme (développement des communications horizontales au détriment des communications verticales, hypertrophie des communications descendantes, respect de la routine, lenteur des adaptations).

Certes l'entreprise publique avait un objectif de maximisation du bien-être social, représenté par une fonction de surplus qui comprend à la fois son profit et les externalités engendrées et aussi d'éventuels objectifs de redistribution. C'est la théorie des « choix publics ». Mais la mesure dans laquelle cet objectif est effectivement poursuivi peut être analysé comme fondamentalement à travers l'examen des intérêts divergents des acteurs du processus : les électeurs, les hommes politiques élus, les dirigeants de l'entreprise publique. C'est l'analyse à laquelle se livre l'école des « choix sociaux ». Le point de départ de l'analyse est de remarquer qu'un des objectifs majeurs des élus est leur réélection. Dans ces conditions le comportement par les électeurs a-t-il des chances de les conduire à maximiser le bien-être social issu de l'entreprise publique en cause ? Les élections n'ont lieu qu'à intervalles distants, entre lesquels l'oubli s'installe ; elles portent sur de nombreux sujets, et la gestion de l'entreprise publique n'est que l'un

d'eux ; par ailleurs l'effet de l'amélioration du bien être social est inégalement réparti : pour l'atteindre, il faudra peut-être pénaliser durement les employés de l'entreprise, sans bénéfice électoral sensible auprès des usagers nombreux qui chacun en tireront un faible bénéfice. Il en résulte un mauvais contrôle de la qualité de la gestion.

Ces mécanismes ont abouti à des déficits lourds pour le budget de l'État. Depuis 30 ans, de nombreux efforts ont été faits pour pallier ces inconvénients. Les objectifs des multiples réformes qui ont été entreprises étaient déjà inscrits dans le premier des rapports effectués sur le sujet, le rapport Nora (1968) :

- clarifier la comptabilité des entreprises publiques en faisant apparaître le coût des obligations de service public, et en identifiant les subventions nécessaires à leur financement,
- séparer dans ces entreprises les activités soumises aux lois du marché et celles soumises aux contraintes de service public, et gérer chacune d'elles selon les principes correspondants,
- donner aux entreprises une autonomie de gestion en les débarrassant de la tutelle administrative.

L'ensemble de ces orientations se traduisait dans les contrats de Plan. Signés entre l'État et l'entreprise publique pour une durée de 4 à 5 ans, ils devaient canaliser et organiser les rapports entre l'État et l'entreprise publique en fixant les objectifs de l'entreprise (productivité, part de marché, résultat financier), ses obligations de service public, et les moyens que l'État lui accorderait dans ce but.

La procédure n'a qu'imparfaitement rempli ses objectifs pour ce qui est de la SNCF : le contrat n'avait ni véritable valeur juridique, ni sanction en cas de non-respect, et les partenaires étaient inégaux. L'État a plusieurs fois renié ses engagements lorsque des considérations conjoncturelles incitaient à revenir sur des clauses du contrat (par exemple en refusant des hausses de tarifs pour lutter contre l'inflation). Par ailleurs le contrat ne pouvait prévoir toutes les éventualités (il était incomplet) et devenait de facto caduc lorsque par exemple une dépression conjoncturelle rendait inatteignables les objectifs qu'il fixait. Enfin, il a certes canalisé les interventions de l'État, mais insuffisamment, et la SNCF se trouve grevée d'obligations de service public cachées, notamment pour les services de voyageurs de petites lignes nationales.

3.2. La concession

Le système de la concession est très utilisé dans les transports. On a vu au chapitre VII son usage dans le cadre des infrastructures. Leur objet est alors essentiellement le financement des investissements. On va en revanche s'attacher ici au cas des concessions d'exploitation.

La formule est particulièrement attirante lorsque le régulateur veut imposer au monopole des obligations de service public. Elle prend alors souvent la forme d'une concession à la suite d'un appel d'offre. C'est ce

système qui est utilisé dans de nombreux pays pour les transports collectifs urbains et régionaux. Et beaucoup de pays qui se sont engagés dans la voie d'une forte fragmentation des transports ferroviaires l'ont utilisé : Grande-Bretagne pour les transports de voyageurs interurbains, Suède pour les transports régionaux par exemple.

Les difficultés sont : risque de collusion entre les candidats, risque de corruption, lié notamment à la subjectivité dans le jugement des offres, incomplétude du contrat, difficultés du contrôle de son exécution, avec risques d'opportunisme de la part du concédant.

Le renouvellement de la concession est un point spécialement délicat, l'entreprise en place dispose d'une supériorité d'information (la « malédiction du vainqueur » si celui-ci est un nouvel opérateur qui n'a gagné que par erreur de calcul résultant d'une moins bonne information sur les conditions du service). Les investissements du concédant posent un problème particulier, en dehors du risque d'insuffisance par la crainte d'opportunisme de la part du régulateur. Il réside dans leur sort à l'achèvement de la concession ; la transmission au nouveau concessionnaire est en général la solution la plus efficace, et souvent la seule possible lorsque les investissements sont spécifiques. Mais l'incertitude sur leur valeur pèse sur le compétiteur éventuel et peut, si ces investissements pèsent beaucoup dans les coûts, le détourner de soumissionner. Il en est de même si les coûts de transmission des actifs correspondants sont substantiels. Ces considérations conduisent à allonger les durées des concessions si celles-ci comportent des investissements durables. Ainsi en France les concessions de transport urbains dans lesquelles le matériel est possédé par la commune ont des durées de quatre à cinq ans, alors que celles qui portent sur la construction et l'exploitation d'infrastructures durent plusieurs dizaines d'années.

Ce système de concession se traduit par des contrats qui abandonnent une plus ou moins grande rente informationnelle à l'opérateur, et qui comportent une plus ou moins forte incitation à l'effort, selon l'ampleur de l'asymétrie d'information entre le concessionnaire et le concédant.

L'asymétrie d'information peut se présenter ex-ante, car le régulateur connaît moins bien que l'opérateur les conditions du marché et les coûts. Alors un contrat à prix fixe devra laisser à l'opérateur une rente pour l'inciter à effectuer le service, selon un mécanisme qui peut être modélisé de manière similaire à celle qui est présentée au chapitre précédent (encadré 3-5) pour expliquer les conditions de différenciation des prix du monopole : une rente importante est abandonnée à l'opérateur efficace, de même qu'un surplus élevé est laissé au consommateur le plus important.

L'asymétrie d'information peut aussi se présenter ex-post, par impossibilité pour le régulateur d'observer l'effort effectué par l'opérateur, par exemple parce que le résultat dépend également d'un hasard. Alors le régulateur devra payer un coût d'autant plus élevé ou obtenir un résultat d'autant plus faible que l'entreprise est aversive au risque. Ce résultat se déduit du même raisonnement que celui développé au chapitre VII, encadré 7-13.

Le système de la concession est fréquemment utilisé dans les transports urbains et dans les ports, et donne l'occasion d'aborder ces sujets, en analysant successivement les transports urbains en France et en Grande-Bretagne et la gestion portuaire.

3.2.1. Les transports urbains en France

Costa (1996) fournit un panorama des différents modèles de gestion rencontrés en Europe dans l'organisation des transports publics urbains. Il aboutit à 4 modèles, caractérisés par l'intégration ou la non-intégration de l'autorité organisatrice et de l'opérateur, par l'existence d'un ou de plusieurs opérateurs, et par la liberté d'entrée sur le marché ; il représente les résultats dans le tableau suivant :

Modèles	Intégration régulateur-opérateur	Séparation régulateur-opérateur ; un seul opérateur	Séparation plusieurs opérateurs	Liberté d'entrée et de concurrence
Caractéristiques principales				
Différenciation des produits	Faible	Faible	Faible	Oui
Barrières à l'entrée	Oui	Oui	Moyenne	Non
Intégration fonctionnelle	Oui	Non	Non	Oui
Intégration verticale	Oui	Oui	Variante	Variante
Stratégie de gestion	Extension maximale du réseau	Extension maximale du réseau	Extension maximale du réseau	Couverture des coûts
Politique de prix	Objectifs divers	Taux de récupération des coûts satisfaisants	Taux de récupération des coûts satisfaisants	Profit
Efficacité allocative et productive	Efficacité liée aux rendements d'échelles	Efficacité liée aux rendements d'échelles	Efficacité productive	Efficacité productive et dynamique
Progrès technique	Lent	?	?	Rapide

Dans ce spectre, les agglomérations françaises se situent presque toutes dans le modèle 2, à base de contrats de concession.

On peut définir ces contrats en termes économétriques de la façon suivante : Étant donné :

- un prix p fixé par l'AO (en première approximation),
- les efforts de commercialisation développés par l'exploitant (marketing et communication), notés a ,
- des paramètres structurels de la demande θ_q

Le volume de trafic résultant est alors :

$$q = D(p, a, \theta_q)$$

et les recettes effectives sont aléatoires :

$$\tilde{r} = \bar{R}(pq)$$

De même, étant donnés :

- les efforts de productivité développés par l'exploitant notés e ,
- les paramètres de sa technologie θ_e ,
- la réalisation de volume d'activité q implique des coûts effectifs aléatoires :

$$\tilde{c}(q, e, \theta_e)$$

La mise en place des contrats peut se formaliser sous la forme du processus technique suivant, faisant référence à la littérature sur la théorie des contrats en information incomplète (voir Caillaud *et al.*, 1988) : d'abord, une phase de discussion et de proposition de la part de l'exploitant sert de phase de révélation, où les paramètres sont annoncés à l'AO. Celle-ci détermine ensuite le niveau d'effort qu'elle désire induire, soit : (a^0, e^0) ce qui détermine implicitement les objectifs :

- r^0 , objectifs de recette
- c^0 , objectifs de coût

L'AO détermine aussi le niveau espéré de rémunération qu'elle désire laisser à l'exploitant pour la réalisation exacte de ces objectifs : $U^0 = r^0 + c^0 - c^0$. Suivant alors le type de contrat retenu, la rémunération effective de l'exploitant est déterminée comme suit :

- *Risques et périls (RP)* : subvention $t = r^0$

L'exploitant supporte donc tout le risque (marginal), mais l'AO supporte le risque d'une subvention fixe correspondant au déficit structurel espéré du secteur de transport étant donné les objectifs définis. Le profit de l'exploitant est :

$$u = U^0 + (\tilde{r} - r^0) - (\tilde{c} - c^0)$$

- *Garantie de recette (GR)* : subvention $t = r^0 + \max(0, r^0 - \tilde{r})$.

L'exploitant supporte le risque industriel marginal et seulement les bons risques commerciaux. Son profit est : r^0 .

- *Prix forfaitaire (PF)* : subvention $t = r^0 + (r^0 - \tilde{r})$

L'exploitant ne supporte que le risque industriel, l'AO le risque commercial.

Le profit de l'exploitant est :

$$U = U^0 - (\tilde{c} - c^0)$$

- *Gérance (G)* : subvention $t = r^0 + (\tilde{r} - r^0) - (\tilde{c} - c^0)$.

L'exploitant ne supporte plus aucun risque et son profit est :

$$U = U^0$$

Dans tous les cas, c'est l'AO qui fixe le prix (les tarifs) du transport. En fait l'AO fixe aussi bien les prix que les conditions de l'offre qu'elle demande (qualité, fréquence, etc.). Mais l'exploitant est parfois invité à faire des suggestions pour que l'AO bénéficie de son expérience. De plus il conserve une certaine autonomie sur des tarifs spéciaux, représentant parfois un pourcentage important de son chiffre d'affaire.

On peut rassembler ces expressions dans une formulation unique :

$$t = T + A(\tilde{r} - r^0) - B(\tilde{c} - c^0)$$

avec comme terminologie caractérisant les types de conventions :

- gérance : $A = B = -1$
- risques et périls : $A = B = 0$
- prix forfaitaires : $A = -1, B = 0$
- recette garantie : $B = 0, A = -1$ si $r \leq r^0$
 $A = 0$ si $r > r^0$

L'étude économétrique des performances des systèmes de transport urbain français a été effectuée par plusieurs auteurs, en particulier Croissant (1996), Kerstens (1996), Caillaud et Quinet (1992).

Croissant s'intéresse à la manière dont les transports urbains des diffèrentes agglomérations satisfont à des objectifs d'optimum économique ; il évalue les conséquences possibles d'une déréglementation. Pour cela il estime des fonctions de coût qui dépendent des volumes de trafics, de la nature, en quantité et en qualité, de l'offre, et des caractéristiques de l'agglomération ; il évalue aussi les fonctions de demande, ce qui lui permet d'en tirer une expression du surplus collectif, qui dépend évidemment du système de tarification, et donc de subventions, choisis par chaque agglomération.

Il résulte de son étude que les subventionnements actuels devraient être augmentés dans une optique d'optimum de premier rang. Ensuite, avec le niveau actuel de subvention, une amélioration de la qualité de service accompagnée d'une augmentation des tarifs serait bénéfique. Une déréglementation, qui rapprocherait les tarifs des coûts moyens, réduirait le bien-être social, mais s'accompagnerait aussi d'une réduction sensible des coûts. Cette analyse suppose que l'efficacité des opérateurs n'est pas influencée par le mode de régulation : leur fonction de coût serait indépendante de leur statut et du contrat qui les lie à l'autorité organisatrice. Cette hypothèse est testée par Kerstens (1996), qui analyse les effets de la propriété et du type de contrat sur l'efficacité productive¹.

L'auteur trouve que la propriété du réseau a une légère influence sur l'efficacité productive, la propriété privée étant plus efficace. Mais le facteur le plus important est la forme du contrat, à travers trois caractéristiques : la longueur du contrat, qui réduit l'efficacité ; le montant de la subvention, qui réduit également l'efficacité ; le partage du risque prévu dans le contrat : plus une part importante du risque est supportée par l'opérateur, plus son efficacité est grande, mais aussi plus importante doit être la rente que le contrat lui laisse. Gagnepain (1998) fait apparaître le même effet en ce qui concerne l'asymétrie d'information : plus elle est élevée, plus le contrat à prix fixe laisse une rente élevée et est inadapte.

Caillaud et Quinet (1992) donnent un autre éclairage à ces contrats en analysant, non pas leurs effets, mais les raisons de leur choix, en cherchant

1. On détermine par une procédure économétrique la frontière de production, enveloppe des productions réalisées à partir des inputs par les opérateurs du réseau. Ensuite, l'efficacité de chaque opérateur est mesurée par sa distance à la frontière de production. Cette efficacité est corrélée avec les caractéristiques du réseau et du contrat.

à expliquer les valeurs des coefficients d'intéressement aux recettes et d'intéressement aux coûts. Il ressort de leur analyse économétrique que ceux-ci peuvent assez bien être expliqués par des considérations issues de la théorie des contrats.

Il semble que l'intéressement aux recettes soit dicté par des considérations d'asymétrie d'information. L'autorité organisatrice est en déficit informationnel important quant aux paramètres caractéristiques de la demande, et se voit donc contrainte d'assurer l'exploitant de manière plus importante vis-à-vis du risque commercial pour éviter d'avoir à lui laisser une rente informationnelle trop importante par l'intermédiaire des subventions fixes. La composante d'incitation à l'effort de commercialisation semble beaucoup plus réduite, quoique non complètement absente.

La détermination du coefficient d'intéressement aux coûts semble tenir en grande partie à un effet de trend, un effet de groupe, et une problématique de type incitation à la productivité. Les problèmes de sélection ne semblent par contre pas jouer de rôle fondamental.

3.2.2. Les transports urbains en Grande-Bretagne

La Grande-Bretagne a choisi, au moins en dehors de l'agglomération londonienne, une solution différente, fondée sur la concurrence classique par le marché, et mise en œuvre à partir du milieu des années 80 : les monopoles ont été détruits, les entreprises publiques privatisées, l'entrée sur le marché totalement libre.

Ce qui s'est passé est voisin de ce qu'on a constaté pour les transports par autocars : après une période de concurrence effective, des structures de monopoles se sont reconstituées. Elles résultent, selon les analyses, des fortes économies d'échelle attachées aux effets de maillage des réseaux de transport. De ce fait, les baisses de coûts obtenue durant la période transitoire se sont arrêtées. En outre, les gains d'efficacité se sont effectués assez largement au détriment des usagers : tarifs accrus, qualité de service réduite. Enfin, les opérateurs ont abandonné les lignes non rentables, que les autorités organisatrices ont dû subventionner lorsqu'elles voulaient les maintenir pour des raisons de service public. L'efficacité productive a été obtenue certes, mais partiellement seulement car l'aiguillon de la concurrence a vite disparu, et au détriment du service public, détérioré par la baisse des tarifs, une moins bonne information des usagers et une coordination des horaires réduite.

A Londres a été utilisé le système de concession après appel d'offre pour une bonne part des bus, les résultats ont été jugés très bons par la plupart des auteurs (Glaister, 1997 ; Mackie, 1997) : réduction des coûts, amélioration de la qualité de service. Ces effets bénéfiques se sont en outre propagés, par une sorte de concurrence par comparaison, aux lignes non concédées, dont la gestion a également été améliorée. Il reste à voir si cette efficacité perdurera avec le renouvellement des concessions ; certes, dans une concurrence pour le marché s'exerce ; mais dans cette compétition l'opérateur déjà en place a un avantage considérable, tenant à sa

sur le contrôle de l'entreprise, assuré non plus par la puissance publique, mais par les actionnaires.

L'hypothèse la plus simple est que ceux-ci sont mus par l'objectif du profit maximum. Mais cet objectif général peut avoir différentes traductions concrètes, par exemple profits à court terme ou à long terme. En outre, le contrôle des actionnaires peut être lâche, et ce pour plusieurs raisons : lorsque l'actionnariat est dispersé, le contrôle d'un actionnaire engendre des externalités positives dont il ne profite pas, et a donc des chances d'être sous optimal ; ce contrôle s'exerce en situation d'asymétrie d'information.

Néanmoins, quelques raisons permettent de lui attribuer une certaine efficacité : l'existence d'un niveau intermédiaire de contrôle entre les actionnaires et l'entreprise, constitué par les dirigeants dont les rémunérations élevées sont liées au profit de l'entreprise, le fait que le contrôle exercé sur une firme peut par comparaison influencer sur le comportement des autres firmes.

Enfin, le risque d'OPA est souvent vu comme une incitation au profit pour l'entreprise et ses managers : un profit faible ferait baisser la valeur des actions et inciterait un acheteur à les reprendre à bas prix pour améliorer la gestion et, faisant monter le cours de l'action, enregistrer des profits. Mais si la perspective de profit après rachat existe, pourquoi les actionnaires anciens vendraient-ils leurs actions ? Cet argument a lui-même ses contre-arguments : les propriétaires de portefeuilles diversifiés peuvent vouloir se faire donner une réputation de vendeur, et ainsi inciter au profit les autres entreprises dont il sont actionnaires ; ceux qui restent minoritaires après le rachat seront vulnérables vis-à-vis du majoritaire. Par ailleurs, on peut montrer que, pour certaines valeurs des paramètres, la menace d'OPA incite les dirigeants, non pas à l'effort, mais à l'abandon, si l'effort est difficile à exercer, et si son résultat est faible et aléatoire et ne peut guère modifier la probabilité de rachat.

