HAL
open science

Réguler les chemins de fer sur une proposition de la nouvelle économie de la réglementation : "la concurrence par comparaison" (yardstick competition)

Julien Lévêque

To cite this version:

Julien Lévêque. Réguler les chemins de fer sur une proposition de la nouvelle économie de la réglementation : "la concurrence par comparaison" (yardstick competition). Economies et finances. Université Lumière - Lyon II, 2005. Français. NNT : . tel-00012043

HAL Id: tel-00012043

https://theses.hal.science/tel-00012043

Submitted on 27 Mar 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Lyon 2 Lumière
Faculté de Sciences économiques et de gestion

RÉGULER LES CHEMINS DE FER SUR UNE PROPOSITION DE LA NOUVELLE ÉCONOMIE DE LA RÉGLEMENTATION : LA « CONCURRENCE PAR COMPARAISON » (*YARDSTICK COMPETITION*)

Thèse pour le doctorat de Sciences Économiques
mention Économie des Transports

présentée, soutenue et obtenue le 5 décembre 2005 par

Julien LÉVÊQUE

avec la mention Très Honorable et les Félicitations du Jury

Directeur de recherche : **M. le Professeur Yves CROZET**

Jury :
Emmanuelle AURIOL (rapporteur), Professeur à l'Université de Toulouse 1
Alain BONNAFOUS, (président), Professeur à l'Université de Lyon 2
Yves CROZET, (directeur), Professeur à l'Université de Lyon 2
Antonio ESTACHE, *Senior Economic Advisor* à la Banque Mondiale
Patrick PERSUY, Directeur financier de Réseau Ferré de France
Saïd SOUAM (rapporteur), Professeur à l'Université de Tours

Avant-propos

> « La SNCF est un cas douloureux, et même les plus ardents défenseurs des entreprises publiques sont désarçonnés... Le gaspillage atteint un seuil difficile à croire et il y a vraiment une appropriation de l'entreprise par ses employés. Sans aller jusqu'à la privatisation, je pense qu'il faudrait accentuer la séparation entre le réseau ferré (RFF) et le transport, et introduire de la concurrence sur le transport. Il faudrait organiser la concurrence en décentralisant et en poussant à la comparaison d'une région à l'autre. »
>
> Jean-Jacques Laffont (1999),
> interview à *l'Expansion*, n° 603.

Élève-ingénieur à l'École Centrale de Lyon, les problèmes de dynamique ferroviaire tels que le contact roue-rail ou l'interaction caténaire-pantographe m'étaient apparus parmi les plus subtils et les plus beaux de la mécanique des solides. Préjugeant de la bienveillance des agents publics, et admirant certaines performances techniques du monopole ferroviaire, la dé-intégration entre le gestionnaire d'infrastructure et l'entreprise ferroviaire, pronée par les économistes, ne pouvaient alors que me heurter. J'ai donc commencé à douter : et si les problèmes d'une industrie aussi technique que les chemins de fer ne pouvaient être résolus par les ingénieurs ? En effet, comment ignorer que les véritables enjeux du transport ferroviaire trouveraient une amorce de solution du côté des incitations et de la régulation économique, plutôt que dans la résolution des problèmes d'instabilités mécaniques et de vibrations des systèmes ?

Le 3e cycle que j'ai mené au sein du Laboratoire d'Économie des Transports, dont cette thèse marque l'aboutissement, m'a offert l'opportunité de saisir l'essentiel du jeu des acteurs du système ferroviaire. A l'issue de ce travail de recherche, j'ai l'intense sentiment d'être moins naïf quant aux motivations qui guident le comportement des agents, quant à leurs intérêts respectifs, leurs moyens d'action et l'information dont ils disposent. Dès lors, je pense avoir assimilé quelques points élémentaires de l'analyse microéconomique.

Cette thèse a été financée par le CNRS, ce qui m'a offert la plus grande liberté pour la mener à bien. Après trois ans de recherches passionnantes, et malgré l'intimité que j'ai pu entretenir avec le sujet, je pense avoir exercé un effort de neutralité suffisant, afin de préserver le caractère scientifique de ce travail doctoral. Cette thèse doit d'ailleurs beaucoup aux remarques et commentaires enrichissants de ceux aux-

quels j'ai présenté mes résultats. De ce fait, elle contribue à restituer, non pas *la* réalité, mais bien un « réel variable des étapes de la connaissance ».

Bien que ce travail demeure individuel, il doit beaucoup à tous ceux qui m'ont aidé à le mener à son terme. Je souhaite ainsi renouveler mes remerciements à Yves Crozet pour avoir proposé et encadré cette thèse. J'exprime également ma gratitude à Alain Bonnafous pour ses conseils et son aide, ainsi qu'aux membres du jury qui ont accepté de s'investir en vue de la soutenance. Parmi les collègues du LET qui m'ont aidé dans ce travail, je tiens à remercier tout particulièrement Dominique Bouf pour le plaisir que j'ai eu à travailler avec lui, et Iragaël pour l'animation motivante qu'il a introduite dans le bureau. Je suis également redevable envers Philippe Adam et les cheminots qui ont contribué par leur apport à ce travail. Enfin, je remercie chaleureusement mes relecteurs familiaux, Odette et Sébastien.

Table des matières

Avant-propos **3**

Introduction générale **11**

I Réguler les chemins de fer **15**

Introduction : la nouvelle économie des chemins de fer **17**

1 Analyse économique des marchés ferroviaires **19**
1.1 Les monopoles naturels 19
1.1.1 Quel(s) monopole(s) naturel(s) sur les infrastructures ? 21
1.1.2 Quelles économies d'échelle sur les services ferroviaires ? . . . 24
1.2 L'émission d'externalités 35
1.2.1 Les externalités positives de réseau 35
1.2.2 Les effets externes 38
1.3 Les propriétés des biens collectifs 40
1.3.1 La non-rivalité 41
1.3.2 La non-excludabilité 41

2 Quelle réglementation pour les chemins de fer ? **43**
2.1 L'apport historique 43
2.1.1 L'intégration verticale 44
2.1.2 L'intégration horizontale 44
2.2 Les limites du monopole public intégré 49
2.2.1 La bienveillance des autorités de tutelle 49
2.2.2 L'absence de risque 52
2.2.3 L'excédent organisationnel 53
2.3 La libéralisation des marchés ferroviaires 57
2.3.1 La concurrence intermodale 57
2.3.2 La concurrence pour le marché 62
2.3.3 La concurrence sur le marché 67
2.3.4 La concurrence potentielle 70