Une autre incitation à l'effort vient du risque de faillite. Cette incitation connaît deux limitations ; la première est analogue à l'argument précédent : si le risque de faillite est élevé et peu sensible à l'effort, alors autant réduire l'effort, dont la probabilité de succès est faible. L'autre limite vient de ce que le risque de faillite dépend en fait du niveau de la dette de l'entreprise, qui est contrôlée par le Conseil d'Administration et les dirigeants, qui maîtrisent ainsi pour une part l'outil de contrainte s'exerçant sur eux.

Ces considérations théoriques sont confirmées par les études statistiques¹ : les entreprises privées sont dans l'ensemble plutôt plus efficaces que les entreprises publiques, la différence étant d'autant plus forte que

les entreprises opèrent sur un marché compétitif. Bien évidemment leur efficacité externe, c'est-à-dire la mesure dans laquelle elles atteignent l'objectif de maximisation du bien être social, dépend de l'ampleur des externalités.

On peut illustrer ces considérations générales par quelques exemples qui montrent comment elles jouent dans la réalité. La plupart de ces exemples sont d'origine britannique, car c'est la Grande-Bretagne qui a le plus utilisé la procédure dans les années récentes.

a) La privatisation de British Airport Authority eu lieu en juillet 87. Les traits de cette privatisation et des mesures qui l'accompagnaient sont d'abord relatifs à la concurrence à laquelle est soumise BAA. En fait BAA regroupe les aéroports londoniens et les aéroports écossais, c'est-à-dire l'écrasante majorité du trafic aérien britannique. La concurrence internationale existe, mais Londres y a une position dominante. Le contrôle des prix s'est effectué selon une procédure de price-cap, ou (RPI-X). Mais ce contrôle n'a pas remédié à une inefficacité majeure du système de prix antérieur. À savoir la sous-tarification du trafic au regard des activités commerciales. Pour celles-ci, l'attribution des concessions s'est poursuivie en évitant la concurrence, ce qui augmente les profits des concessionnaires, donc les rentes qu'on peut en extraire. Quant à la politique d'investissements majeurs furent réalisés avant la privatisation. Au total la privatisation de BAA semble surtout avoir eu des conséquences financières sur le partage du profit de l'entreprise dont le comportement est celui d'un monopole animé par la recherche du profit et dont les contraintes semblent relativement limitées.

b) British Airways a été dénationalisée en 1987, alors que la situation financière mauvaise en 1982 (perte de 108 G £) avait été rétablie (gain de 185 G £ en 1984 ; de 162 en 87). Cette privatisation s'opérait dans le cadre d'une politique affichée d'inspiration libérale et destinée à développer la concurrence. Les recommandations émises à l'occasion de cette privatisation étaient que certaines liaisons internationales soient transférées à British Caledonian, que toutes les autres puissent être transférées si besoin à d'autres opérateurs, que certaines liaisons nationales soient transférées à d'autres opérateurs. En fait, cette politique destinée à promouvoir la concurrence n'a pas été suivie. BA a gardé ses lignes, a bénéficié d'une position dominante sur Heathrow, le plus grand aéroport londonien ; et finalement BA absorba la compagnie British Caledonian.

c) La privatisation de Railtrack est d'un genre très particulier. D'abord, s'il existe monopole naturel c'est bien celui-là, et on pourrait s'attendre à ce qu'une gestion privée conduise à une perte économique élevée. Mais le mécanisme financier est complexe ; en principe Railtrack tarifie en vue de couvrir ses dépenses et de faire un profit (lorsque Railtrack était public, la tarification avait été fixée de façon à rémunérer le capital qui avait été largement évalué ; mais la privatisation n'a rapporté qu'une somme très inférieure aux évaluations du capital¹ en raison d'un prix d'émission faible

1. « Progress with rail privatisation in Great Britain » N. Montagu, World Railway Update 1996-97.

2. Ou une diminution de la redevance que les concessionnaires lui versent.

meilleure information sur les conditions d'exploitation, tenant aussi à ce que son maintien supprime les coûts liés au changement d'opérateur, tenant enfin aux relations qu'il a pu établir avec l'autorité organisatrice.

3.2.3. La gestion portuaire

Les ports maritimes sont fréquemment le siège de rendements croissants, qui trouvent leur origine dans la présence de fortes dépenses fixes d'infrastructures (digues, quais, ouvrages de protection à la mer), et de superstructures (grues, portiques de déchargement et manutention des conteneurs), ainsi que dans les conditions d'exploitation : un grand port accueille des navires de plus grande taille engendrant des économies ; le coût du remorquage et le pilotage augmentent moins que proportionnellement ; pour décharger un conteneur, il faut en déplacer d'autres mais ces derniers sont en nombre d'autant plus réduit que le nombre de conteneurs à décharger est grand, c'est-à-dire que la taille du port est plus grande. Un phénomène analogue se produit pour l'acheminement terrestre en ce qui concerne l'acheminement et le triage des wagons et la possibilité de constituer des trains complets, et pour l'acheminement maritime avec la possibilité pour un grand port d'accueillir des navires plus gros bénéficiant de rendements d'échelle. Les limites aux rendements d'échelle sont les limites du site naturel (largeur, profondeur des postes à quai, taille de l'aire protégée...).

Ces considérations permettent d'abord de tirer quelques conclusions en ce qui concerne la structure du réseau portuaire. Ainsi, pour une collectivité isolée – par exemple une île – la structure portuaire devrait, comme d'ailleurs la structure aéroportuaire, être assez concentrée avec un arbitrage entre les économies d'échelle (dans le port, sur la chaîne maritime et sur la chaîne terrestre) permises par la concentration, et les allongements de parcours de transport terrestre que cette concentration implique. La problématique est tout à fait similaire à celle des réseaux étoilés des autres modes de transport. D'ailleurs une structure portuaire à deux niveaux, où la cargaison des grands navires est éclatée dans un port principal entre plusieurs navires de taille plus réduite pour être finalement déchargée dans des ports secondaires, est de plus en plus fréquente.

L'existence de rendements croissants dans les ports ainsi que dans les deux maillons de la chaîne de transport dont ils sont la jonction, la présence d'effets de seuil (seuil de profondeur pour l'accueil des grands navires, seuil de massification pour les conteneurs), ainsi que leur importance pour le commerce extérieur plaident pour que le système portuaire ne soit pas abandonné à la concurrence, mais fasse l'objet d'un contrôle public s'exerçant par exemple sur le niveau des droits de port perçu (régulation des prix d'un monopole) et l'attribution de subventions incitatives (régulation des investissements), pour lesquelles l'État permet aux différents sports d'intérioriser les effets externes qu'ils engendrent. Ces effets externes ne concernent d'ailleurs pas seulement la chaîne de transport, mais aussi les activités économiques qui se localisent et développent en leur voisinage.

On a jusqu'ici raisonné sur le cas d'une collectivité isolée, une île. Une situation fréquente est celle où tout l'hinterland d'une même façade maritime est desservi par plusieurs ports de pays différents : Hambourg, Rotterdam, Anvers, le Havre sur l'Atlantique ; Marseille, Barcelone en Méditerranée. On trouve alors dans une situation analogue à celle décrite au chapitre VII, § 3.1.8.d), où les politiques nationales diffèrent en général de la politique optimale commune, et se traduisent en général par des surinvestissements. Le recours à la concurrence entre ports, facteur d'efficacité, connaît donc des limites. Mais les situations et les statuts sont très divers d'un pays à l'autre ou d'une période à l'autre. Le Royaume-Uni est passé en 1981 d'une politique marquée par une coordination par un échelon central à une libéralisation conduisant à la privatisation des plus grands ports. Les effets de cette réforme ont été assez semblables à ceux constatés dans les autres modes de transports (Goss, 1998) : des regroupements financiers amenant à la constitution d'oligopoles, une concurrence accrue, mais moins qu'espéré, des réductions de coûts dues peut-être plus à la modification du statut des dockers qu'à la privatisation proprement dite.

Le recours à la concurrence est plus fréquent en matière infra-portuaire ; elle est susceptible de s'exercer pour la fourniture des multiples services présents à l'intérieur de chaque port (Commission des Communautés Européennes 1998). En ce qui concerne les services au navire, le remorquage, le pilotage et le lamanage mettent fortement en jeu la sécurité, et doivent donc faire l'objet d'un contrôle public étroit. Dans la plupart des États de l'Union Européenne, ils sont actuellement assurés soit par la fonction publique, soit par des associations ou coopératives étroitement contrôlées ; mais on pourrait imaginer, au moins dans les grands ports, des structures plus incitatives à l'efficacité. En revanche pour les services à la marchandise (manutention, transit, passage en douane, auxiliaires de transports divers, stockage et logistique), la mise en concurrence sous ses multiples formes (concurrence par le marché, concession) est une pratique fréquente, gage d'efficacité.

3.3. La propriété privée

La propriété privée est souvent jugée :

- supérieure à la propriété publique pour ce qui est de l'efficacité interne, c'est-à-dire la minimisation des coûts,
- inférieure à la propriété publique pour ce qui est de l'efficacité externe, c'est-à-dire la poursuite de l'intérêt collectif, surtout lorsqu'il s'accompagne d'obligations de service public.

Penchons-nous spécialement sur le problème de l'efficacité interne. De ce point de vue, la caractéristique essentielle de la propriété privée porte

1. Par exemple Gathon et Pestieau (1995), Kersten (1996) et Gomez-Ibanez et Meyer (1993).

sur le contrôle de l'entreprise, assuré non plus par la puissance publique, mais par les actionnaires.

L'hypothèse la plus simple est que ceux-ci sont mus par l'objectif du profit maximum. Mais cet objectif général peut avoir différentes traductions concrètes, par exemple profits à court terme ou à long terme. En outre, le contrôle des actionnaires peut être lâche, et ce pour plusieurs raisons : lorsque l'actionnariat est dispersé, le contrôle d'un actionnaire engendre des externalités positives dont il ne profite pas, et a donc des chances d'être sous optimal ; ce contrôle s'exerce en situation d'asymétrie d'information.

Néanmoins, quelques raisons permettent de lui attribuer une certaine efficacité : l'existence d'un niveau intermédiaire de contrôle entre les actionnaires et l'entreprise, constitué par les dirigeants dont les rémunérations élevées sont liées au profit de l'entreprise, le fait que le contrôle exercé sur une firme peut par comparaison influencer sur le comportement des autres firmes.

Enfin, le risque d'OPA est souvent vu comme une incitation au profit pour l'entreprise et ses managers : un profit faible ferait baisser la valeur des actions et inciterait un acheteur à les reprendre à bas prix pour améliorer la gestion et, faisant monter le cours de l'action, enregistrer des profits. Mais si la perspective de profit après rachat existe, pourquoi les actionnaires anciens vendraient-ils leurs actions ? Cet argument a lui-même ses contre-arguments : les propriétaires de portefeuilles diversifiés peuvent vouloir se faire donner une réputation de vendeur, et ainsi inciter au profit les autres entreprises dont il sont actionnaires ; ceux qui restent minoritaires après le rachat seront vulnérables vis-à-vis du majoritaire. Par ailleurs, on peut montrer que, pour certaines valeurs des paramètres, la menace d'OPA incite les dirigeants, non pas à l'effort, mais à l'abandon, si l'effort est difficile à exercer, et si son résultat est faible et aléatoire et ne peut guère modifier la probabilité de rachat.

Une autre incitation à l'effort vient du risque de faillite. Cette incitation connaît deux limitations ; la première est analogue à l'argument précédent : si le risque de faillite est élevé et peu sensible à l'effort, alors autant réduire l'effort, dont la probabilité de succès est faible. L'autre limite vient de ce que le risque de faillite dépend en fait du niveau de la dette de l'entreprise, qui est contrôlée par le Conseil d'Administration et les dirigeants, qui maîtrisent ainsi pour une part l'outil de contrainte s'exerçant sur eux.

Ces considérations théoriques sont confirmées par les études statistiques¹ : les entreprises privées sont dans l'ensemble plutôt plus efficaces que les entreprises publiques, la différence étant d'autant plus forte que

les entreprises opèrent sur un marché compétitif. Bien évidemment leur efficacité externe, c'est-à-dire la mesure dans laquelle elles atteignent l'objectif de maximisation du bien être social, dépend de l'ampleur des externalités.

On peut illustrer ces considérations générales par quelques exemples qui montrent comment elles jouent dans la réalité. La plupart de ces exemples sont d'origine britannique, car c'est la Grande-Bretagne qui a le plus utilisé la procédure dans les années récentes.

a) La privatisation de British Airport Authority eu lieu en juillet 87. Les traits de cette privatisation et des mesures qui l'accompagnaient sont d'abord relatifs à la concurrence à laquelle est soumise BAA. En fait BAA regroupe les aéroports londoniens et les aéroports écossais, c'est-à-dire l'écrasante majorité du trafic aérien britannique. La concurrence internationale existe, mais Londres y a une position dominante. Le contrôle des prix s'est effectué selon une procédure de price-cap, ou (RPI-X). Mais ce contrôle n'a pas remédié à une inefficacité majeure du système de prix intérieur, à savoir la sous-tarification du trafic au regard des activités commerciales. Pour celles-ci, l'attribution des concessions s'est poursuivie en présence de la concurrence, ce qui augmente les profits des concessionnaires, donc les rentes qu'on peut en extraire. Quant à la politique d'investissement de l'entreprise privatisée, il est difficile d'en juger, car les investissements majeurs furent réalisés avant la privatisation. Au total la privatisation de BAA semble surtout avoir eu des conséquences financières sur le partage du profit de l'entreprise dont le comportement est celui d'un monopole animé par la recherche du profit et dont les contraintes semblent relativement limitées.

b) British Airways a été dénationalisée en 1987, alors que la situation financière mauvaise en 1982 (perte de 108 G £) avait été rétablie (gain de 125 G £ en 1984 ; de 162 en 87). Cette privatisation s'opérait dans le cadre d'une politique affichée d'inspiration libérale et destinée à développer la concurrence. Les recommandations émises à l'occasion de cette privatisation étaient que certaines liaisons internationales soient transférées à British Caledonian, que toutes les autres puissent être transférées si besoin à d'autres opérateurs, que certaines liaisons nationales soient transférées à d'autres opérateurs. En fait, cette politique destinée à promouvoir la concurrence n'a pas été suivie. BA a gardé ses lignes, a bénéficié d'une position dominante sur Heathrow, le plus grand aéroport londonien ; et finalement BA absorba la compagnie British Caledonian.

c) La privatisation de Railtrack est d'un genre très particulier. D'abord, il s'agit d'un monopole naturel c'est bien celui-là, et on pourrait s'attendre à ce que la gestion privée conduise à une perte économique élevée. Mais le caractère financier est complexe ; en principe Railtrack tarifie en vue de couvrir ses dépenses et de faire un profit (lorsque Railtrack était public, la gestion avait été fixée de façon à rémunérer le capital qui avait été largement évalué ; mais la privatisation n'a rapporté qu'une somme très inférieure aux évaluations du capital¹ en raison d'un prix d'émission faible

1. « Progress with rail privatisation in Great Britain » N. Montagu, World Railway Update 1996-97.

2. Ou une diminution de la redevance que les concessionnaires lui versent.

considéré comme un succès de la privatisation par l'opinion publique, ce résultat constitue en fait un échec pour les finances publiques et a un effet de répartition massif et passé sous silence). Mais les principaux clients de Railtrack sont les concessionnaires de lignes voyageurs ; et par une subvention accrue à ces concessionnaires², l'Etat en fait subventionne Railtrack.

d) L'évolution d'Eurotunnel, déjà présentée, peut faire douter des vertus de la privatisation. Mais on a vu précédemment que les difficultés de l'opération tiennent essentiellement aux conditions de lancement de l'opération. Eurotunnel semble maintenant avoir mis à jour tous les pas-sifs résultant de ce début, et il sera intéressant de voir maintenant si les caractéristiques de la gestion privée se font jour conformément aux prévisions théoriques.

e) L'étude de synthèse la plus complète et la plus approfondie sur l'effet des privatisations dans le secteur des transports à travers le monde, est celle de Gomez-Ibanez et Meyer (1993). Les auteurs aboutissent à 5 conditions pour le succès d'une privatisation : la présence d'une compétition, un gisement de gains d'efficacité possibles, de faibles transferts, des effets externes peu importants, et une rentabilité financière suffisante, mais pas trop forte. D'une manière générale la privatisation des services est plus aisée que celles des infrastructures.

4. Conclusion

Notre connaissance de la gestion des monopoles et des conditions d'exercice du service public s'est considérablement enrichie à la suite des expériences de libéralisation qui se sont produites un peu partout, et des réflexions qu'elles ont suscitées. Mais les solutions et les problèmes qu'elles visent à résoudre sont loin d'être nouvelles.

Sans remonter à des périodes trop reculées, le 19^e siècle et le début du 20^e siècle ont connu des alternances, différenciées selon les pays, entre l'intervention publique et l'initiative privée. Les chemins de fer français se sont construits sous la forme de monopoles privés, et c'est à la suite de leurs difficultés financières que l'Etat est intervenu de plus en plus lourdement, jusqu'à en prendre la gestion directe à la suite de la deuxième guerre mondiale. Les transports urbains et régionaux ont connu des évolutions plus différenciées, mais là aussi on a assisté à des alternances d'exploitation en régie et de concessions.

Actuellement, la tendance est à une réduction de l'intervention publique et du champ du monopole et à un retour à la gestion privée et au jeu des forces du marché. Ce mouvement est tout à fait net à l'étranger, notamment dans les pays anglo-saxons et de l'Europe du Nord, où il a atteint une sorte de point d'achèvement. Il est également très perceptible en France, même si nous avons un retard sur nos voisins sur certains points tels que la libéralisation des chemins de fer, ou la séparation entre l'autorité politique et le régulateur.

Ce changement de doctrine par rapport à ce qui prévalait il y a une cinquantaine d'années est certainement dû à une connaissance améliorée des mécanismes économiques. Il est aussi entraîné par les idéologies libérales qui prévalent actuellement, et qui ont tendance à amplifier et à renforcer en doctrine les acquis de la science. Mais on ne peut pas s'empêcher de penser que cette évolution est également le résultat des transformations qu'a connues la société, et d'abord de la perte de pouvoir de l'Etat. Le pouvoir en matière publique s'est émietté ; il se partage de plus en plus entre l'Etat, les collectivités locales et les institutions supra-nationales, notamment européennes. Corrélativement, l'influence des acteurs politiques dans les décisions s'est accrue, et a fait apparaître les failles des processus démocratiques qui nous régissent, en particulier le poids des pressions électorales, la rigidité qu'imprime leur rythme, les aspects démagogiques des luttes auxquelles ils conduisent, et les défauts du contrôle exercé sur les administrations par le système politique. Parallèlement, les décisions se prennent selon des processus de plus en plus complexes impliquant des intervenants en nombre élevé, et prennent davantage la forme de négociations et de recherche de consensus que celle de décisions imposées d'en haut.

Enfin l'initiative économique a changé de main. En France, entre les années 30 et 50, elle venait du secteur public. C'est de lui qu'est parti, après la deuxième guerre mondiale, le mouvement qui a abouti à la modernisation de l'appareil productif français, à travers, par exemple, les missions de productivité qui permettaient à des industriels français de s'initier aux méthodes de fabrication et de gestion en vigueur aux USA ; c'est le secteur public qui a pris en charge le contrôle et la dynamisation de notre économie à travers le Commissariat général du Plan. C'était l'époque où les grandes structures faisaient merveille et où leur gestion publique permettait d'assurer des orientations favorables à l'exécution du plan national. Mais cette époque est révolue, et les organisations lourdes et hiérarchiques qui en subsistent ne correspondent plus aux techniques et à l'état d'esprit d'un secteur productif moderne, obligé de s'adapter à des situations changeantes et de faire face aux incertitudes liées à la mondialisation, à la concurrence accrue, et aux aléas d'une croissance ralentie.

Parallèlement, la notion de service public, et notamment celle de service public de transport, s'est transformée avec l'élévation du niveau de vie, la généralisation de l'automobile et la diversification de l'offre de transport. C'est de moins en moins une affaire de masse. Les exclus de la mobilité sont des catégories ciblées, pas forcément les moins aisées. On y rencontre pêle-mêle les personnes âgées, les ruraux à faible revenu, les handicapés, les chômeurs, les illettrés... Paradoxalement, à l'époque où est née la notion de service universel, le service public est devenu moins universel. Sa satisfaction devrait mettre en œuvre des moyens plus diversifiés et plus subtils, et les grands organismes traditionnels de service public ne permettent que difficilement l'adaptation qui serait nécessaire.

Alors que les objectifs d'efficacité externe perdent de leur spécificité et de leur prégnance, l'efficacité interne prend une importance croissante, avec la nécessité de soulager les budgets publics et de réduire les frais généraux de la Nation. L'analyse économique a permis de dégager quelques remèdes en ce sens. Mais ces remèdes s'appliquent à une situation qui comporte de multiples facettes, et qui est dominée par de fortes contraintes politiques et sociales. Ces contraintes ont été surmontées dans de nombreux pays, souvent à travers des changements rapides. La voie choisie en France est implicitement celle d'une pression sur les coûts et les structures, pas trop forte pour ne pas créer de conflit, mais suffisamment continue et permanente pour aboutir à la longue à des résultats. Cette voie est plus douce, mais plus lente. Elle traduit un arbitrage entre efficacité externe et interne plus favorable à la première. La question se pose de savoir si cette position est tenable dans une Europe où les échanges internationaux se développent et où les politiques économiques des pays tendent à s'unifier.

BIBLIOGRAPHIE

- D. Banister, J. Berechman et G. de Rus (1992), « Competitive regimes within the european bus industry : theory and practice », *Transportation Research*, part A, vol. 26, n° 2.
- P. Bauchet (1991), *Le transport international dans l'économie mondiale*, Paris, Economica.
- Baumol, Panzar et Willig (1982), *Contestable markets and the theory of industrial structure*, Harcourt Brace Jovanovitch, San Diego.
- B. Bradshaw (1996), « The privatization of railways in Britain », *Japan Railway and Transport Review*, September.
- C.J. Brewer and C.R. Plott (1996), « A binary conflict ascending price auction (BICAP) mechanism for the decentralized allocation of the right to use railroad tracks », *International Journal of Industrial Organization*, 14.
- B. Caillaud et E. Quinet (1993a), « Analyse du caractère incitatif des contrats de transport urbain », document CEPREMAP, n° 9307.
- B. Caillaud et E. Quinet (1993b), « Faut-il pratiquer des tarifs dégressifs pour l'usage de l'infrastructure ferroviaire », Rapport pour la SNCF, CERAS.
- A. Costa (1996), « The organization of urban public transport systems in western European metropolitan areas », *Transportation Research*, part A, vol. 30, n° 5.
- Y. Croissant (1996), « Les performances des firmes françaises de transport urbain », *Revue d'Économie Politique*, mai-juin.
- N. Curien et M. Gensollen (1992), *L'économie des télécommunications : ouverture et réglementation*, Economica-ENSPTT, Paris.
- M. Debene et O. Raymondie (1996), « Sur le service universel : renouvellement du service public ou nouvelle mystification », *Actualité juridique - Droit Administratif*, 20 mars.

- R. Denoux de Saint-Marc (1996), *Rapport à Monsieur le Premier Ministre sur la service public*. La Documentation française, Paris.
- H.J. Gathon et P. Prestieau (1995), « La performance des entreprises publiques : une question de propriété ou de concurrence ? », CREPP 95/06, Université de Liège.
- J. Gomez-Ibanez et J. Meyer (1993), *Going private : the international experience with transport privatization*, Brooking, Washington.
- R. Goss (1988), « British Ports Policies since 1945 », *Journal of Transport Economics and Policy*, January.
- P. Häfner (1996), « The effects of Railroad reform in Germany », *Japan Railway and Transport Review*, September.
- D. Hahn et D. Thompson (1991) « Privatized transport infrastructure and incentives to invest », *Journal of Transport Economics and Policy*, September.
- C. Henry et E. Quinet (1996), « Service public, efficacité et concurrence dans le système ferroviaire français », *Transports*, juillet-août.
- C. Henry (1997), *Cours d'économie publique*, École Polytechnique.
- B. Jönke (1996), « L'expérience suédoise », Table Ronde 103, CEMT.
- A. Joffe-Dreyfus et M. Mumizaga (1992), « The effects of network density on European Railway costs », Communication à WCTR, Lyon.
- A. Kuriyama (1996), « Technical efficiency measurement and exploration of private urban transit companies », *Transportation Research A*, vol. 30, n° 6.
- J. Kogan et D. Thompson (1994), « Reshaping Argentina's Railways », *Japan Railway and Transport Review*, June.
- I. Kugel (1997), *The new structure of Argentine Railways*, World Railways, Keesington Publications.
- L.L. Laferrière et J. Tirole (1986), « Une théorie normative des contrats État-entreprise », *Annales d'Économie et de Statistiques*, n° 1.
- L.L. Laferrière et J. Tirole (1993), *A theory of incentives in regulation and procurement*, MIT Press.
- L.L. Laferrière et J. Tirole (1994), *Libéralisation et charge d'accès*, W.P. Idei.
- A. Lindbeck (1996), « Restructuring of the Swedish State Railways », *Japan Railway and Transport Review*, September.
- C. Nakanishi (1996), « Privatization and beyond ; the J.R. case », *Japan Railway and Transport Review*, September.
- Z. Maza (1997), « Réglementation ou concurrence », communication au 14^e colloquium de la CEMT, Innsbruck.
- E. Muraoka et K. Nakamura (1997), « Privatization of the Japan National Railways : a review of performance changes », *International Journal of Transport Economics*, February.
- S. Mutsaers et C. Winston (1997) « Empirical implications and tests of the contestability hypothesis », in Oum, Dodgson et al., *Transport Economics*, Harwood Academic Publishers.
- C. NEST (1997), « Séparation de l'infrastructure et de l'exploitation des chemins de fer : l'expérience britannique », Table ronde 103, CEMT.
- A. Murrells (1997), *Progress with rail privatization in Great Britain*, World Railways Update.
- NERA (1995), *An examination of rail infrastructure charges*, Rapport pour la Commission de l'Union Européenne, Londres.

- J.E. Nilsson (1995), « Allocation of track capacity », Working paper 1995-1, CTS.
- Y. Okano (1994), « The backdrop to privatization in Japan », *Japan Railway and Transport Review*, June.
- T.M. Oum et C. Yu (1994), « Economic Efficiency of railways and implications for Public Policy » *Journal of Transport Economics and Policy*, May.
- T. Oum, A. Zhang et Y. Zhang (1996), « A note on optimal airport pricing in a hub and spoke system », *Transportation Research*, part B, vol. 30, n° 1.
- J. Preston (1994), « The economics of rail privatization », Communication au séminaire commun, École Polytechnique, ENPC, 14 octobre.
- A. Rémy (1996), « Performances économiques des principaux opérateurs ferroviaires de l'O.C.D.E. », SES Ministère de l'Équipement, Paris.
- T. Suga (1997), « La cas du Japon », Table ronde 103, CEMTI.
- J. Vickers et G. Yarrow (1988), *Privatization : an economic analysis*, the MIT Press.
- B. Winston et C. Kuranami (1997), *Effects of Rail privatization in Japan*, World Railways, Kensington Publications.

Chapitre X

De l'économie à la politique des transports

On a vu dans les chapitres précédents l'ampleur des connaissances accumulées en économie des transports. Presque tous les thèmes de la théorie y sont mis en œuvre. Certains ont connu des approfondissements importants grâce à leur application à ce secteur ; on pense ici à la théorie des choix publics, à travers le surplus de Dupuit, à la valeur du temps, à la théorie des choix discrets, à certains aspects de la théorie de la production et des fonctions de coût, à l'analyse des effets de réseaux. C'est dans les transports que les études appliquées sont les plus fréquentes, soutenues par un appareil statistique considérable, comme on l'a vu pour les trafics, les coûts et les choix d'infrastructures.

Mais on a vu aussi les sujets d'insatisfaction que suscite la conduite du secteur pour l'économiste : des outils de gestion peu ou mal utilisés, comme l'action sur les prix ; des préconisations peu suivies, comme c'est le cas pour les choix d'investissement et la tarification infrastructures ; des coûts élevés, par exemple ceux qui résultent de la production de certains services publics ; et des préoccupations pour l'opinion publique en matière d'encombrements et d'atteintes à l'environnement.

Ces constatations conduisent à s'interroger d'abord sur l'écart entre recommandations de l'économie et pratiques, ensuite sur la possibilité de mettre en œuvre une politique cohérente des transports.

1. Réduire l'écart entre les recommandations et la pratique

Pour réduire cet écart, il faut d'abord en analyser les raisons qui, semble-t-il, peuvent être classées en plusieurs catégories.

La première tient à la nature du processus décisionnel qui est, dans les transports plus que dans tout autre secteur, marqué par la multiplicité des décideurs. Du côté des décideurs publics, on est passé en quelques décennies d'une situation où l'État était tout puissant, à une situation où il partage de plus en plus son pouvoir avec les institutions supra nationales et

avec les collectivités locales. Du côté des décideurs privés, toutes les entreprises, et pas seulement les entreprises du secteur transport stricto-sensu, sont opérateurs ; et les individus eux-mêmes sont non seulement consommateurs, mais aussi producteurs de transport privé à travers l'usage de l'automobile.

On est loin de la situation où un despote éclairé maîtriserait les commandes et pourrait imposer une situation optimale. Si l'économiste met au point une politique optimale, il ne suffit pas qu'il en convainque une personne, ou un groupe limité de décideurs comme c'est le cas lorsqu'il s'agit de jouer sur les taux d'intérêt en termes de politique macro-économique, ou sur les prix d'intervention dans le cadre de la politique agricole. L'économiste des transports devrait atteindre une myriade de décideurs aux intérêts divergents (voir par exemple Bernat et Olliver-Trigalo, 1997 ; Fourniau, 1994 pour une analyse des processus de décisions et des jeux d'intérêt).

À défaut de convaincre du bien fondé d'une politique, on pourrait penser à mettre au point des mécanismes automatiques permettant d'en assurer la réalisation. Pour l'économiste, le marché et le marchandage sont des mécanismes qui, sous certaines conditions, permettent justement de réaliser la conciliation des intérêts divergents, de faire concourir les égoïsmes à la réalisation d'un optimum collectif.

Il se trouve malheureusement que les conditions de bon fonctionnement de ces mécanismes ne se rencontrent guère dans les transports, en raison de la place qu'y prennent les externalités et les rendements croissants.

La décentralisation des décisions devient alors beaucoup plus complexe. Tout ce qui concerne les externalités est bien plus difficile à appréhender que ce qui concerne les biens marchands, comme on l'a vu au chapitre V, relatif aux coûts de transport, ou au chapitre VII, à l'occasion de la tarification et du choix des investissements ; et la même réflexion peut être faite relativement à la difficulté de traiter les situations de production à rendement croissant, avec l'arbitraire du partage des frais fixes notamment.

En outre le secteur est, plus qu'aucun autre, le siège de la diversité et de la multiplicité (Winston, 1988) : dans les transports, toutes les situations sont particulières ; il n'y a pas d'usager moyen ; il n'y a pas d'entreprise représentative ou de coût standard. Le secteur est aussi le lieu du mélange : la plupart des infrastructures supportent des trafics dont les motifs de déplacements et les distances de transport sont très diverses. Du coup, les règlements généraux sont forcément aveugles et maladroits, ils n'atteignent pas leur but, contraignant trop ou trop peu, mais rarement juste comme il convient. Les taxes uniformes souffrent du même défaut, comme on l'a vu à l'occasion de l'examen des tarifications optimales, au chapitre VII.

L'inadéquation des instruments économiques est accentuée par notre insuffisante connaissance des phénomènes qu'ils déclenchent. On a vu, en analysant les interrelations entre les transports et le développement éco-

nomique, notre faible capacité à prédire le futur lointain, ou les conséquences éloignées géographiquement, qui pourtant pèsent d'un poids important. On sait aussi que, si nous disposons de méthodes pour calculer le surplus global issu d'un investissement, nous ne savons pas dire grand-chose sur la manière dont ce surplus se répartit entre les membres de la collectivité.

Les éléments de diagnostic qui viennent d'être posés peuvent permettre de prendre une vue différente du rôle possible de l'économiste et des modalités de son action. Jusqu'ici l'économiste des transports a centré son intervention sur la recherche de l'optimum collectif. On connaît ses limites de mise en œuvre : incertitudes des calculs, possibilité de manipulations stratégiques ; caractère peu compréhensible et peu convainquant des résultats. Il faut ajouter que le calcul économique ne permet pas à lui seul d'éclairer les choix stratégiques telles que l'équilibre entre efficacité et service public, le choix du style de vie, le poids donné aux préoccupations de développement durable. De ce point de vue, le calcul économique doit être au service des options politiques et non l'inverse.

Mais surtout la référence à l'optimum collectif, indispensable pour rechercher les décisions optimales, n'est plus suffisante quand il s'agit de les concrétiser. Cette concrétisation devrait s'accompagner d'études plus détaillées, analysant les effets de répartition de la mesure envisagée, les moyens éventuels de compenser les agents lésés, ou les mesures d'accompagnement permettant d'en assurer l'acceptabilité sociale¹.

Pour cela nous devons d'abord combler les lacunes de notre connaissance, et deux champs de recherche semblent spécialement importants : les conséquences en termes de répartition, dont on a vu combien nos outils sur le sujet étaient pauvres ; les effets lointains, soit dans le temps, soit dans l'espace.

Il faudrait aussi réduire la manipulabilité des études appliquées et rendre leurs résultats plus fiables, transparentes et reproductibles : définir des règles d'exécution précises des études, comme cela est en cours à la suite de l'impulsion donnée par le rapport Boiteux, dont les principes sont maintenant déclinés par le Ministère des Transports dans des circulaires d'application coordonnées mode par mode ; il conviendrait aussi d'effectuer des contrôles de ces études, et pour cela de prévoir des audits extérieurs qui en rendraient les résultats mieux admis, parce que mieux compris et plus sûrs ; et aussi, pourquoi pas, soumettre les méthodes au débat public non seulement les résultats des études - ce qui est souvent fait maintenant - mais aussi leur méthodologie, de façon à leur assurer une autorité morale accrue².

¹ Verboef, Nijkamp et Rietveld (1997) ont par exemple analysé par enquête auprès des managers potentiels les réticences au péage urbain et mis en évidence quelques conditions à remplir pour les diminuer.

² Cette procédure est appliquée dans certains pays nordiques pour ce qui est des effets d'investissement, dont la valorisation monétaire est décidée au Parlement.

Le soucis d'assurer une mise en œuvre correcte des recommandations implique aussi une attention accrue au cadre institutionnel dans lequel elles s'appliquent. On a vu au chapitre IX l'influence du statut des opérateurs sur le coût des services qu'ils procurent. On a vu aussi au chapitre VII les écarts aux règles de choix des investissements que pouvait entraîner la structure dans laquelle elles s'appliquaient et les clauses auxquelles elles étaient soumises : administration ou concession, propriété publique ou privée. Tout ceci milite pour que l'analyse économique porte attention, non seulement aux formules des calculs, mais aussi aux cadres institutionnels dans lesquels elles s'appliquent.

2. Une politique des transports est-elle possible ?

L'économiste peut améliorer ses connaissances, en rendre l'application plus convaincante ; il restera toujours d'importantes zones d'ombre. Il peut porter plus d'attention aux dispositions institutionnelles, il ne pourra pas les transformer radicalement. Il peut analyser les différences de points de vue des acteurs, rechercher les conditions de l'acceptabilité sociale, il n'empêchera pas pour autant les conflits d'intérêts. On peut se demander si une politique cohérente dans les transports est possible. Éclairons cette question par l'analyse de quelques problèmes concrets, examinés sous des aspects particuliers au cours des chapitres précédents.

2.1. Organisation du marché et volonté politique : le cas des transports de marchandises

Le transport de marchandises est un bien intermédiaire qui concerne les entreprises, les problèmes d'équité et de répartition y sont donc moins importants que ceux d'efficacité. Les recommandations de l'économiste qui s'attachent à ce critère devraient donc être particulièrement pertinentes.

Ces recommandations reposent sur un diagnostic de la situation actuelle, qui se caractérise par une concurrence excessive, souvent destructrice, dans beaucoup de segments du transport routier de marchandises, notamment dans le transport non spécialisé. Elle est aussi marquée par une distorsion des conditions de concurrence entre le fer et la route, résultant d'une sous-tarifcation de l'activité routière, et du non respect de la réglementation dans ce mode, ce non respect portant essentiellement sur les horaires de travail et sur les surcharges. Cette distorsion des conditions de concurrence est accentuée par l'inefficacité dynamique du chemin de fer, dont la position de monopole assisté conduit à un service coûteux et de qualité médiocre. L'ensemble de ces dysfonctionnement se traduit par un coût collectif élevé, matérialisé par des gaspillages et une dépense publique excessive, et par des transferts en faveur des chargeurs, via les bas prix des transporteurs routiers, et en faveur du personnel de la SNCF ;

les perdants sont les employés du transport routier, soumis à une pression sur les conditions de travail et les salaires, et les citoyens-contribuables.