Conclusion : Comment réguler de façon incitative ? 75

II Analyse théorique de la concurrence par comparaison 77

Introduction : Genèse de la concurrence par comparaison 79

3 Analyse unidimensionnelle : les contraintes informationnelles 81
3.1 L'émergence des pratiques comparatives dans un cadre d'aléa moral . 82
3.1.1 Les modèles d'aléa moral 82
3.1.2 Le modèle d'Holmström (1982) 85
3.2 Le cas particulier des tournois de rang 89
3.2.1 Principe général 89
3.2.2 Quelques résultats supplémentaires 90
3.3 Le modèle originel de Shleifer et le problème de l'hétérogénéité . . . 91
3.3.1 Le modèle de Shleifer (1985) 91
3.3.2 Corriger l'hétérogénéité 94
3.4 L'apport des comparaisons dans un cadre de sélection adverse 96
3.4.1 Les modèles de sélection adverse 96
3.4.2 Le modèle de Crémer & McLean (1988) 98
3.4.3 Le modèle d'Auriol & Laffont (1992) 99
3.5 Concurrence par comparaison en aléa moral et sélection adverse . . . 104
3.5.1 Les modèles en aléa moral et sélection adverse 104
3.5.2 Le modèle d'Auriol (2000) 106
3.5.3 Le modèle de Boyer & Laffont (2003) 108

4 Analyses bidimensionnelles 111
4.1 Les comportements d'entente explicite 111
4.1.1 La menace de collusion 112
4.1.2 La capture du régulateur 116
4.2 Les effets dynamiques 118
4.2.1 La collusion tacite 118
4.2.2 L'effet de cliquet 120
4.3 Les incitations à investir 124
4.3.1 Les investissements de productivité 124
4.3.2 Les investissements de qualité 127

Conclusion : Des progrès théoriques favorables à la mise en œuvre 131

III La concurrence par comparaison appliquée aux chemins de fer : justification et résultats 135

Introduction : Les mises en œuvre de mécanismes comparatifs 137

5 La concurrence par comparaison appliquée aux chemins de fer 141
5.1 L'expérience japonaise 141
5.1.1 Technique d'application 143
5.1.2 Principe du mécanisme incitatif 145
5.1.3 Une mise en œuvre pragmatique 147
5.2 Une proposition d'application aux TER français 147
5.2.1 Quelques options pour une application en France 148
5.2.2 Perspective d'application régionale 149
5.2.3 Applicabilité des modèles théoriques 152
5.2.4 Mesure de la performance des opérateurs ferroviaires 155

6 Estimation de la performance par frontière 163
6.1 Le modèle économique 163
6.1.1 Structure du modèle complet 163
6.1.2 Correction de l'hétérogénéité 169
6.2 Le modèle économétrique 171
6.2.1 La frontière *stochastique* de coût 171
6.2.2 Intégration des variables environnementales 174
6.3 Les résultats 175
6.3.1 Résultats économétriques 175
6.3.2 Interprétation économique 180

Conclusion : De la théorie à la pratique 185

Conclusion générale :
Réguler les TER par comparaison : mode d'emploi 187

Bibliographie 193

Annexes 202

A Rappels de méthodologie en économétrie des données de panel 205

B Programmes d'estimation des fonctions de coût 209
B.1 Données régionales 209
B.2 Données sous-réseaux 212

C Résultats estimés des rendements d'échelle et de densité 217
C.1 Résultats régionaux 217
C.2 Résultats des sous-réseaux 221

D Chronologies des régimes concurrentiels 225

E Formulation et résolution du modèle économique complet 229
E.1 Forme et spécification des fonctions 229
E.2 Résolution du problème de maximisation 231

F Estimation des élasticités du coût aux facteurs d'hétérogénéité 233

G Programme d'estimation de la frontière de coût 235

H Autres classements des opérateurs régionaux 241
H.1 Classement suivant l'efficience brute 241
H.2 Classement suivant les coûts unitaires 242

Table des figures

1.1 Nomogramme des rendements de taille pour les sous-réseaux 33
1.2 Nomogramme détaillé des rendements de taille pour les sous-réseaux 34

2.1 Évolution de paramètres financiers du système ferroviaire français . . 51
2.2 Évolution de la production par agent en Europe 54
2.3 Évolution des charges de personnel et des effectifs à la SNCF 55
2.4 Évolution de la répartition modale du trafic intérieur de marchandises 58
2.5 Évolution de la répartition modale du trafic intérieur de voyageurs . 59

3.1 Variation de l'apport des comparaisons selon le rapport $\mathrm{var}(\eta)/\mathrm{var}(\epsilon)$ 87
3.2 Distribution des caractéristiques des entreprises (Auriol et Laffont) . 101
3.3 Relâchement des distorsions de productions grâce aux comparaisons 102
3.4 Distribution des caractéristiques des entreprises (Auriol) 103

5.1 Compte de résultat reconstitué des entreprises ferroviaires japonaises 143
5.2 Mécanisme d'incitation statique 145
5.3 Illustration de l'efficience coût 157

6.1 Variation du coût en fonction de la production 166
6.2 Structure du modèle 168
6.3 Frontière stochastique de coût : décomposition du terme résiduel . . 173
6.4 Frontière stochastique de coût estimée 177
6.5 Variation des scores d'efficience nette normés 182

CG.1 Déroulement chronologique 189

D.1 Intégration horizontale en France, Grande-Bretagne et Allemagne . . 226
D.2 Intégration horizontale en Hollande, Belgique et Suède 227
D.3 Intégration horizontale en Italie, Espagne et Suisse 228

Liste des tableaux

1.1 Statistiques descriptives régionales 27
1.2 Estimations des fonctions de coût régionales 30
1.3 Statistiques descriptives des sous-réseaux 31
1.4 Estimations de la fonction de coût des sous-réseaux 32

C2.1 Récapitulatif des modèles analysés 133

I3.1 Principaux mécanismes comparatifs. 138

5.1 Variables mobilisées pour calculer le coût de référence 144

6.1 Prix moyen des circulations en fonction de variables d'hétérogénéité . 170
6.2 Statistiques descriptives régionales 175
6.3 Estimation de la frontière stochastique 176
6.4 Scores d'efficience nette normés pour 1998 179
6.5 Comparaison de trois mesures normées de la performance pour 1998 180

CG.1 Gains de productivité annuels selon le score d'efficience nette normé 191

B.1 Rendements d'échelle et de densité des réseaux régionaux 217
B.2 Rendements d'échelle et de densité des sous-réseaux 221

H.1 Scores d'efficience brute normés pour 1998 241
H.2 Coûts unitaires normés pour 1998 242

Introduction générale

> « Le système étatique et administratif français repose sur une conception idéaliste du pouvoir politique et de la vie démocratique, sur un postulat général de bienveillance des hommes politiques, de l'administration et de tous les fonctionnaires et personnels assimilés. Cette vision optimiste de l'appareil d'État a conduit à des structures qui ont fonctionné avec un certain succès pour l'essentiel du XXe siècle. »
>
> Jean-Jacques Laffont (1999),
> « Étapes vers un État moderne : une analyse économique »,
> *rapport au Conseil d'Analyse Économique.*

L'intervention de la puissance publique au sein des économies de marché perdure au delà des traditionnelles missions régaliennes, tout en se renouvelant. Ceci est particulièrement avéré avec les industries de réseau (transports, communications, énergie, *etc.*) dont les marchés sont caractérisés par trois défaillances : des rendements d'échelle qui conduisent au monopole naturel, des externalités et des propriétés de biens collectifs. De ce fait, les marchés seuls ne peuvent prétendre à une allocation optimale des ressources : le pouvoir de marché, les distorsions de consommation et les désincitations à investir qui en résulteraient s'y opposent. Pour cette raison, les États établissent la *réglementation* de ces industries, c'est-à-dire les règles légales qui ordonnent et définissent le fonctionnement des marchés associés.