Restaurer l'efficacité comporte d'abord des actions des pouvoirs publics pour rétablir des conditions de concurrence normales : faire respecter la réglementation routière, relever les taxes qui frappent la route pour les amener au bon niveau de couverture des coûts. Mais ceci ne suffit pas ; il faut aussi assurer le bon fonctionnement des marchés : dresser des barrières à l'entrée plus élevées pour les transports routiers de marchandises, à travers par exemple les conditions d'accès à la profession (compétence technique et managériale, surface financière).

À l'inverse, pour le transport ferroviaire de marchandises, il faut introduire davantage de compétition dans le transport ferroviaire pour en assurer l'efficacité dynamique : la réforme récente aboutissant à la création de RFF permet d'isoler la gestion des infrastructures et d'améliorer l'efficacité de cette gestion, tout en ouvrant la porte à un traitement plus égal du fret et des voyageurs dans l'allocation des sillons. Il faudrait aussi (Henry et Quinet, 1997) isoler le fret au sein de la SNCF en substituant dans l'organisation de l'entreprise une logique de produits dont la gestion serait intégrée et autonome, à une logique de moyens. Il conviendrait enfin de favoriser une certaine concurrence sur le marché, en suscitant l'entrée de concurrents même si ceux-ci n'opèrent que sur une frange du marché. On peut à cet égard regretter que la France ne s'associe pas à l'opération « Freeways » de la Commission, qui aurait eu pour résultat d'introduire une plus grande liberté d'accès sur certains axes du réseau.

Ces recommandations sont simples. Mais on mesure les difficultés de leur mise en œuvre et on voit là un exemple où la logique économique se heurte à la logique sociologique. Pour dépasser cette opposition, on conçoit qu'il faille à la fois une attention accrue aux modalités de mise en œuvre des mesures, avec la définition et l'annonce de mesures d'accompagnement et d'un calendrier de réalisations progressives, et aussi une volonté politique tenace.

2.2. La gestion des infrastructures : norme ou discrétion ?

On oppose souvent deux modes de conduite des affaires publiques : la norme ou la discrétion. La discrétion consiste à prendre dans chaque situation la décision la plus adaptée aux circonstances particulières. Cette manière de faire est apparemment optimale, mais elle peut dégénérer en un abandon devant la facilité, elle présente les risques d'incohérence et de moindre résistance aux pressions, et ne permet pas de prévisibilité. La norme, qui correspond à l'annonce du respect de certaines règles, présente les caractéristiques inverses. La gestion des infrastructures peut s'analyser à travers cette opposition.

La situation actuelle est surtout marquée par la discrétion. Pour les routes ordinaires et les autoroutes hors péage, il y a dissociation entre les procédures de fixation des tarifs et les procédures de décisions de

exemple une tarification générale du type Ramsey-Boiteux, c'est-à-dire dont la structure recoupe le coût marginal social, et dont le niveau soit adapté pour satisfaire des contraintes de couverture globale de dépenses. Cet organisme devrait se voir donné une norme de dotation budgétaire, fixée par une procédure législative sur le moyen terme, de façon à assurer grosso-modo, un investissement jugé optimal. Enfin, un régulateur indépendant aurait pour mission de surveiller les tarifs, et, soit de les approuver, soit de fournir à leur égard une évaluation dont le pouvoir résulterait de l'autorité morale du régulateur, selon un mécanisme analogue à celui prôné par Henry (1997).

On se rend compte aisément de la distance qui sépare ce schéma de la situation existante, et là aussi on imagine l'attention qu'il faudrait porter aux moyens par lesquels il serait possible de réaliser la transition

2.3. Le contrôle de la mobilité ou les difficultés de coordonner la multiplicité

Les problèmes de la mobilité ont été évoqués à différentes occasions au cours des chapitres précédents sur des aspects partiels. Il est maintenant possible d'en prendre une vue de synthèse et d'en dégager des lignes de conduite politique.

On connaît les faits : la mobilité urbaine s'est développée au cours des dernières années, non pas tant par le nombre des déplacements par personne, que par un allongement des distances de déplacement, qui s'est effectué à travers un étalement des villes – c'est le phénomène de la périurbanisation –, et à budget-temps approximativement constant.

Cette évolution fait l'objet d'appréciations divergentes. Pour certains, elle constitue un progrès, dans la mesure où elle permet aux citoyens de profiter d'une plus grande diversité de biens grâce à une accessibilité accrue ; et où elle permet aux entreprises de bénéficier de plus larges aires de marché, et de mieux développer leurs économies d'échelles. Pour de nombreux autres analystes cependant, elle est le signe d'un dysfonctionnement, entraînant une hypertrophie urbaine, et des atteintes à la qualité de l'environnement et du cadre de vie.

Le point de vue adopté ici sera, non pas de juger le bien-fondé de telle ou telle politique, ni d'en recommander une, mais simplement d'évaluer les lois dont dépend la mobilité, et les possibilités que nous avons de l'influencer dans un sens ou dans un autre, selon les orientations que pourrait vouloir imprimer une volonté politique.

La mobilité a fait l'objet de nombreuses études et a été l'occasion de rassembler une masse considérable d'informations statistiques, plus peut-être qu'aucun autre phénomène social. Ce savoir considérable laisse des impressions très différentes selon l'échelle de temps que l'on considère.

Pour ce qui concerne les effets à court terme, assimilables à ce que les modèles de trafic traitent au titre du choix d'itinéraire et du choix modal,

les lois sont assez bien connues. On sait que les comportements individuels sont diversifiés ; mais on maîtrise à peu près cette diversité, on connaît les paramètres qui la caractérisent, on peut quantifier leur modalité d'intervention. Les modèles de prévision de trafic prévoient assez fidèlement, on l'a vu, la réalité ultérieure sous réserve de quelques précautions de mise œuvre.

Il n'en est pas de même pour les effets à long terme, ceux qui dans les modèles de trafic correspondent aux phases de génération et distribution. On dispose certes d'une masse de connaissances et d'information en ces domaines. Mais on en reste à l'état descriptif, comme par exemple la loi de Zahavi, ou la formule gravitaire, qui est analogue avant d'être explicative. La raison en est compréhensible : ces effets, que l'on pourrait percevoir à la suite de la réalisation d'une nouvelle infrastructure par exemple, se font sentir dans la durée. Ils sont peu perceptibles sur courte période, et sur longue période ils sont noyés au milieu des nombreuses autres causes de changements. En outre ils ne s'exercent pas seulement dans le domaine des transports, mais s'étendent, à travers des chaînes causales complexes, à l'utilisation des sols, à la répartition spatiale des activités, à l'organisation des entreprises et des programmes d'activité des individus. Cette implication de domaines et disciplines diverses n'est pas faite pour faciliter l'évaluation des phénomènes. Elle montre aussi que les changements dans l'offre de transport ont à terme des conséquences qui s'étendent bien au-delà du secteur.

Le moyen de contrôle serait, pour l'économiste, facile à trouver et à administrer si l'on pouvait mettre en place un système efficace de prix. Or, le marché des transports, et notamment des transports automobiles, connaît plusieurs causes de déficiences : d'abord il n'y existe pas de prix d'usage ; l'automobiliste paie son transport indirectement à travers de nombreux actes liés au transport : l'achat du véhicule, sa possession, la consommation de carburant, le péage parfois ; mais il ne paie pas le transport lui-même. En outre le coût du déplacement, si l'on veut qu'il représente le véritable coût pour la collectivité, ne peut pas résulter du simple jeu du marché, mais doit intégrer de nombreuses externalités, qui présentent deux caractéristiques gênantes : elles sont mal connues, chiffrées avec une imprécision plus grande que celle qui frappe les coûts privés ; le peu qu'on sait d'elles donne à penser qu'elles sont importantes par rapport aux coûts privés.

Encore ces chiffrages supposent-ils que les secteurs voisins concernant la localisation des activités et l'urbanisme sont tarifés de façon correcte. Il faudrait donc que les charges pesant sur ces secteurs soient optimisées également, ce qui est loin d'être le cas (Martin, 1997) : subventions au logement, taxes sur les transactions. Mais même ceci ne suffirait pas, car on ne peut faire l'hypothèse que le jeu du marché conduise à un optimum dans l'interaction transport-urbanisme. On y rencontre des effets externes, souvent positifs, et des rendements croissants, comme on l'a vu au chapitre III.

Les mécanismes en jeu dans ces domaines ont d'ailleurs aussi des conséquences en matière de répartition et d'équité : certaines zones, certaines tranches de population sont avantagées, d'autres désavantagées. Et la collectivité a légitimement des préférences sur ces points ; elle ne peut laisser un mécanisme aveugle et maladroit trancher pour elle.

L'enseignement de l'analyse économique, que d'ailleurs le bon sens approuve, c'est qu'alors il convient de mettre en œuvre de façon coordonnée tous les instruments – prix, taxes, réglementation... – dont on dispose. Or si les instruments qui permettent d'agir sur la mobilité sont nombreux, leur coordination semble difficile à assurer. Ils sont en effet aux mains d'acteurs multiples, dont les intérêts sont en général divergents. La taxation des véhicules, le prix des carburants sont dans la main de l'état central, qui les fixe en fonction de considérations budgétaires dans lesquelles la politique des transports tient une place réduite. Et lorsque cette dernière s'exprime c'est plutôt à la suite de pressions que pour satisfaire à des considérations d'intérêt général, comme le montre l'exemple du prix du gazole. Le stationnement et la police de la circulation dépendent des communes, et les décisions le concernant sont donc morcelées au sein d'une même agglomération. L'urbanisme, qui constitue une sorte de variable duale de la mobilité, est lui aussi dans la main d'autorités communales morcelées même sur le territoire d'une même agglomération. La réalisation d'infrastructures dépend des différents niveaux de pouvoirs politiques, dont les intérêts sont souvent divergents.

En outre chacun des acteurs se trouve dans une situation naturelle d'incitation à la resquille (« free-rider ») : on a intérêt à ce que les autres suivent une discipline collective, mais on n'a pas intérêt à la suivre soi-même.

Enfin chaque action aboutit souvent à long terme à un effet contraire à celui qui était voulu. L'exemple le plus frappant est lié à la réalisation d'infrastructures, faites pour décongestionner, et qui finalement appellent un trafic nouveau rétablissant, partiellement au moins, la congestion qu'on avait initialement supprimé.

Ce dernier exemple illustre aussi les conséquences d'une mauvaise combinaison des instruments de contrôle : là où il aurait fallu utiliser l'instrument tarifaire, le péage, on s'en prive et on est alors amené à traiter le problème par la réalisation de nouvelles infrastructures qui coûtent cher et ne sont qu'un palliatif temporaire.

De ce point de vue, le mélange des charges qui touchent la circulation et des politiques de stationnement et d'urbanisme pourrait être combiné de façon à optimiser leurs effets, comme cela est tenté et assez bien réussi, semble-t-il, à Singapour, et dans certaines villes européennes, dont Zurich est le meilleur exemple. Mais dans bien d'autres pays, l'usage des instruments est tiraillé entre la protection de l'industrie, l'alimentation des finances publiques, et la cohérence de la politique des transports qui n'a souvent que la portion congrue.

Notre relative ignorance des phénomènes, surtout des conséquences à long terme, la faiblesse de nos instruments d'action, les difficultés qu'il y a

à les mobiliser de façon cohérente, expliquent nos problèmes pour contrôler la mobilité, et la surprise qu'on peut ressentir devant des évolutions non voulues et non maîtrisées.

Mais quand on analyse les tendances émergentes actuellement et qu'on les projette – avec un peu d'optimisme – dans l'avenir, on trouve quelques raisons de changement, et quelques possibilités de progrès.

La première réside dans les modifications que va introduire l'information à travers le péage électronique. Celui-ci fournira un moyen unifié de faire payer la circulation routière et d'éviter les distorsions entraînées par l'existence de multiples moyens gérés de façon contradictoire. Certes il faudra que le péage électronique intègre bien tous les effets externes liés aux transports. Certes aussi il n'est pas encore accepté politiquement, mais on peut penser qu'il ne s'agit que d'une affaire de temps, et qu'une analyse sérieuse des raisons de réticence des usagers permettra de mettre en œuvre les mesures d'accompagnement nécessaires (Verhoef, Nijkamp et Rijnveld, 1997).

Les télécommunications vont apporter d'autres changements : l'information des usagers, la voiture intelligente, les transports collectifs vont révolutionner les pratiques de mobilité ; de nouvelles formes d'adaptation vont apparaître, comme la modulation des heures de déplacement, les changements d'itinéraire en temps réel ; enfin le développement des télécommunications va changer notre manière d'exercer les activités, avec le télétravail, la transformation des temps de déplacement en temps masqué. Le savoir accumulé sur la mobilité va devoir être renouvelé.

Un autre impératif pour la recherche est de mieux appréhender les liens entre transports et aménagement. Il faut pouvoir identifier ces liens pour les prévoir. La démarche de planification usuelle est de se fixer des objectifs d'urbanisme, puis d'en déduire les conséquences en termes de transports. Il faut être capable de dérouler la causalité inverse, de façon à prévoir les conséquences en termes de tendances d'urbanisation des décisions de transports, et de contrôler leur cohérence avec les objectifs de long terme.

Si cela était réalisé, on disposerait alors d'un outil performant de tarification pour contrôler les effets de court terme, de possibilités accrues de gestion en temps réel, des connaissances nécessaires pour mieux apprécier, donc contrôler, les effets de long terme sur l'urbanisme.

Mais encore faudrait-il que ces moyens soient coordonnés dans la perspective d'une politique des transports. Deux grandes conditions pour cela : sur le plan organisationnel, une réduction du nombre des acteurs et décisions sur les plans sectoriels et géographiques. C'est en quelque sorte, à l'inverse des tendances ambiantes, plaider pour une planification intégrée et centralisée à des niveaux adéquats, un problème que les auteurs anglo-saxons traitent du nom de fédéralisme.

Il restera quand même une pluralité d'outils pour contrôler la mobilité, et leur coordination impliquera une volonté affirmée et continue. Com-

ment se créera cette volonté, quelles forces la maintiendront ? Le développement des processus démocratiques, la généralisation de débats, la diffusion de l'information y suffiront-ils ?

3. Conclusion

Les recommandations de l'économiste sont d'ambitions diverses. Elle peuvent définir un objectif précis, et cette position, caractéristique de l'impartialisme souvent reproché aux économistes, est surtout valable lorsque l'efficacité seule est en jeu. Elles peuvent aussi, plus modestement, définir les conditions d'obtention d'une politique dont les objectifs sont définis par ailleurs.

Elles impliquent en tout cas un changement des situations existantes et nécessitent la volonté collective de le mettre en œuvre. Cette volonté, l'économiste n'a pas le pouvoir de la créer. Mais si elle existe, même ténue, son rôle est de la guider et de l'aider à se réaliser. Et si ses recommandations sont comprises et prouvent leur bien-fondé, peut-être pourra-t-il contribuer à la renforcer, et par là même à accroître sa propre efficacité.

BIBLIOGRAPHIE

- V. Bernat et M. Ollivier-Trigalo (1997), « Politique des transports en France. Processus de décision : discours et pratiques », Projet TENASSESS, Work package 6, INRETS.
- K. Button et E. Verhoef (1996), *Road pricing, traffic congestion and the environment issues in efficiency and social feasibility*, Aldershot ; Edward Elgar.
- CEMT (1997), « Projet de recommandations politiques », Rapport préparé par A. Bleijenberg.
- Commission des Communautés Européennes (1998), « Livre Vert relatif aux ports et aux infrastructures maritimes ».
- J.M. Fourniau (1994), « Analyse et évaluation des politiques de transport », Séminaire INRETS 9 février.
- C. Henry (1997), *Efficacité et concurrence dans les services publics*, PUF, Paris.
- C. Marchetti (1991), « On mobility », First Progress Report IIASA, Laxenburg.
- C. Marchetti (1992), « What they need is speed », IISA, Laxenburg.
- OCDE-CEMT (1995), « Transports urbains et développement durable », CEMT.
- J.B. Opschoor, A.F. de Savornin Lohman et H.B. Vos (1994), « Gérer l'environnement. Le rôle des instruments économiques », Rapport préparé pour l'OCDE.
- J.P. Orfeuil (1997), « Les difficultés de l'évaluation des effets externes et les problèmes posés par leur internalisation », in « Environnement, véhicules et mobilité urbaine », Actes 59 INRETS.

- Cl. Marinand (1993), *L'expérience française du financement privé des équipements publics*. Economica.
- Y. Martin (1995), « Transport et développement durable », note ronéo de 1911.
- E. Quinet et B. Walliser (1998), « L'économiste, l'ingénieur, le médecin », *Le Marché*, 6 janvier.
- Report 251 (1997), « Toward a sustainable future », TRB Special Report, Washington D.C.
- Royal Commission on Environmental Pollution (1994), *Transport and the Environment* 18th Report HMSO, Londres.
- E. Verhoef, P. Nijkamp et P. Rietveld (1997), « The social feasibility of road pricing », *Journal of Transport Economics and Policy*, September.
- A.G. Wilson (1998), « Land use/Transport interaction models : past and future », *Journal of Transport Economics and Policy*, January.
- C. Winston (1985), « Conceptual developments in the economics of transportation : an interpretative survey », *Journal of Economic Literature*, vol. XXIII.
- World Bank (1995), *Sustainable transport : a sector policy review*, World Bank, August.

Annexes

Annexe 1. Outils de la maximisation sous contrainte

On cherche à résoudre le problème suivant :

Soient :

- x un vecteur de composantes x_1, x_2, \dots, x_n
- f une fonction réelle de x
- g_i des fonctions réelles de x avec i variant de 1 à m
- f et $(-g_i)$ sont des fonctions concaves ou quasi-concaves.