En Europe, à l'issue de la Seconde Guerre mondiale, cette réglementation a largement institué des organisations monopolistiques publiques, comme solution aux défaillances des marchés. En supposant ces organisations bienveillantes, les décideurs publics se sont dispensés de les *réguler* de façon incitative, c'est-à-dire d'intervenir finement (par voie de contrats, d'arbitrages, d'avis, d'autorisations ou d'évaluations) afin d'assurer un meilleur fonctionnement des marchés dans le cadre de la réglementation.

Cette défaillance des pouvoirs publics a conduit un peu partout au développement d'inefficiences qui ont conduit nombre d'entreprises, ferroviaires notamment, à un état de faillite virtuelle. Depuis, cette réglementation, que les États n'ont pas su accompagner d'une régulation incitative, est progressivement réorientée vers les marchés dans le cadre de la libéralisation. Il s'agit non pas de « déréglementer », mais bien de « re-réglementer », de façon compatible avec l'ouverture des marchés. Cette démarche conduit alors à des interventions plus efficaces des pouvoirs publics,

désormais contraints de réguler leurs industries de réseau.

Mais, les mécanismes de régulation incitatifs et adéquats sont rares, tant les contraintes, notamment informationnelles, sont fortes. En effet, contrairement aux entreprises, le régulateur ignore les informations relatives aux coûts, aux aléas ou à la demande, alors même qu'elles sont nécessaires à une régulation optimale des marchés. C'est à partir de ce constat que la nouvelle économie de la réglementation propose des mécanismes incitatifs, parmi lesquels la concurrence par comparaison.

Ce mécanisme repose sur le fait que lorsque plusieurs entreprises opèrent sur des marchés monopolistiques semblables et géographiquement distincts, elles dégagent des externalités informationnelles. Le régulateur peut alors en tirer parti grâce à des comparaisons inter-entreprises. Les modèles théoriques montrent qu'un mécanisme incitatif permet d'extraire l'information stratégique, et induit des comportements concurrentiels. En pratique, un travail de collecte et de traitement des données relatives à la production et aux coûts permet d'évaluer les performances relatives des entreprises. Ces comparaisons permettent au régulateur de réduire l'asymétrie informationnelle, et donc d'accroître sa capacité d'expertise et d'améliorer la régulation.

Littéralement, la « *yardstick competition* » proposée par Shleifer correspondrait à une « concurrence à l'étalon », que la plupart des auteurs français traduisent en termes de « concurrence par comparaison ». Nous verrons toutefois que ce concept recouvre de nombreuses formes (théoriques et pratiques) de mécanismes comparatifs. Dans les faits, la mise en œuvre de ces mécanismes s'apparente bien plus à un outil stratégique d'amélioration de la régulation qu'à une forme de mise en concurrence. Ainsi, il paraîtrait plus pertinent de parler de « régulation par comparaison », afin de suivre la littérature anglophone récente qui traite, à juste titre, de « *yardstick regulation* ».
Pour autant, la concurrence par comparaison ne doit pas être confondue avec le *benchmarking* [1], ne procédant pas du même esprit. D'abord parce que la concurrence par comparaison est conduite par le régulateur, et non par l'entreprise, de sorte qu'elle s'impose à celle-ci de façon plus contraignante. Ensuite parce que la concurrence par comparaison, contrairement au *benchmarking*, se limite à l'évaluation relative des performances ; au-delà, l'identification des points faibles de l'entreprise et la mise en œuvre de solutions appropriées relèvent d'une autre initiative, interne à l'entreprise.

C'est aux chemins de fer que nous proposons d'appliquer la concurrence par comparaison. Dans cette industrie, plus que les autres, la mise en œuvre d'une régulation incitative portant sur les coûts s'avère nécessaire. C'est la raison pour laquelle nous focalisons l'analyse sur l'évaluation économique de l'efficience-coût. Cette démarche nous conduit à mobiliser plusieurs fonctions de coût, microéconomiques et économétriques, afin de modéliser la structure des coûts, de façon à estimer les rendements d'échelle et l'inefficience.
Nous montrons notamment que la seule inefficience interne à la SNCF dans le trans-

[1] Le *benchmarking* (ou « parangonnage ») est la démarche de recherche des meilleures pratiques, d'évaluation par comparaison, et de mise en œuvre d'approches similaires visant à optimiser la performance d'une entreprise ou d'une entité publique.

port régional de voyageurs représente un surcoût de 10 %, supérieur à 100 M€.
Cette recherche illustre aussi la transition de l'économie publique à la nouvelle économie de la réglementation, en mettant l'accent sur les innovations introduites par cette dernière. Elle met également en valeur l'interaction permanente et féconde entre l'analyse économique théorique et les mises en œuvre pratiques qui en résultent.
Ce travail est organisé en suivant le « tiercé » suivant :

- la problématique est précisée en première partie : il faut réguler (économiquement) les chemins de fer ;
- la donne est ensuite développée : il s'agit des résultats théoriques de la nouvelle économie de la réglementation, relatifs à la concurrence par comparaison ;
- l'analyse est enfin menée : un mécanisme comparatif adéquat est proposé, afin de réguler les trains express régionaux.

Première partie

Réguler les chemins de fer

Introduction : la nouvelle économie des chemins de fer

> « Deux grands types de solutions doivent être explorés. Le développement des marchés, chaque fois que les conditions pour l'existence de marchés concurrentiels sont réunies. L'anonymat des marchés concurrentiels résout en effet, à la fois les problèmes d'incitation et les problèmes de corruption.
>
> Quand ce n'est pas possible, l'organisation de l'État doit être repensée en termes de contre-pouvoirs, de « checks and balances », qui permettent de garantir une autonomie d'action suffisante tout en limitant les effets des groupes d'intérêts, et des procédures de concurence interne doivent être multipliées pour créer des incitations individuelles et éviter la corruption. »
>
> Jean-Jacques Laffont (1999),
> « Étapes vers un État moderne : une analyse économique »,
> *rapport au Conseil d'Analyse Économique.*

Durant la majeure partie du xxe siècle, les États européens ont organisé leurs chemins de fer sous la forme de monopoles publics intégrés. Au delà des performances d'une telle organisation, la connaissance économique des activités ferroviaires a été freinée par les structures mises en place. Les organisations se sont substituées aux marchés. Progressivement, les entités historiques ont conforté leur rôle en laissant croire que toutes leurs activités constituaient un monopole naturel. Malgré les progrès de l'économie publique, des arguments techniques ont préservé les monopoles intégrés, jusqu'aux crises financières des années 1980-90. Depuis, la libéralisation des marchés ferroviaires permet d'entrevoir la sortie de ce « siècle des ténèbres » du chemin de fer. La restauration des marchés, à la place d'organisations sclérosées, ouvre en effet un champ nouveau de réflexions économiques.