Trouver le vecteur x tel que :

$$f(x) \text{ soit maximum}$$

avec :

$$g_i(x) \leq 0 \quad \forall i$$

La solution est obtenue en formant le Lagrangien :

$$L(x, \lambda) = f(x) - \sum_i \lambda_i g_i(x)$$

et en annulant les dérivées de L par rapport aux composantes de x et λ :

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_i \lambda_i \frac{\partial g_i}{\partial x_j} = 0$$

et :

$$\lambda_i > 0 \quad g_i(x) = 0$$

ou :

$$\lambda_i = 0 \quad \text{et} \quad g_i(x) < 0$$

Un théorème utile est le théorème de l'enveloppe :

Si f et les g_i dépendent d'un paramètre a , le maximum de f dépend également de a ; c'est une fonction $F(a)$ et :

$$\frac{dF}{da} = \frac{\partial f}{\partial a} - \sum_i \lambda_i \frac{\partial g_i}{\partial a}$$

Nota :

- Une fonction $f(x)$ est concave si l'ensemble :

$$Z = \{x, y\}$$

tel que :

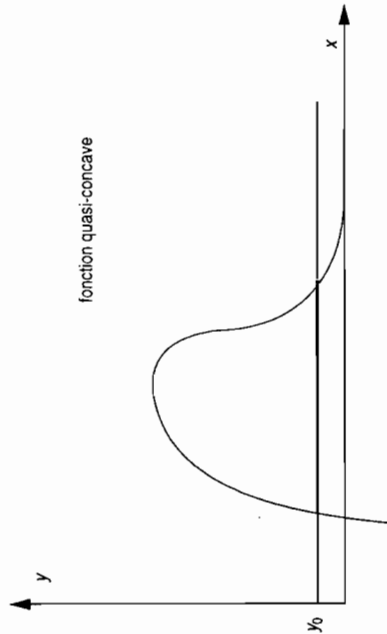
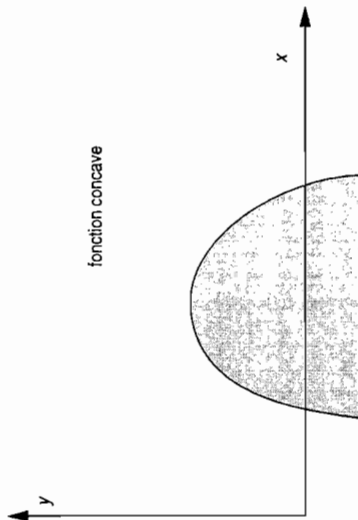
$$y \leq f(x)$$

est convexe.

- Une fonction $f(x)$ est quasi-concave si :

$$X(y_0) = \{x : y_0 \leq f(x)\}$$

est convexe.



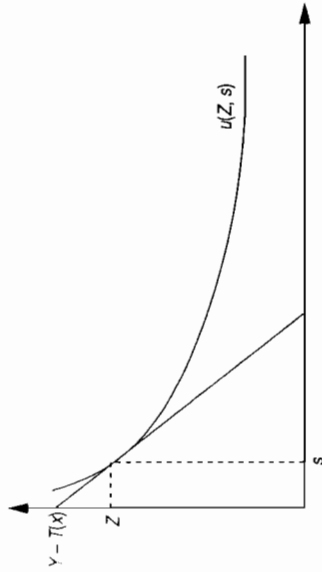
Annexe 2. Le mécanisme de la rente foncière¹

On va ici présenter un modèle de formation de la rente foncière analogue à celui présenté dans le chapitre 3 à cette différence près que l'hy-

1. Cette présentation classique est inspirée de Fujita (1989), Henderson (1985) et Ponsard (1988).

pothèse de constance de la surface de logement par individu faite dans cette présentation est abandonnée. La démarche est alors plus complexe, et avant de déterminer l'équilibre, il faut d'abord déterminer le prix maximum que chaque individu est prêt à proposer pour une unité de surface de logement. Ce prix maximum dépend du revenu Y (supposé d'abord le même pour tous), de la quantité de bien commun Z acquise (on suppose que le prix du bien commun est l'unité), de la distance au centre, de la surface du logement s que l'individu va choisir, et du niveau d'utilité $u(Z, s)$ que l'individu en question atteint (cf. figure A2.1).

FIGURE A2-1
Prix d'enchère maximale



Ce prix d'enchère maximale est donné par le programme suivant :

$$\max_{s, Z} \frac{Y - T(x) - Z}{s}$$

avec :

$$U(Z, s) = U_0$$

On peut résoudre en Z la dernière égalité

$$Z = Z(u_0, s)$$

Et le prix d'enchère maximale s s'exprime par :

$$\Psi(Y - T(x), u_0) = \max_s \left(\frac{Y - T(x) - Z(u_0, s)}{s} \right) \quad (\text{A 3-1})$$

On en tire l'expression de la surface unitaire du logement correspondant à cette enchère maximale : $s(Y - T(x), u_0)$. De l'expression (A 3-1), on tire quelques relations remarquables ; par le théorème de l'enveloppe,

$$\frac{\partial \Psi}{\partial x} = -\frac{1}{s} \frac{dT}{dx}$$

$$\frac{\partial \Psi}{\partial u} = -\frac{1}{s} \frac{\partial Z}{\partial u}$$

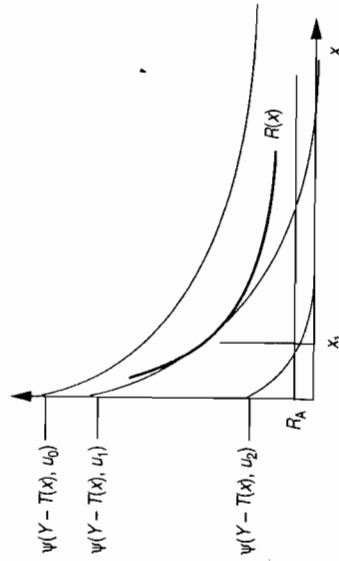
$$\frac{\partial \Psi}{\partial Y} = \frac{1}{s} \geq 0$$

On démontre¹ aussi que :

$$\frac{\partial s}{\partial x} \geq 0, \quad \frac{\partial s}{\partial u} \geq 0, \quad \frac{\partial s}{\partial Y} \leq 0$$

On démontre aussi que si $T(x) < 0$, alors les courbes Ψ sont convexes en X . La famille de ces courbes, pour un revenu Y donné et lorsque u varie, présente l'allure figurée sur la figure A2-2.

FIGURE A2-2
Choix de la localisation



Appelons alors $R(x)$ la valeur, supposée connue, de la rente foncière en x . L'individu dont le comportement est analysé et représenté par le réseau des courbes Ψ se localisera en x_1 , point où son utilité est maximale (elle vaut u_1), et occupera la surface

$$s(Y - T(x_1), u_1)$$

On aura alors, à l'équilibre, une généralisation de la relation 3-1 du chapitre III sous la forme :

$$\frac{dR}{dx} = \frac{\partial \Psi}{\partial x} = -\frac{1}{s} \frac{dT}{dx} \quad (\text{A3-2})$$

1. Avec l'hypothèse que la terre est un bien normal (la quantité consommée croît avec le revenu), et en notant que s est égale à la demande hicksienne du sol $s(\Psi, u) = s(Y - T(r), u)$ lorsque $\Psi = \Psi(Y - T(r), u)$.

Mais comment déterminer la courbe de la rente foncière à l'équilibre ? Deux cas sont simples à analyser : celui où les agents sont identiques, non seulement par leur fonction d'utilité mais aussi par leur revenu ; et celui où il y a un nombre fini de types différents d'agents, chacun d'eux se divisant en deux cas selon que le nombre d'habitants de la ville est fixe ou non.

A.2.1. Agents identiques

Ce cas se divise lui-même en deux cas polaires.

A.2.2.1. Nombre d'habitants fixé

Supposons d'abord que le nombre d'habitants N_0 de la ville soit fixé. Alors chaque agent maximise son utilité pour choisir sa localisation. L'équilibre exige que en chacune des localisations, l'utilité soit la même (sinon certains agents auraient intérêt de se déplacer). La courbe de rente foncière est donc un résultat du programme suivant :

$$\max U$$

tel que :

$$\int_0^{x_M} \frac{2\pi x}{S(Y - T(x), U)} dx = N_0$$

$$\Psi(Y - T(x_M), U) = R_A$$

Soit U^* cette solution. La courbe de rente foncière est donc :

$$R(x) = \Psi(Y - T(x), U^*) \quad \text{si } x \leq x_M$$

$$= R_A \quad \text{dans le cas contraire}$$

A.2.2.2. Nombre d'habitants variable

On a supposé précédemment que le nombre N_0 d'habitants était fixé de façon exogène, c'est-à-dire qu'aucune migration n'était possible ; cette situation de stabilité de la population est proche de celle qui se rencontre dans les agglomérations ayant atteint une certaine maturité, par exemple, l'ensemble de la région parisienne. Elle ne reflète toutefois pas les situations fréquentes où se produisent des migrations entre ville et campagne, ou entre villes de tailles différentes. On peut modéliser cette situation en faisant l'hypothèse comportementale que les migrants sont attirés par un niveau d'utilité attendu supérieure à celui de leur localisation présente. Dans ces conditions les migrations se prolongent jusqu'à ce que l'utilité dans la ville devienne égale à l'utilité dans le reste du pays, noté \bar{u} . L'équilibre foncier est alors fixé par :

$$R(x) = \Psi(Y - T(x), \bar{u}) \quad \text{si : } x \leq x_M$$

$$= R_A \quad \text{dans le cas contraire}$$

avec x_M tel que :

$$\Psi(Y - T(x_M), \bar{u}) = RA$$

et :

$$N = 2\pi \int_0^{x_M} \frac{x dx}{s(Y - T(x), \bar{u})}$$

Jusqu'ici on a supposé implicitement que les revenus des propriétaires fonciers n'étaient pas réinjectés dans l'économie de la ville : c'était l'hypothèse des « propriétaires absents ». Dans l'autre hypothèse, celle des « propriétaires présents », un cas simple à traiter est celui où la rente différentielle est également répartie entre tous les habitants. Les solutions de l'hypothèse « propriétaires absents » sont applicables en ajoutant au niveau du revenu Y jusque là considéré la part de rente différentielle afférent à chaque individu ; le revenu Y est à remplacer par $(Y + RDT/N)$; mais le processus de détermination de l'équilibre est plus complexe car la rente différentielle totale n'est pas connue a priori, elle résulte de l'équilibre.

À noter qu'on démontre assez facilement, à partir de (A3-2.) la relation établie au chapitre 3 dans le cas où S est fixé :

$$RDT = \frac{\alpha}{1 - \alpha} CTT$$

On démontre aussi que ces équilibres sont des optimums, dans la mesure où on se limite aux optimums égalitaires, dans lesquels chaque agent bénéficie du même niveau d'utilité et est donc indifférent à la localisation ; la courbe donnant la rente foncière en fonction de la distance au centre est une courbe Ψ ; on démontre enfin que réciproquement tout optimum égalitaire peut être obtenu par le jeu des enchères foncières, assorti d'une taxe ou subvention uniforme des agents pour atteindre le niveau d'utilité désiré (cf. Fujita, 1989).

A.2.2. Agents différents

On supposera ici qu'il y a deux catégories d'agents répertoriés par les indices 0 et 1, et qu'ils ne se différencient que par leur revenu $Y_0 < Y_1$; on se placera dans le cas d'une ville fermée (il y a N_0 agents du type 0 et N_1 du type 1) et de propriétaires absents.

La détermination de l'équilibre résulte de la remarque que lorsque deux courbes Ψ se coupent, la plus pentue correspond à l'individu ayant le revenu le plus élevé¹.

1. En effet, comme la terre est bien normal, l'individu ayant le revenu le plus élevé en loue une plus grande superficie et donc, comme

$$\Psi_r = \frac{\partial \Psi}{\partial r} = -\frac{T}{s}$$

la pente de la courbe Ψ , qui est négative, sera pour lui plus élevée en valeur absolue.

L'équilibre s'en déduit immédiatement :

- les individus de revenus Y_0 se localisent le plus près du centre, la rente qu'ils sont prêts à payer est donné par :

$$R_0(x) = \Psi(Y_0 - T(x), u_0)$$

Ils s'étendent jusqu'à la distance x_1 , telle que

$$\int_0^{x_1} \frac{2\pi x}{s(Y_0 - T(x), u_0)} dx = N_0$$

- les individus de revenus Y_1 se localisent autour ; la rente qu'ils paient est donnée par :

$$R_1(x) = \Psi(Y_1 - T(x), u_1)$$

avec :

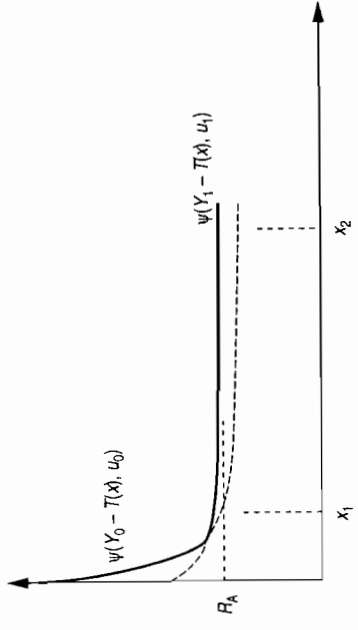
$$\Psi(Y_1 - T(x_0), u_1) = \Psi(Y_0 - T(x_0), u_0)$$

$$\int_{x_1}^{x_2} \frac{2\pi x}{s(Y_0 - T(x_0), u_1)} dx = N_0$$

$$\Psi(Y_0 - T(x_2), u_0) = R_A$$

Ceci définit un système de 4 équations à quatre inconnues : u_0, u_1, x_1, x_2 qui définit l'équilibre, dont on démontre qu'il est unique, comme le fait comprendre la figure A2-3. La courbe donnant la rente foncière est composée de deux courbes Ψ .

FIGURE A2-3
Équilibre avec 2 catégories d'agents



Annexe 3. La théorie du comportement du consommateur

Dans sa présentation la plus classique¹, la théorie du consommateur part de l'existence, pour chaque consommateur, d'une fonction d'utilité² qui dépend de la quantité de chaque bien consommé ; cette fonction d'utilité, qui n'est définie qu'à une fonction monotone croissante près, est une fonction croissante (dans une conception moins stricte, une fonction non décroissante) de ses arguments :

$$U = U(x_1, x_2, \dots)$$

Le consommateur est indifférent entre deux paniers de biens auxquels sont attachées des utilités égales. Entre deux paniers présentant des utilités différentes, il préférera celui qui a l'utilité U la plus forte. Placé devant un système de prix p_1, \dots, p_m pour chacun des biens, et limité par son revenu r , il cherche à maximiser son utilité :

$$\max U(\dots, x_i, \dots) \quad \text{avec :} \quad \sum p_i x_i \leq r$$

Avec l'hypothèse que U est concave ou quasi-concave (voir Annexe 1 sur la maximisation sous contrainte) la solution de ce programme fait intervenir la variable duale λ du Lagrangien :

$$L = U(\dots, x_i, \dots) - \lambda \left(\sum_i (p_i x_i - r) \right)$$

et on obtient un premier résultat, qui est la proportionnalité des utilités marginales aux prix :

$$\frac{\partial U}{\partial x_i} = \lambda p_i$$

V , valeur de l'utilité maximisée, est une fonction des paramètres p_i et r :

$$V = V(\dots, p_i, \dots, r) \quad (\text{A3-1})$$

Le théorème de l'enveloppe apprend que :

$$\frac{\partial V}{\partial r} = \lambda$$

et λ apparaît ainsi comme l'utilité marginale de la monnaie. V est appelée fonction d'utilité indirecte. Les relations précédentes et la contrainte de revenu constituent un système d'équations permettant de définir la demande de chaque bien en fonction des prix et du revenu :

$$x_i = x_i(\dots, p_i, \dots, r)$$

Cette demande est appelée demande Marshallienne, par opposition avec la demande Hicksienne, celle qui correspond à une minimisation de la dépense nécessaire pour obtenir un niveau d'utilité donné et qui correspond au programme :

$$\min r = \sum p_i x_i \quad \text{tel que :} \quad U(\dots, x_i, \dots) \geq V^*$$

La solution de ce programme aboutit à définir en fonction des paramètres p_i et V^*

$$r = E(\dots, p_i, \dots, V^*) \quad (\text{A3-2})$$

et on peut voir que A3.1 et A3.2 sont équivalents. On déduit de ce dernier programme les demandes hicksiennes :

$$x_i^*(p_1, \dots, p_i, \dots, V^*)$$

On a :

$$x_i^*(\dots, p_i, \dots, V^*) \equiv x_i(\dots, p_i, \dots, E(\dots, p_i, \dots, V^*))$$

D'où :

$$\frac{\partial x_i^*}{\partial p_j} = \frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial E} \frac{\partial E}{\partial p_j}$$

On peut déduire de ces définitions deux identités remarquables, les relations de Slutsky et de Roy. En appliquant le théorème de l'enveloppe au programme de minimisation de r , on a :

$$\frac{\partial E}{\partial p_j} = x_j^*$$

d'où l'équation de Slutsky :

$$\frac{\partial x_i^*}{\partial p_j} = \frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial r} x_j^*$$

On a par ailleurs :

$$\frac{\partial x_i^*}{\partial p_j} = \frac{\partial x_j}{\partial p_i}$$

puisque les 2 membres expriment la dérivée seconde de E par rapport à p_i et p_j

Enfin, l'égalité de Roy s'obtient en appliquant deux fois le théorème de l'enveloppe au premier programme de maximisation :

$$\frac{\partial V}{\partial p_i} = x_i(\dots, p_i, \dots, r) = - \frac{\partial p_i}{\partial V} \frac{\partial V}{\partial r}$$

1. Voir Varian (1992).

2. Dont l'existence et les propriétés peuvent être démontrées à partir d'un certain nombre d'axiomes auxquels répondent des choix cohérents : transitivité, répétitivité...

Les fonctions de demande x_j , définies précédemment par :

$$x_j = x_j(\dots, p_j, \dots, r)$$

sont des fonctions de demande individuelles, spécifiques à chaque individu j : marquons cette spécificité en affectant les variables propres à chaque individu d'un indice j :

$$x_j^i = x_j^i(\dots, p_j, \dots, r^j)$$

La fonction de demande la plus usuellement appréhendée est relative à l'ensemble des individus de la collectivité. Elle a comme expression :

$$X_j = \sum_j x_j^j$$

Elle dépend des prix p_j , variable commune à tous les individus, et des n revenus r^j de chaque individu. Ce dernier point est particulièrement gênant pour les applications, et on voudrait souvent remplacer ces n variables r^j par une seule variable R qui en serait la somme :

$$R = \sum_j r^j$$

Sauf si l'on contraint les r^j (par exemple en supposant que chaque individu dispose d'une fraction constante du revenu total, ce qui implique des mécanismes forts de contrôle des revenus), ceci est réalisé dans l'hypothèse où les demandes individuelles sont de la forme :

$$x_j^i = \alpha_j^i(\dots, p_j, \dots) + r^j \beta_j^i(\dots, p_j, \dots)$$

L'identité de Roy permet de voir que cette relation découle d'une fonction d'utilité indirecte de la forme :

$$v^j(\dots, p_j, \dots, r^j) = a^j(\dots, p_j, \dots) + r^j b^j(\dots, p_j, \dots)$$

Une forme particulière de fonction d'utilité souvent utilisée en raison de sa commodité analytique est la fonction dite quasi-linéaire, écrite ici avec 3 biens :

$$U(x_0, x_1, x_2) = x_0 + u(x_1, x_2)$$

On utilise souvent cette expression en mettant sous x_0 un composé de l'ensemble des biens sauf les biens 1 et 2 sur lesquels se focalise l'attention.