Le renouveau organisationnel qui se met lentement en place, manque encore de repères, tant les décennies de monopole intégré ont influencé la donne à l'origine de l'investigation. C'est la raison pour laquelle nous consacrons cette première partie à rétablir certains résultats économiques que les idéologies de tous bords ont tenté de négliger. Parce que, comme l'explique Bonnafous (1989) [5], « la singularité même des faits économiques et tout particulièrement leur historicité rendent nécessaire l'inscription de la donne dans un contexte historique déterminé. Négliger ce contexte

implique une lecture superficielle du discours théorique, porteuse des erreurs les plus grossières. »

Afin d'établir correctement notre problématique de la régulation des chemins de fer, nous consacrons un premier chapitre à l'analyse économique des marchés ferroviaires. Il s'agit essentiellement de caractériser dans quelle mesure les chemins de fer font partie des industries de réseau. Plus précisément, nous établissons, par une étude très fine des rendements d'échelle, l'existence de multiples monopoles naturels, correspondant à autant de marchés distincts. Ces résultats conduisent à envisager les contours des différents marchés ferroviaires, et à justifier la régulation publique de ceux-ci.

Dans un second chapitre, nous discutons des formes que cette intervention publique est susceptible de prendre. L'analyse est développée à partir des enseignements retirés des diverses expériences historiques. Elle s'enrichit des formes de concurrence que la libéralisation invite à mettre en œuvre. Ayant démontré la faillite du monopole intégré non régulé, nous retenons que la libéralisation ne dispense en rien les pouvoirs publics d'améliorer la régulation des chemins de fer. Cette première partie nous fournit donc des bases saines, avant de poursuivre notre investigation d'une forme opérationnelle de régulation incitative pour les chemins de fer.

Chapitre 1

Analyse économique des marchés ferroviaires

Afin de préciser notre problématique de la régulation des chemins de fer, ce premier chapitre vise à réactualiser certaines caractéristiques économiques des marchés ferroviaires. Nous reconsidérons ainsi la déclinaison ferroviaire des trois *market failures* qui caractérisent les industries de réseau. Il s'agit des rendements croissants, des externalités et des propriétés de non-rivalité et non-excludabilité. Ces défaillances des marchés tiennent en échec l'allocation optimale des ressources rares à travers la « main invisible » d'Adam Smith, et nécessitent ainsi l'intervention des pouvoirs publics.

Au cours du xxe siècle, ceux-ci ont organisé les chemins de fer sous la forme de monopoles publics intégrés. L'organisation au sein de ces entreprises semblait alors résoudre la plupart des défaillances du marché en faisant disparaître ce dernier. Mais l'échec de ces organisations conduit à libéraliser les systèmes ferroviaires en réintroduisant des marchés en leur sein. Cette libéralisation nous invite donc à reconsidérer sous une autre approche les marchés ferroviaires.

Dans ce chapitre, nous nous intéressons plus particulièrement à l'existence de monopoles naturels. Nous démontrons dans la première section qu'ils sont multiples (et non uniques), notamment dans les activités de transport. Dans la deuxième section, nous rappelons que les chemins de fer émettent des externalités, aussi bien positives que négatives. Nous étudions enfin les chemins de fer en tant que bien collectif mixte, avant de conclure à la nécessité de réguler ces marchés ferroviaires.

1.1 Les monopoles naturels

L'analyse des marchés au moyen de fonctions de coût renseigne sur le périmètre d'activité optimal d'une industrie de réseau comme les chemins de fer. Les économies d'échelle ont d'ailleurs été longtemps associées à l'idée de monopole naturel. Elles n'en sont pourtant qu'une condition suffisante (voir Braeutigam (1989) [8]), dans le cas où la structure de coût est supposée exogène à la structure du marché. Gasmi, Laffont

& Sharkey (2002) [26] montrent en effet que les rendements d'échelle « peuvent être compensés par les bénéfices d'un marché duopolistique, en termes de surplus social, de production plus importante et d'incitations plus fortes à l'efficience [1]. »

L'analyse théorique des rendements d'échelle a été fortement développée dans les années 1975 à 1985, période correspondant aux débuts de la libéralisation des industries de réseau aux États-Unis. Il s'agissait alors essentiellement de rendre compte des avantages structurels des compagnies de transport, liés au volume de leur production et à l'étendue de leur réseaux.

Depuis les travaux fondateurs de Caves, Christensen et Tretheway (Caves, Christensen & Tretheway (1984) [12] et Caves, Christensen, Tretheway *et al.* (1985) [13]), l'approche adoptée dans la littérature repose sur une analyse macroéconomique, agrégée, à l'échelle des différents réseaux de transport. À la suite des travaux américains conduits par ces auteurs, les rendements d'échelle des compagnies européennes intégrées de chemins de fer ont été analysés, notamment par Cantos (2001) [10]. Enfin, Ida & Suda (2004) [31] ont récemment étudié les structures de coût des six principales compagnies japonaises. Ces travaux concluent généralement à l'existence d'une certaine forme de rendements d'échelle : les rendements de densité, c'est-à-dire que pour un réseau de services identiques, une hausse du trafic engendre une augmentation proportionnellement moindre des coûts totaux. En revanche, la plupart des études ne permettent pas de prouver qu'il y a un gain, sous forme d'économies d'envergure, à multiplier les différents types de trafic, qu'il s'agisse de voyageurs et de marchandises en Europe, ou de voyageurs à grande vitesse et à vitesse classique au Japon.

Cette approche traditionnelle n'est cependant plus très pertinente en Europe, compte tenu du contexte de libéralisation des marchés ferroviaires. En effet, les réformes engagées ont conduit à rendre les entreprises européennes de transport ferroviaire significativement hétérogènes. Il en résulte que leurs fonctions de coût ne peuvent plus être estimées à partir d'une même spécification agrégée, ceci pour les deux principales raisons suivantes :

- d'une part, l'agrégation verticale des données, correspondant à l'intégration des différentes activités ferroviaires (transport de marchandises, transport de voyageurs, gestion de l'infrastructure), n'est plus possible. En effet, avec la disparition des entreprises ferroviaires intégrées et l'ouverture des marchés, les périmètres de l'exercice de ces activités ne se superposent plus : certains opérateurs historiques ont vendu leur activité de transport de marchandises, d'autres ont dû céder leur activité de gestion de l'infrastructure.
  Rappelons également que cette agrégation verticale, lorsqu'elle est envisageable, soulève le problème de la pondération entre les produits (quelle équivalence en termes de coût entre un voyageur.kilomètre et une tonne.kilomètre ?).
- d'autre part, l'agrégation horizontale des données, correspondant à l'intégration des trafics à l'échelle nationale (transports régionaux, longue distance,

[1] Dans leur article « *The monopoly test reconsidered* », ces auteurs justifient notamment ce résultat par la possibilité de mettre en œuvre la concurrence par comparaison sur les marchés duopolistiques.

grande vitesse), conduit à estimer des fonctions de coût sur la base d'une spécification internationale, alors même que les situations sont très contrastées. Ceci engendre des biais conséquents du fait de l'hétérogénéité des données en général [2], et des disparités comptables, en particulier.

Ainsi, l'actuelle libéralisation des marchés conduit désormais les économistes à recentrer leur analyse sur les marchés qui émergent au sein des anciens monopoles nationaux intégrés. Il semble notamment particulièrement intéressant de tenir compte dans l'étude des rendements d'échelle de la séparation verticale mise en œuvre en Europe entre les activités de gestion de l'infrastructure et celles des services. Si l'infrastructure est toujours considérée comme un unique monopole naturel, la justification des monopoles dans les services de transport est étudiée avec d'autant plus d'acuité que ces activités s'ouvrent actuellement à la concurrence. Notre analyse s'inscrit donc dans un cadre actualisé, cohérent avec la réglementation européenne.

1.1.1 Quel(s) monopole(s) naturel(s) sur les infrastructures ?

Malgré l'émergence croissante des gestionnaires d'infrastructure dans l'industrie ferroviaire européenne, les données manquent pour évaluer finement les rendements d'échelle dans ces activités. À défaut de pouvoir présenter une estimation pertinente de ceux-ci, nous rappelons les arguments « historiques » qui justifient, en termes d'*indivisibilités* d'offre, l'actuelle organisation de ces activités. Il s'agit donc bien d'observer l'état des choses, plus que d'en restituer l'état théorique. Et ce d'autant qu'il a été prouvé dans l'industrie des télécommunications (en théorie - voir Gasmi, Laffont & Sharkey (2002) [26] -, comme en pratique) que les raisonnements fondés sur les indivisibilités conduisaient à des conclusions erronées.

L'indivisibilité des infrastructures de transport est à l'origine de l'idée de monopole naturel qui leur est associée. Il convient cependant de dépasser cette première impression afin de mieux cerner les contours des monopoles naturels. Nous commençons donc par rappeler les principales activités d'un gestionnaire d'infrastructure ferroviaire, avant de discuter dans quelle mesure les unes et les autres sont susceptibles de relever du monopole naturel.

Les activités de gestion d'une infrastructure ferroviaire

La gestion de l'infrastructure ferroviaire fait référence à deux activités principales : la gestion du cycle de vie de l'infrastructure et la gestion de l'offre d'infrastructure. Le cycle de vie d'une infrastructure comporte trois phases : la construction, l'usage (qui nécessite la maintenance) et le démantèlement. La gestion de l'offre d'infrastructure concerne l'allocation des sillons et la gestion opérationnelle des circulations.

[2] Certains trafics - souvent les moins rentables pour l'opérateur historique - ont été captés par des concurrents au profil très différent.

Détail des activités de gestion d'une infrastructure ferroviaire :

- **La construction de l'infrastructure ferroviaire** (ou sa duplication, sa réhabilitation) fait appel à deux types d'activité :
  - le génie civil qui réalise l'infrastructure au sens propre (terrassements, ouvrages d'art) ;
  - le génie ferroviaire qui réalise la superstructure (voie, ballast, caténaire, signalisation...).
- **La maintenance de l'infrastructure ferroviaire** est elle-même composée de trois types d'activité :
  - la maintenance préventive qui consiste à surveiller (par observations et vérifications) l'état de l'infrastructure, à assurer l'entretien courant (nettoyage, graissage, petits réglages et remplacements) et à intervenir de façon plus ou moins légère (réglages, remises en état, remplacements) ;
  - la maintenance corrective qui peut comporter des opérations de dépannage et de réparation suite à l'observation d'une défaillance ;
  - la maintenance régénératrice est une remise en état des installations qui s'accompagne d'un remplacement significatif d'éléments constitutifs vitaux par des éléments neufs ou remis à neuf. Il peut s'agir de renouvellement, reconstruction ou d'une grosse opération de remise en état.
- **Le démantèlement de l'infrastructure ferroviaire** (dépose des voies) peut avoir lieu après prononciation de la fermeture de la ligne correspondante.
- **L'allocation des sillons** concerne la répartition des capacités d'infrastructure disponibles pour la circulation des trains.
  - définition et étude des capacités ;
  - réception des demandes ;
  - attribution des sillons (réalisation du graphique de circulations) ;
  - contractualisation avec les entreprises ferroviaires.
- **La gestion des circulations** comprend les activités de régulation des circulations ferroviaires :
  - la gestion opérationnelle des circulations ;
  - les mesures propres à assurer leur fluidité, leur régularité et leur acheminement en toute sécurité ;
  - la mise en œuvre des dispositions nécessaires pour assurer le rétablissement de la situation normale en cas de perturbation de la circulation.

La gestion du cycle de vie

Les trois activités constitutives de la gestion du cycle de vie d'une infrastructure (construction, maintenance, demantèlement) sont caractérisées par une indivisibilité technique du fait de la nécessaire continuité du réseau d'infrastructure. Cela conduit à un monopole naturel sur la maîtrise d'ouvrage liée à l'ensemble du réseau interconnecté, opéré par le gestionnaire d'infrastructure. La maîtrise d'œuvre, quant à

elle, peut être divisée à l'échelle des différents marchés indivisibles de génie civil et de génie ferroviaire définis par le maître d'ouvrage.
La question de l'étendue de ces marchés indivisibles se pose concernant les activités de maintenance préventive et correctrice. Il est difficile de déterminer le périmètre optimal de ces deux activités, fonctionnellement liées l'une à l'autre, sans disposer d'une fonction de coût. *A priori*, il ne semble pas que l'indivisibilité technique due à la continuité du réseau d'infrastructure justifie un monopole trop important sur l'ensemble du réseau interconnecté. Il n'est pas interdit d'imaginer plusieurs marchés de maintenance géographiquement distincts sur le réseau classique, ainsi qu'un marché de maintenance du réseau à grande vitesse. Toutefois, seuls les chemins de fer britanniques ont expérimenté, jusqu'en 2004, une telle segmentation avant de réintégrer ces activités nationalement au sein du gestionnaire d'infrastructure.
Enfin, les missions de maîtrise d'ouvrage de ces trois activités de gestion du cycle de vie doivent être intégrées au sein du gestionnaire d'infrastructure en raison des indivisibilités fonctionnelles qui les lient. En effet, les décisions prises pour la construction d'une infrastructure sont lourdes de conséquences, à terme, sur les niveaux de maintenance ultérieurs. De même, les investissements de maintenance régénératrice influent sur le niveau de maintenance préventive et corrective à apporter. Ainsi, compte tenu de la difficulté à établir des contrats incitatifs complets permettant d'en déléguer la gestion, ces marchés restent intégrés, à charge pour le gestionnaire d'infrastructure de minimiser le coût du cycle de vie de l'infrastructure.