Alors les fonctions de demande et d'utilité indirecte ont une expression particulièrement simple. En prenant, selon l'usage, le prix du bien 0 comme unité monétaire, le Lagrangien s'écrit :

$$x_0 + u(x_1, x_2) - \lambda(x_0 + p_1 x_1 + p_2 x_2 - r)$$

L'utilité marginale de la monnaie prend la valeur :

$$\lambda = 1$$

et les lois de demande des biens 1 et 2 ne dépendent pas du revenu (mais la demande de 0 en dépend) :

$$x_1 = x_1(p_1, p_2)$$

$$x_2 = x_2(p_1, p_2)$$

$$x_0 = r - p_1 x_1 - p_2 x_2$$

L'utilité indirecte, obtenue en remplaçant dans U les x_j par leur expressions en fonction des prix et du revenu prend la forme souvent utilisée :

$$V(p_1, p_2, r) = r + v(p_1, p_2)$$

L'hypothèse forte au prix de laquelle est obtenue cette commodité analytique est que la demande des biens x_1 et x_2 soit indépendante du revenu, une condition rarement réalisée dans la pratique.

Annexe 4. La fonction de coût. Rappel théorique

La théorie économique classique ¹ prend le cas d'une entreprise (ou plus généralement, d'un organisme chargé de la production) qui fabrique un bien en quantité q à partir de facteurs de production x, y, z . Ces variables sont reliées par la fonction de production qui définit la quantité maximale de bien q susceptible d'être produite à partir de quantités données des facteurs de production x, y, z :

$$q = f(x, y, z)$$

À partir de là, on définit une fonction de coût comme le coût minimum de production de la quantité q , avec des prix donnés des facteurs de production p_x, p_y, p_z

$$C = \min (xp_x + yp_y + zp_z)$$

tel que :

$$q \leq f(x, y, z)$$

Avec l'hypothèse peu contraignante que la fonction ($-f$) est concave ou quasi concave, on résout ce problème de minimisation de C (c'est-à-dire de maximisation de $-C$) par les outils décrits en Annexe 1. En formant le Lagrangien, qu'on dérive par rapport à x, y , et z , on en tire le coût minimum C , qui dépend des paramètres q, p_x, p_y, p_z :

$$C = C(q, p_x, p_y, p_z)$$

Du théorème de l'enveloppe appliqué au Lagrangien du programme précédent, on déduit (Lemme de Shephard) :

$$x = \frac{\partial C}{\partial p_x}$$

et les égalités similaires pour y et z . Ces égalités fournissent les fonctions de demande des facteurs de production.

1. Voir Varian (1992) ou Malinvaud (1975).

En considérant p_x, p_y , et p_z comme des données, on définit le coût marginal :

$$\frac{\partial C}{\partial q}$$

et le coût moyen :

$$\frac{C(q)}{q}$$

Bien sûr :

$$\frac{dC}{dq} = p_x \frac{dx}{dq} + p_y \frac{dy}{dq} + p_z \frac{dz}{dq}$$

pourvu que dq, dx, dy, dz vérifient :

$$dq = f'_x dx + f'_y dy + f'_z dz$$

Si x représente un facteur de production à ajustement lent, par exemple l'équipement, qui à court terme peut être considéré comme fixe, on définira la fonction de coût à court terme, c'est-à-dire à équipement x donné :

$$C(q, x) = xp_x + \min_{y,z} (yp_y + zp_z)$$

tel que :

$$q = f(x, y, z)$$

On en tire le coût marginal à court terme :

$$CmCT = \frac{\partial C(q, x)}{\partial q}$$

La fonction de coût à long terme, ou de coût adapté sera :

$$C^*(q) = \min (C(q, x) + xp_x)$$

D'où l'on peut calculer ce que l'on appelle le coût marginal de long terme :

$$CmLT = \frac{dC^*}{dq}$$

La minimisation en cause conduit à :

$$\frac{\partial C(q, x)}{\partial x} = p_x$$

relation d'où l'on tire le niveau d'équipement adapté pour un volume de production q donné :

$$x^* = x(q)$$

On a donc :

$$C^*(q) = C(q, x^*(q)) + x^*(q) p_x$$

Le théorème de l'enveloppe dit que :

$$\frac{dC^*(q)}{dq} = \frac{\partial C(q, x^*(q))}{\partial q}$$

ce qu'on exprime en disant qu'à équipement adapté ($x = x^*(q)$), le coût marginal de long terme est égal au coût marginal de court terme.

Deux concepts importants sont ceux d'économie (déséconomie) d'échelle sur la fonction de coût et de rendement croissants (décroissants) sur la fonction de production. Ils revêtent plusieurs aspects selon qu'on en cherche une définition locale ou globale. Appliqué à la fonction de production, on dira qu'on se trouve en présence de rendements croissants au point x, y, z si :

$$f(\lambda x, \lambda y, \lambda z) \geq \lambda f(x, y, z) \quad \text{avec : } \lambda \geq 1$$

Les économies d'échelle sur la fonction de coût s'expriment par la propriété :

$$C(\lambda q) \leq \lambda C(q) \quad \text{avec : } \lambda \geq 1 \quad (\text{A4-1})$$

À côté de cette définition globale, on a une définition locale, exprimée par :

$$\left. \frac{df(tx, ty, tz)}{dt} \right|_{t=1} \geq 1$$

La version locale correspondante pour la fonction de coût est exprimée par

$$\frac{C}{q} \frac{dC}{dq} \geq 1 \quad (\text{A4-2})$$

La quantité : $(C/q)(dC/dq)$ mesure les économies d'échelle.

On démontre que des rendements croissants (décroissants) sur la fonction de production entraînent des économies (des économies) d'échelle sur la fonction de coût et réciproquement : c'est ce que l'on appelle le principe de dualité, qui s'applique aussi bien localement que globalement.

La notion d'économies d'échelle est voisine de la propriété qui exprime qu'une seule entreprise est préférable à deux entreprises identiques pour fabriquer le même produit, situation qu'on qualifie de « monopole naturel ». C'est la sous-additivité des coûts :

$$C(q) \leq C(q_1) + C(q_2) \quad (\text{A4-3})$$

quels que soient q_1 et q_2 tels que :

$$q_1 + q_2 = q$$

Il est facile de voir que (A4-1) ou (A4-2) entraînent (A4-3), mais la réciproque n'est pas vraie.

Les raisonnements, résultats et concepts qui viennent d'être développés dans le cas d'une production unique (q scalaire) peuvent être étendus au cas d'une production multiple (q vecteurs). Les notions de fonction de production et de fonction de coût se généralisent aisément, ainsi que les concepts de coût marginal à court terme et à long terme. En revanche, celui de coût moyen n'a pas d'équivalent.

Les rendements d'échelles sur la fonction de production, ainsi que les économies d'échelle sur la fonction de coût se généralisent aisément ; il suffit de considérer $f(\cdot)$ et q comme des vecteurs et non comme des scalaires. La relation (4-1) reste valable. La relation (4-2) aussi, à condition que q et dC/dq représentent des vecteurs, dont $q \, dC/dq$ est le produit scalaire.

La condition pour qu'une entreprise soit plus efficace que plusieurs pour assurer un niveau de production q donné reste la même que celle donnée par (A4-3), sous réserve de considérer q , q_1 et q_2 comme des vecteurs à composantes positives.

Si l'on recherche des conditions suffisantes pour que (4-3) soit remplie, on trouve des relations plus complexes que (A4-1) ou (A4-2). Il faut d'ailleurs en général des couples de condition. Baumol, Panzar et Willig (1982) ont établi quelques uns de ces couples de conditions suffisantes. Ils reviennent tous à exprimer :

– une idée de rendements d'échelle homothétique (économie d'échelle) ; par exemple :

$$C(\lambda q) \leq \lambda C(q) \quad \text{avec} \quad \lambda \geq 1$$

– et une idée de complémentarité des produits (économies d'envire) ; par exemple, pour toute partition de q en 2 composantes (q_1, q_2) :

$$C(q) = C(q_1, q_2) < C(q_1, 0) + C(0, q_2)$$

Annexe 5. Critères de surplus en cas de transformations finies¹

A.5.1. Comparaison des utilités d'un même individu

Comment passer des transformations marginales aux transformations finies pour un même individu ?

La transformation finie est définie par les caractéristiques de l'état final $S_1 : (p^1, q^1, r^1)$ et celles de l'état initial $S_0 (p^0, q^0, r^0)$. Il serait tentant de l'évaluer par l'indicateur :

$$D = \int_{S_0}^{S_1} dU \quad (\text{A5-1})$$

dU étant fourni par l'expression (7-2) du chapitre 7¹. Mais dU n'est pas une différentielle totale (c'est $\lambda_i dU$ qui l'est) ; (A5-1) dépend du chemin d'intégration et n'a pas une valeur unique. On peut lever cette indétermination par une convention. La plus fréquemment recommandée est celle d'un chemin d'intégration linéaire ; on pose, pour chacune des variables x :

$$x(t) = (1-t)x_0 + tx_1$$

et on définit D par :

$$D = \int_{t=0}^{t=1} \sum_j p(t) (q_{ij}(1) - q_{ij}(0)) dt$$

Pour éviter l'arbitraire de cette définition on définit aussi deux autres mesures de l'intérêt d'une transformation pour un individu : la variation équivalente et la variation compensatoire.

La variation compensatoire est le revenu supplémentaire qu'il faudrait donner (ou prendre) à l'individu pour que, dans le système de prix final, il se trouve avec un niveau d'utilité inchangé.

La variation équivalente, dont la définition est très proche, est le revenu supplémentaire qu'il faudrait donner (ou prendre) à l'individu pour lui permettre, dans le système de prix initial, de maintenir son utilité au niveau initial. La variation équivalente présente l'avantage que, lorsqu'il s'agit de comparer plusieurs projets, elle n'utilise qu'un seul système de prix (le système de prix initiaux) et aboutit donc à un classement cohérent, ce qui n'est pas forcément le cas de la variation compensatoire (Kanemoto et Méra, 1985).

En simplifiant les notations et en supprimant l'indice i représentant l'individu, on notera le vecteur des biens par q , et le vecteur des quantités par p . Les utilités indirectes sont notées par la lettre v .

La variation équivalente est :

$$EV = E(p^0, v^1) - E(p^0, v^0)$$

Et la variation compensatoire est :

$$CV = E(p^1, v^1) - E(p^1, v^0)$$

$E(p, v)$ étant la dépense minimale nécessaire pour obtenir l'utilité v avec le vecteur de prix p . On a :

$$v^1 = v(p^1, p^1 q^1)$$

$$v^0 = v(p^0, p^0 q^0)$$

$$E(p^0, v^0) = p^0 q^0 = r^0$$

$$E(p^1, v^1) = p^1 q^1 = r^1$$

1. Voir Varian (1992), Kanemoto et Méra (1985), Morisugi (1983).

1. $dU = \sum_j p^j dq_j$

Ces quantités peuvent être rapprochées du surplus de Dupuit. Prenons le cas où le revenu ne change pas, seul un prix se modifiant. Alors EV peut s'écrire :

$$\begin{aligned} EV &= E(p^0, v^1) - r^0 - E(p^1, v^1) + r^1 \\ &= E(p^0, v^1) - E(p^1, v^1) = - \int_{p^0}^{p^1} \frac{\partial E}{\partial p}(p, v^1) dp \end{aligned}$$

comme :

$$\begin{aligned} \frac{\partial E}{\partial p} &= q^*(p, v) \\ EV &= \int_{p^1}^{p^0} q^*(p, v^1) dp \end{aligned}$$

et de même :

$$CV = \int_{p^1}^{p^0} q^*(p, v_0) dp$$

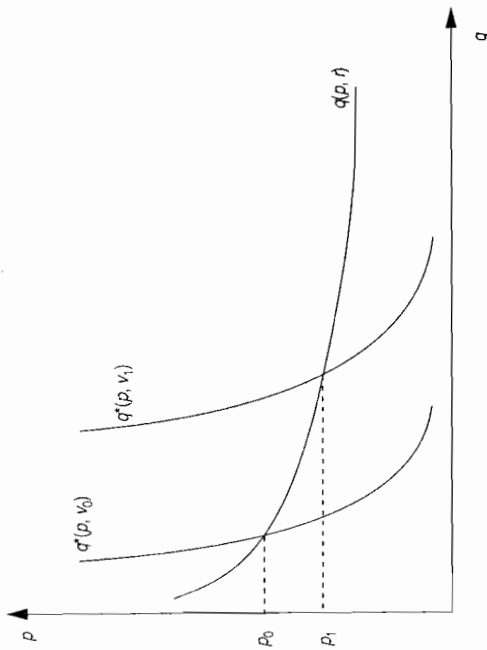
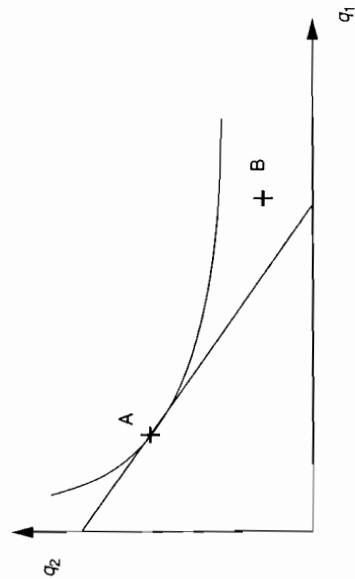
Ainsi EV et CV sont représentées par des intégrales de la fonction de demande compensée $q^*(p, v)$, alors que le surplus de Dupuit SD est représenté par une intégrale de la fonction de demande marshallienne $q(p, r)$.

Si le (ou les) bien(s) considérés sont normaux ($\partial q / \partial r \geq 0$), les fonctions de demande sont l'une par rapport à l'autre dans les positions indiquées sur la figure jointe, comme le fait voir l'équation de Slutsky :

$$\frac{\partial q}{\partial p} = \frac{\partial q^*}{\partial p} + \frac{\partial q}{\partial r}$$

On voit que EV et CV encadrent le surplus de Dupuit.

Notons enfin que EV , en tant que fonction monotone croissante de v , est en soit une fonction d'utilité, mais que la variation de revenu à prix constants n'est pas un bon indicateur de l'utilité, comme le montre l'exemple classique figuré ci-dessous, où l'individu consomme 2 biens 1 et 2, et où l'on compare deux états A et B. Le panier B ne peut être obtenu aux prix de A qu'avec un revenu supérieur à celui correspondant à A, mais présente une utilité moindre que A.



A.5.2. Aggrégation des utilités individuelles

Le critère admis est le critère de compensation : les gagnants ont de quoi compenser les perdants. Sa mise en œuvre rigoureuse est très délicate dans le cas de transformations finies. Il connaît deux expressions :

- Le test de Kaldor : le changement proposé est accepté s'il est possible de redistribuer les biens dans la situation finale de telle façon que la situation après redistribution soit Pareto-supérieure (préférée à la situation initiale pour au moins une personne, et équivalente pour les autres).
- Le test de Hicks : le changement est accepté si on ne peut pas redistribuer les biens de la situation initiale de telle façon que la situation initiale ainsi modifiée soit Pareto-supérieure à la situation finale.

A noter que ces règles ne sont pas transitives. Il y a des cas où, pour chacun des deux tests, la situation initiale est préférable à la finale et réciproquement (paradoxe de Scitowski).

On démontre que lorsque l'on n'est pas dans ce cas, alors les deux tests sont équivalents, et que la positivité de la somme des variations équivalentes sur l'ensemble des individus est une condition suffisante pour la satisfaction des tests ci-dessus (voir Morisugi, 1983 ou Kanemoto et Méra, 1985 pour une présentation adaptée au cas particulier des transports).

A.5.3. Cas de fonction d'utilité quasi-linéaire

Une situation particulière existe où les propriétés des transformations marginales restent valables pour les transformations non marginales, c'est celle de fonction d'utilité directe quasi linéaire. Dans l'exemple d'une économie à trois biens, elle s'écrit :

$$U = x_0 + f(x_1, x_2)$$

On se trouve souvent dans une situation de ce type lorsque les biens x_1 et x_2 sont liés mais indépendants du bien 0. Le bien 0 peut d'ailleurs représenter un bien composite, agrégat de l'ensemble des autres et il est alors commode de prendre comme prix unitaire le prix de ce bien, qui représente alors en quelque sorte le numéraire. Alors l'équation du budget s'écrit :

$$x_0 + p_1 x_1 + p_2 x_2 = r$$

On voit facilement que l'utilité marginale de la monnaie est constante et égale à un. Les fonctions de demande de x_1 et x_2 , définies par

$$\frac{\partial f}{\partial x_1} = p_1$$

$$\frac{\partial f}{\partial x_2} = p_2$$

sont indépendantes du revenu. Alors le surplus de Dupuit et la variation équivalente sont égaux, et l'analyse de transformation finies peut s'opérer avec les mêmes facilités que l'analyse marginale.

Annexe 6. Le traitement de l'incertitude ¹

A.6.1. Comportement des agents économiques vis-à-vis du risque

On se place dans la situation d'un agent confronté à des événements incertains munis de probabilités, et amené à choisir entre des loteries impliquant ces événements. On démontre alors, sous des hypothèses admissibles, que la rationalité de l'agent implique qu'il attache à ces événements une utilité définie, non plus seulement à une fonction monotone croissante près, mais plus précisément à une fonction affine près, et que la règle de choix de l'agent soit fondée sur l'espérance d'utilité des loteries ²

$$E(U) = \sum_i p_i u_i$$

- i représentant les issues possibles,
- p_i leur probabilité d'apparition,
- u_i leur utilité.