La gestion de l'offre d'infrastructure

La gestion des circulations ferroviaires (définition des capacités lors de l'allocation des sillons et suivi depuis les postes de contrôle) fait appel à une gestion monopolistique sur un ensemble de lignes interconnectées en raison de l'indivisibilité technique due à la continuité des circulations. Selon la nature des trafics utilisant l'infrastructure, ces activités peuvent être déléguées localement ou, au contraire, être traitées ou coordonnées en amont, à un niveau national ou supranational. Ainsi, le suivi des circulations sur une ligne à grande vitesse peut être assuré par un gestionnaire d'infrastructure distinct de celui qui suit les circulations sur les lignes classiques interconnectées, à condition que leurs agents puissent s'échanger leurs informations efficacement [3]. L'allocation des sillons est réalisée à l'échelle européenne pour la prise en compte des demandes de transport international, à l'échelle nationale pour les circulations longue distance et à l'échelle régionale pour le transport régional et les trains de service. La compatibilité des demandes est généralement assurée en donnant la priorité aux circulations les plus longues.

Il ressort de cette analyse que les indivisibilités propres aux marchés de la gestion d'infrastructure permettent de déterminer assez aisément l'étendue des monopoles naturels relatifs à ces activités. Nous nous interrogeons à présent sur l'existence, *a priori* moins évidente, de monopoles naturels dans les services de transport.

[3] Voir l'exemple britannique d'interconnexion des réseaux grande vitesse (CTRL), classique (Network Rail) et Eurotunnel.

1.1.2 Quelles économies d'échelle sur les services ferroviaires ?

La justification d'éventuels monopoles naturels dans l'activité de transport requiert d'analyser les rendements d'échelle à un niveau suffisamment fin. En effet, s'il est imaginable que l'optimisation des rotations (du matériel roulant ou des personnels roulants) et la mutualisation de certaines tâches (achats, maintenance, administration) conduisent à certaines économies d'échelle, l'ampleur de celles-ci reste à estimer. *A priori*, l'optimum de production économique n'a aucune raison de coïncider avec les divisions administratives du pays, chargées de réguler les services ferroviaires dont elles ont la charge. Nous procédons donc à une étude actualisée des rendements d'échelle afin de remettre en cause certains préjugés concernant ce sujet.

Précisions théoriques

Les rendements d'échelle sont classiquement définis comme étant l'inverse de l'élasticité apparente du coût à la prodution, estimée à partir d'une fonction de coût. Les rendements d'échelle peuvent toutefois également être calculés à partir des élasticités de la production aux facteurs de production, estimées par une fonction de production. Toutefois, il est plus pertinent de travailler sur une fonction de coût, dès que le volume de production est exogène, ce qui est bien souvent le cas dans les industries de réseau. Dans le transport ferroviaire de voyageurs, la production est évaluée par le trafic des personnes transportées. On la mesure par le nombre de voyageurs.kilomètres.

Dans les industries de réseau, les développements théoriques ont conduit à distinguer les rendements d'échelle selon qu'ils sont dus à une variation de trafic à réseau inchangé, ou à une variation de trafic induite par une modification du réseau. Ainsi, Caves, Christensen & Tretheway (1984) [12], suivis par Pels & Rietveld (2000) [42] et Seabright (2003) [48], considèrent une fonction de coût total $C(y_i, n)$ dont les variables exogènes sont :

- les y_i : les quantités de produit i ;
- n : une variable traduisant l'étendue du réseau de transport.

Ils définissent alors :

- **Les rendements de taille (RTS) :** ils correspondent au sens courant des rendements d'échelle :

$$RTS = \frac{1}{\sum_i \epsilon_{y_i} + \epsilon_n}$$

  où les ϵ_{y_i} et ϵ_n correspondent respectivement aux élasticités du coût aux produits et à celle du coût à la taille du réseau.
  Les rendements de taille sont croissants (ce qui correspond à la présence d'économies d'échelle) lorsque leur valeur est supérieure à 1. Ils sont constants en cas d'égalité et décroissants, sinon. Lorsque ces rendements sont croissants sur un réseau caractérisé par des trafics y_i et une étendue n, il n'y a aucun gain à partitionner ce réseau, toutes choses égales par ailleurs.

- **Les rendements de densité (RTD) :** ils ne font intervenir que l'élasticité du coût au trafic, toutes choses égales par ailleurs, c'est-à-dire notamment à réseau fixé :

$$RTD = \frac{1}{\sum_i \epsilon_{y_i}} > RTS$$

  Pour cette raison, les rendements de densité ne peuvent être valablement estimés qu'à partir de données de panel, c'est-à-dire de séries de données temporelles pour plusieurs réseaux.
  Les rendements de densité sont croissants (ce qui correspond à la présence d'économies d'échelle sur le trafic) lorsque leur valeur est supérieure à 1. Ils sont constants en cas d'égalité et décroissants, sinon. Lorsque ces rendements sont croissants, il y a des surcoûts à multiplier le nombre d'entreprises sur le réseau correspondant, toutes choses égales par ailleurs. Il y a au contraire un gain à massifier les trafics sur ce réseau.

- **Les économies d'envergure :** elles sont susceptibles d'apparaître dans les entreprises multiproduits. Il s'agit du gain qui peut résulter de synergies entre les différentes activités (transport de marchandises et de voyageurs ou bien transport de voyageurs à grande vitesse et régional) :

$$Scope(i,j) = \frac{\partial^2 C(y_i, y_j)}{\partial y_i \partial y_j}$$

  Il y a des économies d'envergure entre les deux produits i et j, lorsque ce terme (égal au coefficient du terme croisé d'une spécification translog) est négatif.

Le cas du transport régional

Nous proposons d'analyser les rendements d'échelle dans le transport régional de voyageurs, à une échelle microéconomique. Il ne s'agit donc pas d'une évaluation supplémentaire des économies d'échelle entre les grands réseaux européens, laquelle n'aurait plus beaucoup de sens, ainsi que nous l'avons rappelé dans l'introduction de cette section. Ce travail prétend au contraire produire des résultats, dans un cadre actualisé, susceptibles d'orienter les choix des décideurs dans l'actuel contexte de libéralisation des marchés ferroviaires.

Données : Les données que nous mobilisons proviennent de la comptabilité de la SNCF pour les services de transport régionaux de voyageurs [4]. À la source, nous disposons des coûts totaux et du trafic, par section de lignes régionales et par région administrative, sur la période 1992-1998 (hormis 1996). Les données par sections de ligne ne reflètent pas de logique de production de service de transport ferroviaire et n'ont, en ce sens, pas d'intérêt à l'état brut. Ainsi, il n'y a aucune pertinence

[4] Il s'agit des résultats fournis par le logiciel comptable FC12K.

à considérer séparemment les coûts sur Strasbourg-Colmar, de ceux sur Colmar-Mulhouse, dans la mesure où une importante part du trafic est opérée par des trains Strasbourg-Colmar-Mulhouse. Une agrégation des données brutes a donc été menée afin de regrouper en sous-réseaux pertinents les différentes sections de lignes. À l'issue de ce travail, nous disposons donc des deux échantillons de données de panel suivants :

- un échantillon de 33 sous-réseaux répartis en Alsace, Bourgogne, Franche-Comté et Rhône-Alpes, qui conduit à 169 observations sur 1992-1998 (l'année 1996 est manquante et le panel est non-cylindré) ;
- un échantillon des 20 régions administratives[5] qui correspond à 120 observations sur la période 1992-1998 (l'année 1996 est manquante).

Les fonctions de coût définies dans la littérature (voir Braeutigam (1999) [9]) résultent de la minimisation du coût sous contrainte de production. Elles intègrent donc en variables explicatives un vecteur des produits $\boldsymbol{y}$ de l'entreprise et un vecteur des prix unitaires des facteurs de production $\boldsymbol{w}$. Dans notre cas, la quasi-identité des prix unitaires des facteurs de production entre les différents services[6] permet de nous affranchir du vecteur de variables explicatives $\boldsymbol{w}$, de toute façon indisponible. Ceci est d'autant plus justifié que notre analyse porte ici exclusivement sur l'évolution de la fonction de coût par rapport aux variables de production ; nous verrons que celles-ci expliquent plus de 94 % de la variance observée. Nous considérons donc les prix des facteurs comme étant fixés à des niveaux prédéterminés, ainsi que le suggère Varian (1992) [52].

Précisons aussi que l'étude des rendements d'échelle nécessite de spécifier une fonction de coût de long terme. Une telle fonction correspond à la faculté qu'ont les entreprises d'ajuster tous leurs facteurs de production. Cette hypothèse est difficilement vérifiable sur un marché ferroviaire. Il y a en effet une suspicion permanente quant à l'adéquation des charges d'infrastructure, avec le maintien durable des performances de celle-ci.

Suivant les recommandations de Caves, Christensen & Tretheway (1984) [12], nous adoptons une fonction de coût qui intègre la dimension du réseau, afin de distinguer économies de taille et économies de densité. Nous initialisons nos estimations avec une fonction de coût de type translog, réputée pour être une forme fonctionnelle flexible :

$$\begin{aligned} \ln C &= \alpha_0 + \beta_{VK} \ln VK + \beta_{LL} \ln LL \\ &\quad + \gamma_{VK} (\ln VK)^2 + \gamma_{LL} (\ln LL)^2 + \delta \ln VK \ln LL \end{aligned} \tag{1.1}$$

où α_0, $\beta_.$, $\gamma_.$ et δ sont les coefficients linéaires,
C correspond au coût total,
VK désigne le trafic exprimé en voyageurs.kilomètres,
LL est la longueur de lignes du réseau exploité.

[5] Hors Corse et Île-de-France dont les transports ferroviaires sont exploités différemment.

[6] Ceci est garanti par la gestion centralisée de la SNCF. Ainsi, à grade et échelon égaux, le coût du travail est le même pour tous les cheminots, à l'exception de la majoration résidentielle, susceptible d'augmenter de 2 à 4 % le traitement suivant la commune d'affectation du poste.

Le traitement de nos données fait appel à l'économétrie des données de panel, qui permet de prendre en compte la double dimension spatio-temporelle de nos échantillons. Nous rappelons brièvement dans les paragraphes qui suivent, les caractéristiques des estimateurs à notre disposition.

Méthodologie économétrique : L'économétrie des données de panel, développée depuis les années 1970, propose de nombreux modèles adaptés aux différents échantillons à analyser. L'objectif de ces modèles consiste à inclure dans les estimateurs, des effets propres aux différents individus [7]. Pour une présentation approfondie des éléments théoriques de l'économétrie des données de panel, voir en annexe page 205 et Sevestre (2002) [49].

Dans notre cas d'analyse, la nature des données (le fait que la variable longueur de lignes, LL, soit constante dans le temps) contraint le choix des modèles et estimateurs à notre disposition. En effet, ni les modèles à effets fixes, ni l'estimateur intra-individuel des modèles à erreur composée ne peuvent être mis en œuvre dans ce cas. En revanche, les Moindres Carrés Quasi-Généralisés (MCQG) permettent d'estimer un modèle à erreur composée avec des variables exogènes constantes.
L'inconvénient majeur de cet estimateur est qu'il ne converge qu'à condition que les régresseurs soient exogènes : les effets spécifiques aléatoires ne doivent pas être corrélés aux autres variables exogènes du modèle [8].

Résultats régionaux : Notre échantillon régional peut être décrit statistiquement par les résultats suivants :

Variable	Moyenne	Écart-type	Minimum	Maximum
Coûts (k€ 2004)	78608	46091	21232	260946
Voyageurs.kilomètres (millions)	348	246	61	1245
Longueur de lignes (km)	1070	326	598	2033

Tab. 1.1 – Statistiques descriptives régionales (120 observations). Source : SNCF.

Nous conduisons nos estimations sous Limdep ; les programmes et les résultats figurent en annexe, page 209. L'estimation de la fonction translog complète (équation 1.1) fournit des coefficients non significativement différents de zéro. Nous estimons donc une fonction réduite dans laquelle les coefficients β_{VK}, β_{LL} et δ sont contraints,

[7] Nous ne présentons ici que les modèles à effets individuels et non temporels, dans la mesure où la présence d'effets temporels dans nos observations n'est pas avérée.

[8] Dans ce cas, la distribution asymptotique de cet estimateur est la même que celle de l'estimateur intra-individuel. Aussi, afin de tester la convergence de l'estimateur des MCQG, Hausman & Taylor (1981) [30] ont proposé un test, dit « de Hausman », qui compare cet estimateur et sa matrice des variances-covariances aux caractéristiques de l'estimateur intra-individuel. Lorsqu'une variable est constante dans le temps, le test d'Hausman porte sur les estimations des coefficients associés aux autres variables.