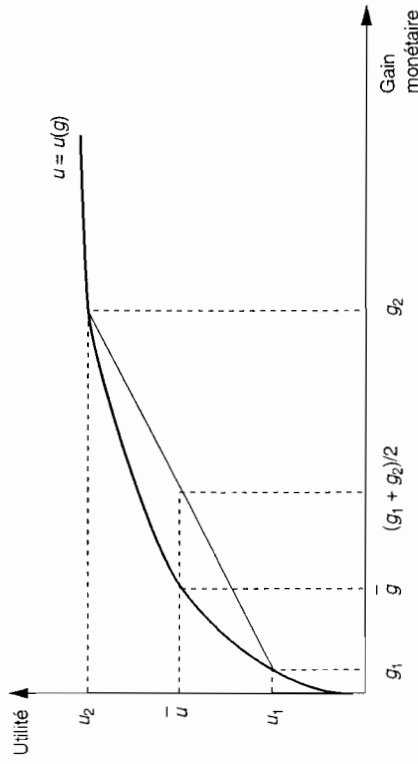
Comment les utilités ainsi définies sont-elles reliées aux gains monétaires des loteries ? La question était sans objet dans le cas d'événements certains puisque l'utilité était définie à une fonction monotone près ; elle

1. Voir Varian (1992), Pindyck (1991).

2. On démontre, en outre, sous des hypothèses à peine plus restrictives, que même s'il n'existe pas de probabilité objectivement connue pour les événements possibles, l'individu doit appliquer ce critère dans lequel les p_i sont des probabilités subjectives.

est ici pertinente puisque l'utilité est définie à une fonction affine. On doit alors se poser la question de la concavité de la courbe qui relie l'utilité au gain monétaire. Dans le cas figuré ci-contre d'une fonction concave, on voit que l'utilité d'une loterie qui donnerait une chance sur deux aux gains g_1 et g_2 serait, d'après la règle précédente :

FIGURE 7-8
L'aversion au risque



$$\bar{u} = (1/2)(u_1 + u_2)$$

\bar{u} est l'utilité d'un gain certain \bar{g} vérifiant :

$$\bar{u} = u(\bar{g})$$

Sur la figure ce gain est inférieur au gain moyen : $1/2(g_1 + g_2)$. On dit que dans ce cas, l'agent présente une aversion au risque, liée à la concavité de la courbe représentative de la fonction $u(g)$. Si cette fonction est linéaire, on dit que l'agent est neutre au risque. On mesure l'aversion au risque, non pas par la dérivée seconde u'' , car cette mesure changerait par une transformation affine, mais par le rapport :

$$r = -u''/u'$$

On démontre que devant une loterie de moyenne nulle et d'écart-type (petit) σ , la somme π que l'individu serait prêt à payer pour éviter de s'y trouver confronté serait :

$$\pi \approx - (u''/u') \frac{\sigma^2}{2}$$

Les modèles introduisant l'aversion au risque d'un agent utilisent souvent la fonction :

$$u(g) = -\exp(-rg)$$

Ces quantités peuvent être rapprochées du surplus de Dupuit. Prenons le cas où le revenu ne change pas, seul un prix se modifiant. Alors *EV* peut s'écrire :

$$EV = E(p^0, v^1) - r^0 - E(p^1, v^1) + r^1$$

$$= E(p^0, v^1) - E(p^1, v^1) = - \int_{p^0}^{p^1} \frac{\partial E}{\partial p}(p, v^1) dp$$

comme :

$$\frac{\partial E}{\partial p} = q^*(p, v)$$

$$EV = \int_{p^1}^{p^0} q^*(p, v^1) dp$$

et de même :

$$CV = \int_{p^1}^{p^0} q^*(p, v_0) dp$$

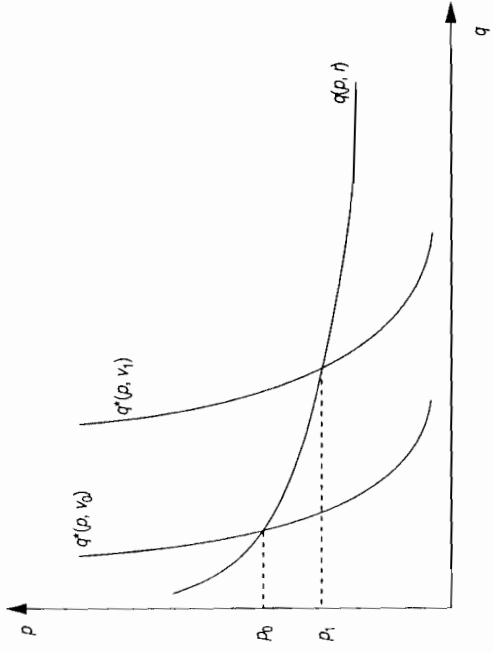
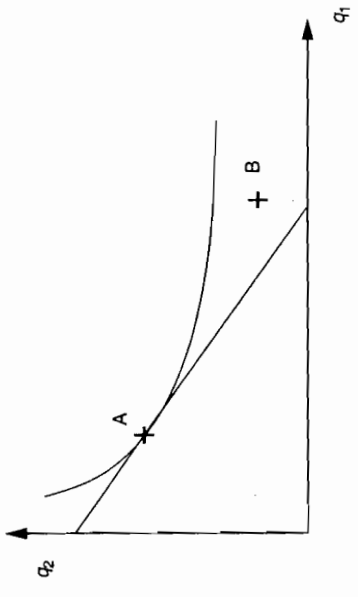
Ainsi *EV* et *CV* sont représentées par des intégrales de la fonction de demande compensée $q^*(p, v)$, alors que le surplus de Dupuit *SD* est représenté par une intégrale de la fonction de demande marshallienne $q(p, r)$.

Si le (ou les) bien(s) considérés sont normaux ($\partial q / \partial r \geq 0$), les fonctions de demande sont l'une par rapport à l'autre dans les positions indiquées sur la figure jointe, comme le fait voir l'équation de Slutsky :

$$\frac{\partial q}{\partial p} = \frac{\partial q^*}{\partial p} + \frac{\partial q}{\partial r} q$$

On voit que *EV* et *CV* encadrent le surplus de Dupuit.

Notons enfin que *EV*, en tant que fonction monotone croissante de v , est en soit une fonction d'utilité, mais que la variation de revenu à prix constants n'est pas un bon indicateur de l'utilité, comme le montre l'exemple classique figuré ci-dessous, où l'individu consomme 2 biens 1 et 2, et où l'on compare deux états A et B. Le panier B ne peut être obtenu aux prix de A qu'avec un revenu supérieur à celui correspondant à A, mais présente une utilité moindre que A.



A.5.2. Aggrégation des utilités individuelles

Le critère admis est le critère de compensation : les gagnants ont dû quoi compenser les perdants. Sa mise en œuvre rigoureuse est très délicate dans le cas de transformations finies. Il connaît deux expressions :

- Le test de Kaldor : le changement proposé est accepté s'il est possible de redistribuer les biens dans la situation finale de telle façon que la situation après redistribution soit Pareto-supérieure (préférée à la situation initiale pour au moins une personne, et équivalente pour les autres).
- Le test de Hicks : le changement est accepté si on ne peut pas redistribuer les biens de la situation initiale de telle façon que la situation initiale ainsi modifiée soit Pareto-supérieure à la situation finale.

À noter que ces règles ne sont pas transitives. Il y a des cas où, pour chacun des deux tests, la situation initiale est préférable à la finale et réciproquement (paradoxe de Scitowski).

On démontre que lorsque l'on n'est pas dans ce cas, alors les deux tests sont équivalents, et que la positivité de la somme des variations équivalentes sur l'ensemble des individus est une condition suffisante pour la satisfaction des tests ci-dessus (voir Morisugi, 1983 ou Kanemoto et Méret, 1985 pour une présentation adaptée au cas particulier des transports).

A.5.3. Cas de fonction d'utilité quasi-linéaire

Une situation particulière existe où les propriétés des transformations marginales restent valables pour les transformations non marginales, c'est celle de fonction d'utilité directe quasi linéaire. Dans l'exemple d'une économie à trois biens, elle s'écrit :

$$U = x_0 + f(x_1, x_2)$$

Alors l'utilité d'une loterie où l'issue est une variable aléatoire tirée d'une loi normale de moyenne \bar{g} et d'écart-type σ sera :

$$u(L) = \bar{g} - (1/2) r \sigma^2$$

A.6.2. Règle de choix en présence d'un risque réductible

Considérons une situation où la somme des bénéfices futurs actualisés à l'instant t est une variable aléatoire qui suit une loi de mouvement brownien :

$$dV = \alpha V dt + \sigma V dz$$

relation dans laquelle α et σ sont des paramètres, et dz est une variable aléatoire suivant un processus de Wiener :

$$dz = \varepsilon(t) (dt)^{1/2}$$

$\varepsilon(t)$ étant une variable normale réduite vérifiant :

$$E[\varepsilon(t) \varepsilon(t')] = 0$$

Pour obtenir la décision optimale, écrivons d'abord la valeur de l'option d'investissement à l'instant t . C'est :

$$F(V) = \max_{T \geq t} E[(V_T - I) \exp(-\mu T)]$$

L'expression de dF , variable aléatoire par l'intermédiaire de F , est donnée par le lemme d'Ito dans le cas d'un processus stationnaire :

$$dF = \frac{\partial F}{\partial V} dV + (1/2) \frac{\partial^2 F}{\partial V^2} (dV)^2$$

ou encore, compte tenu de l'expression de dV

$$dF = \alpha V \frac{\partial F}{\partial V} dt + (1/2) \frac{\partial^2 F}{\partial V^2} \sigma^2 V^2 dt + \sigma V \frac{\partial F}{\partial z} dz$$

D'où :

$$E(dF) = \alpha V \frac{\partial F}{\partial V} dt + (1/2) \sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} dt$$

et l'équation différentielle à laquelle obéit F est :

$$(1/2) \sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} + \alpha V \frac{\partial F}{\partial V} - \mu F = 0$$

La solution générale de cette équation est :

$$F(V) = a_1 V^{\beta_1} + a_2 V^{\beta_2}$$

les valeurs (β_1, β_2) étant racines de l'équation

$$(1/2) \sigma^2 \beta(\beta - 1) + \alpha \beta - \mu = 0$$

comme on doit avoir $F(0) = 0$ seule la racine positive est valable, soit

$$\beta = (1/\sigma^2) \left[-\alpha + (\sigma^2/2) + \sqrt{\left(\alpha - \frac{\sigma^2}{2} \right)^2 + 2\mu\sigma^2} \right]$$

L'investissement doit être réalisé lorsque V^* est tel que :

$$F(V^*) = V^* - I$$

ce qui exprime qu'il est alors indifférent d'investir ou d'attendre.

Enfin, une caractéristique de F au voisinage de cette valeur est que l'on doit avoir :

$$\frac{\partial F}{\partial V}(V^*) = 1$$

condition en l'absence de laquelle une autre valeur d'exercice de l'option serait possible et meilleure, ou encore qui exprime que la fonction $f(V)$ est dans la famille des fonctions possibles, celle qui fournit les valeurs les plus élevées. L'ensemble de ces conditions permet de déduire que :

$$V^* = \frac{\beta I}{\beta - 1}$$

et

$$a = (V^* - I) (V^*)^{-\beta}$$

En revenant à l'expression de β , on voit qu'il est supérieur à l'unité, égal à μ/α en l'absence d'incertitude et égal à l'unité si l'incertitude est infini. On voit alors que V^* , qui prend la valeur $(\mu I)/(\mu - \alpha)$ en l'absence d'incertitude ($\sigma = 0$), est d'autant plus élevé que cette incertitude est grande et devient infini avec cette incertitude. Le règle de décision obtenue en l'absence d'incertitude est, on le vérifie aisément, la même que celle donnée précédemment (taux de rentabilité immédiate = taux d'actualisation) dans le cas où les avantages croissent géométriquement au taux x . L'incertitude peut modifier assez sensiblement cette règle.

Prenons par exemple :

$$\begin{array}{ll} \mu = 8 \% & \text{(taux d'actualisation en vigueur)} \\ \alpha = 3 \% & \text{(taux de croissance moyen des trafics)} \end{array}$$

Alors $\beta(\beta - 1)$ prend les valeurs

$$\begin{array}{ll} \mu/(\mu - \alpha) = 1,6 & \text{pour } \sigma = 0 \\ \text{et } 2,3 & \text{pour } \sigma = 0,2 \end{array}$$

1. Une valeur de l'ordre de grandeur des écarts-types des taux de croissance des trafics ferroviaires et routier sur longue période.

Le seuil de rentabilité est alors d'à peu près 50 % supérieur à celui qui serait justifié en l'absence d'incertitude. Le paradoxe de cet effet est que, conduisant à reculer les décisions (pour atteindre une rentabilité supérieure), il aboutit à des investissements qui, devant répondre à un trafic supérieur, sont également plus capitalistiques : là où en avenir certain on aurait fait dès maintenant une route à 4 voies, on fera dans 10 ans une autoroute (Quinet, 1974).

Annexe 7. Éléments de théorie du duopole¹

Soient deux agents $i = 1, 2$, cherchant chacun à maximiser son profit. Le profit de chaque agent i dépend du choix qu'il fait de la variable stratégique x_i qui est à sa disposition et du choix que fait l'autre de la variable stratégique x_j qui est à sa disposition :

$$\begin{aligned}\Pi_1 &= \Pi_1(x_1, x_2) \\ \Pi_2 &= \Pi_2(x_1, x_2)\end{aligned}$$

Un équilibre de NASH est une situation telle que si on s'y trouve, aucun agent ne regrette sa décision :

$$\begin{aligned}\Pi_1(x_1^*, x_2^*) &= \max_{x_1} \Pi_1(x_1, x_2^*) \\ \Pi_2(x_1^*, x_2^*) &= \max_{x_2} \Pi_2(x_1^*, x_2)\end{aligned}$$

les deux relations précédentes correspondent à un équilibre en stratégies pures.

On définit aussi un équilibre en stratégies mixtes définies par des distributions de probabilités sur les espaces de stratégies, f_1^* sur x_1 et f_2^* sur x_2 vérifiant :

$$\begin{aligned}\int f_2^*(x_2) \left[\int f_1^*(x_1) \Pi_2(x_1, x_2) dx_1 \right] dx_2 \\ = \max_{f_2} \int f_2(x_2) \left[\int f_1^*(x_1) \Pi_2(x_1, x_2) dx_1 \right] dx_2\end{aligned}$$

et l'égalité similaire pour l'entreprise 1. Comme précédemment l'équilibre est une situation sans regret, mais la présence de probabilités dans la définition des stratégies n'est pas d'une interprétation aisée dans la plupart des applications économiques. L'intérêt de l'intervention des stratégies mixtes tient à ce qu'on peut démontrer qu'un jeu a toujours au moins une solution, soit en stratégie pure soit en stratégie mixte.

Dans le duopole de Bertrand la variable stratégique est constituée par les prix. La présentation la plus simple est la suivante : deux entreprises ($i = 1, 2$) ont chacune un coût marginal de production constant c_i (on sup-

pose que $c_1 \leq c_2$). Elle choisissent un prix, et les quantités qu'elles vendent sont déterminées ainsi :

$$\begin{aligned}q_1 &= D(p_1) & \text{si } p_1 < p_2 \\ & \frac{D(p_1)}{2} & \text{ou } p_1 = p_2 & \text{et } c_1 < c_2 \\ q_1 &= \frac{D(p_1)}{2} & \text{si } p_1 = p_2 & \text{et } c_1 = c_2\end{aligned}$$

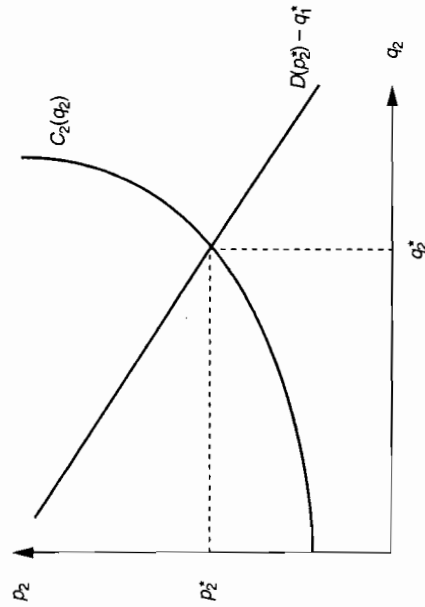
$q_1 = 0$ autrement

On démontre alors que le prix se fixe au niveau de c_2 , et la quantité vendue par chaque entreprise s'en déduit par les relations précédentes.

Le duopole de Bertrand comporte des solutions beaucoup plus compliquées lorsque le coût marginal n'est pas constant. En prenant pour sim- plifier le cas de deux entreprises semblables (figure ci-contre) il est facile de voir que la généralisation de la solution précédente, vérifiant :

$$p_1^* = p_2^* = c_1(q_1^*) = c_2(q_2^*) \quad \text{avec :} \quad q_1^* + q_2^* = D(p_1^*)$$

n'est pas un équilibre de Nash : l'entreprise 1 servant la quantité maximale q_1^* au prix p_1^* (toute quantité supérieure sera vendue à perte) l'entreprise 2 a intérêt à augmenter son prix au-delà du coût marginal puisqu'elle se trouve en situation de monopole vis-à-vis de sa demande : $D(p_2) - q_1^*$.



Plus généralement il faut fixer la manière dont la demande se répartit entre les deux entreprises. L'hypothèse la plus fréquente (celle qui a été implicitement suivie dans l'exemple ci-dessus) est que l'entreprise qui offre le plus bas prix sert les consommateurs dont la disponibilité à payer est la plus élevée, et que l'autre entreprise traite avec la demande résiduelle comme si elle était en situation de monopole.

Enfin, lorsque les deux firmes ont chacun un coût marginal fixe et égal et une capacité limitée et que la somme de leur capacité ne suffit pas à servir la demande avec un prix égal au coût marginal commun, l'équilibre de

1. Voir Tirole (1988), Varian (1992), Dang-Nguyen (1995).

la compétition par les prix est le prix qui permet de satisfaire toute la demande, les deux entreprises travaillant à pleine capacité (Tirole, 1988). Dans le cas général où les coûts marginaux sont variables, les équilibres peuvent impliquer des stratégies mixtes dont la signification économique est l'instabilité du marché.

Dans le duopole de Cournot, les variables stratégiques sont les quantités, et les profits s'écrivent :

$$\Pi_1(q_1, q_2) = q_1 p(q_1 + q_2) - c_1(q_1)$$

et la relation symétrique pour l'entreprise 2.

Si on suppose que chaque entreprise fixe sa quantité indépendamment de l'autre, la solution est donnée par les conditions de maximisation des profits, dont l'écriture est simplifiée en posant : $q_1 + q_2 = Q$

$$p(Q^*) + q_1^* \frac{dp}{dQ} - \frac{dc_1}{dq_1} = 0$$

$$p(Q^*) + q_2^* \frac{dp}{dQ} - \frac{dc_2}{dq_2} = 0$$

Dans le duopole de Stackelberg, la firme 1 est leader et la firme 2 fixe sa quantité en fonction de la quantité mise le marché par la firme 1. Son comportement vise à maximiser le profit, soit :

$$\frac{\partial \Pi_2}{\partial q_2} = p(q_1 + q_2) + q_2 \frac{dp}{d(q_1 + q_2)} - \frac{dc_2}{dq_2} = 0$$

on en tire la courbe de réaction :

$$q_2 = q_2(q_1) \quad (1)$$

Dans une version plus générale, on suppose que chaque entreprise a une courbe de réaction vis-à-vis de l'autre, en fonction de la quantité produite par l'autre :

$$q_2 = q_2(q_1) \quad \text{définissant : } dq_2/dq_1 = v_2$$

$$q_1 = q_1(q_2) \quad \text{définissant : } dq_1/dq_2 = v_1$$

alors les égalités précédentes deviennent :

$$p(Q) + q_1(dp/dQ)(1 + v_2) - (dc_1/dq_1) = 0$$

$$p(Q) + q_2(dp/dQ)(1 + v_1) - (dc_2/dq_2) = 0$$

- $v_1 = v_2 = 0$ donne le duopole de Cournot (chaque firme prend le comportement de l'autre comme donné),

- $v_1 = v_2 = -1$ donne la concurrence parfaite ou le duopole de Bertrand (le prix se fixe au coût marginal),

- $v_1 = 0$,

v_2 défini par la pente de la courbe de réaction (1) donne le duopole de Stackelberg,

- $v_1 = q_1/q_2$,

$v_2 = q_2/q_1$ donne l'équilibre de collusion.

Kreps et Scheinkman (1983) ont établi une équivalence entre l'équilibre de Cournot et l'équilibre d'un jeu dans lequel les deux firmes fixent d'abord leur capacité (à un coût unitaire constant) et ensuite se livrent à une compétition par les prix. En termes imagés :

choix des capacités + Bertrand sur les prix = Cournot

Les hypothèses pour atteindre ce résultat sont toutefois restrictives :

- coûts marginaux constants pour chaque niveau d'équipement
- répartition des consommateurs avec priorité aux plus demandeurs

La formalisation qui précède est relative à un jeu joué une fois pour toute. Dans la réalité les jeux se répètent : ainsi par intervalles, les entreprises redéterminent la valeur de leurs variables stratégiques. Quelle est l'issue des jeux répétés ? Les résultats sont différents selon que le nombre de répétitions est fini ou infini.

Si le nombre de répétition est fini, le résultat reste le même. On le voit facilement par récurrence à rebours : au dernier « round », les joueurs se trouvent dans la situation d'un jeu joué une seule fois, et la solution de NASH s'applique. En remontant le temps, on voit qu'une solution sera composée d'une succession d'équilibres de NASH.

La situation est très différente lorsque le nombre de coups est infini¹. On démontre alors que toute stratégie qui procure plus que le niveau de sécurité (obtenu par la stratégie minimax, celle dont l'issue la plus mauvaise est la plus élevée) est un équilibre de NASH. Notons qu'une stratégie dans un jeu répété est l'annonce du 1^{er} coup et de toutes les actions prises aux coups suivants en fonction des réponses de l'adversaire. C'est donc une définition beaucoup plus complète et complexe que celle qui est en œuvre dans un jeu à un seul coup. Les équilibres sont très nombreux, trop nombreux pour constituer un guide suffisant. C'est le *Folk Theorem* ou théorème de M. tout le monde, en raison de la multiplicité de ses inventeurs.

On peut illustrer ce résultat dans l'exemple simple suivant : deux entreprises 1 et 2 de même coût marginal constituent un duopole de Bertrand dans un jeu où leurs stratégies sont x_1^j et x_2^j , et leurs profits :

$$\Pi_1(x_1^j, x_2^j)$$

$$\Pi_2(x_1^j, x_2^j)$$

1. Il n'y a guère de jeu ayant un nombre infini de coup. Mais on peut penser à un jeu dont la probabilité d'arrêt entre le coup n et le coup $n + 1$ est p . Alors si S_n est l'issue du coup n , l'issue total est

$$S = \sum_{n=1}^{\infty} S_n(1-p)^n$$

Soient Π_1^N et Π_2^N les profits résultants d'un équilibre de NASH, et Π_1^C et Π_2^C ceux résultant d'une collusion, Π_1^M le profit que pourrait s'adjuger 1, s'il cesse de coopérer et joue sa stratégie la plus favorable quand 2 coopère, avec la définition symétrique pour l'entreprise 2 ; et Π_1^R et Π_2^R les profits minimaux que chaque entreprise peut s'assurer. Si le jeu ne se joue qu'une fois, l'équilibre de NASH (éventuellement multiple, ou en stratégies mixtes) aboutit à Π_1^N et Π_2^N . Si le jeu se joue un nombre fini de fois, l'équilibre de NASH du « super jeu » est, par récurrence à rebours, une succession d'équilibres de NASH à chaque période. Si en revanche, le jeu se joue une infinité de fois, les équilibres de NASH dépendent du taux d'actualisation δ . Le profit de 1 pour ce super jeu est :

$$\Pi_1 = \sum_t \delta^t \Pi_{1,t}$$

Considérons la stratégie :

1 joue la coopération (et gagne Π_1^C) tant que 2 fait de même, et si 2 dévie une fois, joue ensuite la stratégie qui limite 2 à Π_2^R (on supposera pour simplifier la notation que cette stratégie assure aussi Π_1^R à 1). Alors si 2 dévie en θ , son profit est :

$$\Pi_2^{\theta} = \sum_{t=1}^{\theta-1} \delta^t \Pi_2^C + \delta^\theta \Pi_2^M + \sum_{t=\theta+1}^{\infty} \delta^t \Pi_2^R$$

la coopération indéfinie rapporterait

$$\Pi_2 = \sum_{t=1}^{\infty} \delta^t \Pi_2^C$$

la comparaison des deux stratégies est celle de

$$\Pi_2^M + \sum_{t=1}^{\infty} \delta^t \Pi_2^R = \Pi_2^M + \frac{\delta}{1-\delta} \Pi_2^R$$

et :

$$\sum_{t=1}^{\infty} \delta^t \Pi_2^C = \frac{\delta}{1-\delta} \Pi_2^C$$

Si δ est assez proche de 1, la collusion est un équilibre ; mais sont aussi des équilibres toutes les stratégies de collusion qui rapportent des profits un peu plus faibles que Π_2^C , et toutes les stratégies de punition qui limitent suffisamment le profit, (par exemple punir fortement quelques coups puis coopérer de nouveau). On en tire deux conclusions : les équilibres sont très nombreux (trop nombreux même), et la coopération et la collusion apparaissent ainsi comme des équilibres possibles.

Par exemple la stratégie qui, pour chaque joueur, consiste à coopérer d'abord, puis à coopérer tant que l'autre coopère et à le punir indéfiniment s'il ne coopère pas est un équilibre de NASH aboutissant à une coopération permanente.

Notons enfin que, lorsque les produits fabriqués par les compétiteurs sont substitués imparfaits, la compétition par les prix a des conséquences adoucies et n'empêche pas des profits pour les deux entreprises. On le voit sur un modèle simple.

Soit deux produits y_1 et y_2 , fabriqués chacun par une firme en situation de concurrence par les prix, qui pratiquent des prix p_1 et p_2 , et dont les coûts marginaux sont nuls.

Supposons que les lois de demande soient du type logit :

$$y_1 = \frac{\exp(-\lambda p_1)}{\exp(-\lambda p_1) + \exp(-\lambda p_2)}$$

et la relation symétrique pour le produit 2.

Les profits sont :

$$\begin{aligned} \Pi_1 &= p_1 y_1 \\ \Pi_2 &= p_2 y_2 \end{aligned}$$

les entreprises 1 et 2 maximise leur profit pour des prix vérifiant :

$$\frac{\partial \Pi_1}{\partial p_1} = \frac{\partial \Pi_2}{\partial p_2} = 0$$

on trouve facilement :

$$\frac{\partial \Pi_i}{\partial p_i} = y_i - p_i \lambda y_i + p_i \lambda y_i^2 = 0$$

la solution est symétrique, et on a :

$$p_1 = p_2 = \frac{1}{2\lambda}$$

Ainsi lorsque les produits sont différenciés, les entreprises font un profit non nul : la concurrence par les prix s'adoucit, et ce d'autant plus que la différenciation est forte. À l'inverse, en cas de non différenciation ($\lambda \rightarrow \infty$) : $p_1 = p_2 = 0$.

BIBLIOGRAPHIE

- G. Dang-Nguyen (1995), *Économie industrielle appliquée*, Vuibert.
M. Fujita (1989), *Urban economic theory*, Cambridge University Press, Mass.
J.V. Henderson (1985), *Economic theory and the cities*, Academic Press, Orlando, Floride.
Y. Kanemoto et K. Mera (1985), « General equilibrium analysis of the benefits of large transportation improvements », *Regional Science and Urban Economics*, 15.
E. Malinvaud (1975), *Leçons de théorie micro-économique*, Dunod, Paris.
H. Morisugi (1983), « A basic definition of transport benefits », *Proceeding of the W.C.T.R.*, Hambourg.
R. Pendyck (1991), « Irreversibility, uncertainty and investment », *Journal of Economic Literature*, Sept.
C. Ponsard (1988), *Analyse économique spatiale*, PUF, Paris.
J. Tirole (1988), *The Theory of industrial organization*, the MIT Press.
H. Varian (1992), *Micro-economic analysis*, Norton International.

Chapitre II. – Transports et localisation	43
1. Transport et localisation des unités économiques.....	43
1.1. Demande prédéterminée en volume et en localisation.....	43
1.2. Constance des rendements et externalités.....	46
2. Transport et développement régional.....	48
2.1. Les transports dans les modèles régionaux.....	48
2.1.1. Un exemple de modèle multirégional transport-économie : Meplan.....	50
2.1.2. Autres modèles.....	51
2.2. Les études empiriques.....	53
2.2.1. Les études françaises.....	53
2.2.2. Études étrangères.....	57
3. Conclusions.....	59
Chapitre III. – Les transports et la ville	63
1. Les transports et les tendances à l'agglomération.....	63
1.1. Localisation de ressources naturelles, indivisibilités.....	64
1.2. Externalités.....	64
1.3. Économies d'échelle.....	65
1.4. L'intérêt pour la diversité.....	66
1.5. Un processus cumulatif d'agglomération.....	68
1.6. Les effets de ségrégation.....	69
2. Les transports, la rente foncière et la localisation.....	70
2.1. Rente foncière et court de transport.....	70
2.2. Les équipements collectifs et les transports de proximité.....	73
3. Les modèles et leur adéquation à la réalité.....	75
4. Politique des transports et politique urbaine.....	77
4.1. La planification urbaine.....	77
4.2. La péri-urbanisation.....	78
4.3. Problèmes de répartition.....	78
4.4. Le fédéralisme fiscal.....	79
DEUXIÈME PARTIE	
La demande et les coûts	
Chapitre IV. – La demande de transport	83
1. Les bases théoriques.....	83
1.1. Le comportement du consommateur.....	83
1.2. Extension aux choix discrets.....	84
1.3. Extension au temps. Le coût généralisé de transport.....	89
1.3.1. Modèle simple.....	89
1.3.2. Modèle d'activité.....	89
1.3.3. Modèle d'opportunité.....	91
1.3.4. Le coût généralisé de transport.....	91
2. L'application aux transports : les modèles de prévision de trafic.....	92
2.1. Les modèles de génération.....	93
2.2. Les modèles de destination.....	93
2.3. Le partage modal.....	93
2.3.1. Modèles agrégés.....	93
2.3.2. Modèles désagrégés.....	93
2.3.3. Le modèle prix-temps.....	93
2.3.4. Les valeurs du temps.....	93
2.4. Le choix d'itinéraires.....	110
2.4.1. La formule d'Abraham.....	111
2.4.2. Le premier principe de Wardrop.....	113
2.4.3. Le deuxième principe de Wardrop.....	115
3. La modélisation en pratique.....	116
3.1. L'information statistique.....	118
3.2. Les logiciels.....	120
3.3. Comparaisons prévisions-réalisations.....	123
3.4. Des voies de progrès.....	123
3.4.1. Les échelles de temps courts et les modèles dynamiques.....	123
3.4.2. Les échelles de temps longues et les modèles transport-urbanisme.....	125
3.4.3. Les transports, demande dérivée et les programmes d'activité.....	126
4. Les modèles de trafic de marchandises.....	127
5. Conclusion.....	128
Chapitre V. – Le coût des transports	135
1. La fonction de coût, rappels.....	137
2. Le coût monétaire pour l'utilisateur.....	138
3. Les dépenses de temps.....	139
3.1. La congestion routière.....	139
3.2. La congestion aérienne.....	141
3.3. La congestion ferroviaire.....	142
3.4. La congestion dans les transports publics urbains.....	143
4. Les dépenses d'infrastructures.....	144
4.1. Infrastructures routières.....	145
4.2. Infrastructures aériennes.....	147
4.2.1. Les aéroports.....	147
4.2.2. Infrastructures de navigation aérienne.....	148
5. Les coûts d'environnement.....	149
5.1. Les méthodes d'évaluation.....	149
5.2. Les résultats d'évaluation.....	152
5.2.1. Bruit.....	152
5.2.2. Pollution locale.....	153
5.2.3. Pollution globale (effet de serre).....	153
6. L'insécurité.....	154
6.1. L'évaluation de la vie humaine.....	155
6.2. Les coûts d'insécurité.....	159

7. Le coût des opérateurs	159
7.1. Les outils d'analyse	159
7.2. Les résultats des évaluations	161
8. Le coût total des transports : résultats numériques	166
8.1. Études américaines	168
8.2. Études européennes	170
8.2.1. Coûts des opérateurs voyageurs	171
8.2.2. Coût des opérateurs Marchandises	174
9. Conclusion	174

TROISIÈME PARTIE L'organisation de l'offre

Chapitre VI. – Nature des marchés et interventions des pouvoirs publics	185
1. Intégration ou fragmentation	186
1.1. Les trois niveaux d'organisation	186
1.2. Coordination hiérarchique ou par le marché	187
2. Entre monopole et concurrence	190
2.1. Le champs du monopole	190
2.2. L'étendue et les modalités de la concurrence	191
3. Entre intervention et libéralisme	193
3.1. Les défaillances du marché	195
3.2. Les missions de service public	197
3.3. Les limites à l'intervention publique	200
4. Conclusion	200

Chapitre VII. – Décisions publiques optimales	203
1. Définition et mesure de l'intérêt collectif	203
1.1. Le calcul des variations d'utilité individuelles	205
1.2. Comment agréger les utilités individuelles ?	206
1.3. Problèmes pratiques de calcul des surplus	207
2. L'exploitation des infrastructures : tarification, réglementation, marché de droit	210
2.1. Règles de gestion optimale	210
2.1.1. Prise en compte des effets externes – Les principes	211
2.1.2. Les pertes de temps	213
2.1.3. L'environnement	219
2.1.4. L'insécurité	221
2.2. Tarification dans un environnement non-optimal	223
2.2.1. Imparfait tarification des modes concurrents	223
2.2.2. Contrainte d'équilibre financier de l'exploitant	224
2.3. La tarification des infrastructures : de la théorie à la réalité	229
3. Les choix d'investissement	231
3.1. Le critère du bénéfice actualisé	231
3.1.1. Le surplus des usagers	232

3.1.2. Les externalités d'environnement	234
3.1.3. Les organismes gérant les transports	234
3.1.4. Les contribuables et l'État	234
3.1.5. Synthèse	235
3.1.6. L'utilisation de ce critère pour la programmation	236
3.1.7. Les conditions de validité du critère du bénéfice actualisé	238
3.1.8. Les lacunes du critère du bénéfice actualisé	242
3.1.9. Le champ de l'analyse multicritère	248
3.2. Les conditions de mise en œuvre des méthodes de choix des investissements	249
3.2.1. La pratique des études de rentabilité	249
3.2.2. Les recommandations des études sont-elles suivies ?	253
3.2.3. Le financement privé des infrastructures	255
4. Conclusion : les décisions publiques dans la pratique	267

Chapitre VIII. – Modalités et effets de la concurrence entre opérateurs

1. Les formes de la concurrence	275
1.1. Le marché du matériel	276
1.2. L'intervention des auxiliaires de transport	278
1.3. Le marché direct entre les opérateurs de transport et les clients	282
1.3.1. Rappel sur les oligopoles	282
1.3.2. Application aux transports	283
1.3.2. Application aux transports	284
2. Les effets de la concurrence	308
2.1. L'intervention des pouvoirs publics dans les marchés concurrentiels	308
2.2. Les expériences de libéralisation et leurs conséquences	309
2.2.1. Les traits communs des dérèglementations	311
2.2.2. La déréglementation du transport aérien aux USA	313
2.2.3. La déréglementation du transport routier de marchandises en Europe et en France	315
2.2.4. La déréglementation des transports par autocar en Grande-Bretagne	318
3. Conclusion	320

Chapitre IX. – Monopole et service public dans les transports	325
1. Les réformes des chemins de fer	325
1.1. En dehors d'Europe	326
1.2. En Europe	328
1.2.1. La directive 91-440 et ses suites	329
1.2.2. Les mises en œuvre selon les pays	330
2. Analyse des expériences. Débats et enjeux autour de la gestion des monopoles	332
2.1. Modalités de fragmentation	334

2.1.1. Fragmentation géographique.....	334
2.1.2. Séparation horizontale	334
2.1.3. Séparation verticale	335
2.2. La concurrence par le marché comme contrôle des monopoles.....	335
2.2.1. La concurrence à la frange et la concurrence par comparaison.....	335
2.2.2. La concurrence potentielle et la résistance des entreprises en place.....	336
2.3. L'accès au réseau.....	337
2.3.1. Les conditions d'accès	338
2.3.2. La tarification de l'accès.....	339
2.3.3. Le cas de l'accès des tiers au réseau.....	340
2.3.4. Considérations pratiques.....	341
2.4. Le contrôle public des monopoles	343
2.4.1. L'organisation du contrôle	343
2.4.2. La régulation des prix.....	345
3. L'exercice du service public.....	348
3.1. La gestion publique.....	348
3.2. La concession	350
3.2.1. Les transports urbains en France	352
3.2.2. Les transports urbains en Grande-Bretagne	355
3.2.3. La gestion portuaire.....	356
3.3. La propriété privée	357
4. Conclusion	360

Chapitre X. - De l'économie à la politique des transports.....	365
1. Réduire l'écart entre les recommandations et la pratique.....	365
2. Une politique des transports est-elle possible ?	368
2.1. Organisation du marché et volonté politique : le cas des transports de marchandises	368
2.2. La gestion des infrastructures : norme ou discrétion ?	369
2.3. Le contrôle de la mobilité ou les difficultés de coordonner la multiplicité	372
3. Conclusion	376

Annexes	379
Annexe 1. Outils de la maximisation sous contrainte	379
Annexe 2. Le mécanisme de la rente foncière	380
A.2.1. Agents identiques.....	383
A.2.2.1. Nombre d'habitants fixé	383
A.2.2.2. Nombre d'habitants variable	383
A.2.2. Agents différents.....	384
Annexe 3. La théorie du comportement du consommateur	386
Annexe 4. La fonction de coût. Rappel théorique.....	389
Annexe 5. Critères de surplus en cas de transformations finies.....	392

A.5.1. Comparaison des utilités d'un ensemble d'individus.....	399
A.5.2. Aggrégation des utilités individuelles.....	400
A.5.3. Cas de fonction d'utilité quasi-linéaire.....	404
Annexe 6. Le traitement de l'incertitude.....	406
A.6.1. Comportement des agents économiques en présence d'un risque.....	411
A.6.2. Règle de choix en présence d'un risque.....	412
Annexe 7. Éléments de théorie du duopole	418

Index