égaux à zéro. La spécification retenue s'écrit donc :

$$\ln C = \alpha_0 + \gamma_{VK}(\ln VK)^2 + \gamma_{LL}(\ln LL)^2 \tag{1.2}$$

Il s'agit de la spécification qui minimise le critère d'information de Akaike et maximise le R^2, parmi celles dont les coefficients sont significatifs. Un test du ratio de vraisemblance appliqué à cette restriction de trois degrés de liberté, la justifie pleinement :
Il s'agit de tester l'hypothèse nulle :

$$\mathrm{H}_0 : \beta_{VK} = \beta_{LL} = \delta = 0$$

contre l'hypothèse englobante :

$$\mathrm{H}_1 : \beta_{VK}, \beta_{LL}, \delta \neq 0$$

Le ratio de vraisemblance $\mathrm{LR} = -2[\ln \mathrm{L}(\mathrm{H}_0) - \ln \mathrm{L}(\mathrm{H}_1)]$ suit une loi du χ^2 à 3 degrés de liberté. Le calcul donne :

$$\mathrm{LR} = -2(77,4059 - 78,2733) = 1,7348 < 7,81 = \chi^2_{3;5\%}$$

Ainsi, pour un risque de 5 % de rejet à tort de l'hypothèse H_0, le χ^2 tabulé est bien supérieur à la valeur du ratio de vraisemblance ; nous retenons donc l'hypothèse H_0.

Nous vérifions la présence d'effets spécifiques, et donc l'opportunité de traiter les données de panel en tant que telles, par le test du multiplicateur de Lagrange (LM-test), proposé par Breush & Pagan :
Il évalue l'hypothèse nulle :

$$\mathrm{H}_0 : \text{absence d'effets spécifiques}$$

contre l'hypothèse :

$$\mathrm{H}_1 : \text{présence d'effets spécifiques.}$$

Sous H_0, la statistique du multiplicateur de Lagrange suit asymptotiquement une loi du χ^2 à 1 degré de liberté :

$$\mathrm{LM} = \frac{NT}{2(T-1)} \left[\frac{\sum_{n=1}^{N} \left(\sum_{t=1}^{T} \hat{\epsilon}_{n,t} \right)^2}{\sum_{n=1}^{N} \sum_{t=1}^{T} \hat{\epsilon}^2_{n,t}} - 1 \right]^2 \rightsquigarrow \chi^2_1$$

Ce test nous conduit à rejeter l'hypothèse de présence d'effets spécifiques en dimension temporelle. Bien que nos observations puissent être influencées par des effets spécifiques comme la mauvaise conjoncture de 1993 ou les mouvements sociaux de 1995, ceux-ci ne sont pas significatifs. En revanche, le même test permet de conclure à l'existence d'effets spécifiques individuels[9]. Les résultats fournis par Limdep vérifient : $\mathrm{LM} = 179,66 \gg 3,84 = \chi^2_{1;5\%}$.

[9] Le test de Honda, présenté par Sevestre, qui est un peu plus restrictif que celui du multiplicateur de Lagrange est tout aussi concluant.

Enfin, le test d'Hausman est appliqué pour tester l'exogénéité des régresseurs. Il évalue l'hypothèse nulle :

$\mathrm{H}_0 : \boldsymbol{\gamma}_{\text{within}}$ et $\boldsymbol{\gamma}_{\text{MCQG}}$ sont convergents
où $\boldsymbol{\gamma}_{\text{within}}$ correspond au vecteur des estimateurs intra-individuels,
et $\boldsymbol{\gamma}_{\text{MCQG}}$ correspond au vecteur des estimateurs des MCQG.

contre l'hypothèse :

$\mathrm{H}_1 : \boldsymbol{\gamma}_{\text{within}}$ est convergent et $\boldsymbol{\gamma}_{\text{MCQG}}$ est non-convergent

Limdep effectue ce test en comparant l'estimateur des MCQG à l'estimateur intra-individuel appliqué au modèle incluant une variable indépendante du temps : $(\ln LL)^2$. Ces résultats n'ont donc pas de sens ; nous programmons donc séparément ce test (voir page 210), afin de respecter le cadre d'application spécifié par Hausman et Taylor. Il s'agit de calculer la statistique de Hausman, qui suit asymptotiquement une loi du χ^2 à $\dim(\boldsymbol{\gamma}) = 1$ degré de liberté :

$$Q_H = (\hat{\boldsymbol{\gamma}}_{\text{within}} - \hat{\boldsymbol{\gamma}}_{\text{MCQG}})'[\widehat{\boldsymbol{V}}(\hat{\boldsymbol{\gamma}}_{\text{within}}) - \widehat{\boldsymbol{V}}(\hat{\boldsymbol{\gamma}}_{\text{MCQG}})]^{-1}(\hat{\boldsymbol{\gamma}}_{\text{within}} - \hat{\boldsymbol{\gamma}}_{\text{MCQG}}) \rightsquigarrow \chi^2_1$$

où $\widehat{\boldsymbol{V}}()$ désigne la matrice des variances-covariances estimées.
Ce test nous conduit à rejeter l'hypothèse de convergence H_0, puisque l'on trouve : $Q_H = 11,4 > 3,84 = \chi^2_{1;5\%}$. Ainsi, la variable $(\ln VK)^2$ est corrélée aux effets spécifiques, ce qui nous amène à renoncer à l'estimateur des MCQG. Cela nous conduit aussi à renoncer à la méthode des variables instrumentales qui impose qu'une variable, au moins, soit strictement exogène. Dans ce travail, compte tenu de nos données, nous choisissons de nous reporter sur une estimation classique par les MCO, sans prendre en compte les effets individuels. Mais compte tenu du résultat du test d'Hausman, cette estimation par les MCO est également biaisée du fait de la corrélation entre variables exogènes et effets individuels ; cela sera corrigé dans de prochains travaux.

Signalons que nos résultats seront biaisés par le fait que notre fonction de coût, C, théoriquement de long terme, sous-estime gravement les coûts de gestion de l'infrastructure. Aussi, une simulation a été développée avec des coûts d'infrastructure multipliés par 2 sur la période. Cela revient à nous rapprocher de la couverture du coût marginal d'usage de l'infrastructure ferroviaire, telle que l'envisage la tarification du réseau depuis 2004. Nous verrons néanmoins que les résultats concernant cette nouvelle fonction de coût, notée $C_{\text{infra}\times 2}$, ne diffèrent qu'à la marge des autres estimations.

Les résultats présentés dans le tableau 1.2 de la page suivante, obtenus avec les MCO, permettent de calculer pour chacune de nos observations :

– les rendements de taille :

$$RTS = \frac{1}{2(0,02870 \ln VK + 0,01354 \ln LL)}$$

	Observations : 120			
	Variable endogène : $\ln C$		Variable endogène : $\ln C_{\text{infra}\times 2}$	
	MCO		MCO	
Paramètre	Coefficient	t-test	Coefficient	t-test
α_0	12,82	88,90	12,88	90,47
γ_{VK}	0,02870	33,73	0,02863	34,07
γ_{LL}	0,01354	9,975	0,01427	4,243
R^2	0,940		0,942	

TAB. 1.2 – Estimations des fonctions de coût régionales (1.2) par les MCO.

– les rendements de densité :

$$RTD = \frac{1}{2.0,02870 \ln VK}$$

Les résultats des calculs, présentés en annexe page 217, montrent que les rendements de densité sont croissants et concaves, et ce, sans ambiguité : ils varient avec le trafic de 1,580 à 1,241, ce qui correspond à une moyenne de 1,392, cohérente avec les résultats de la littérature [10]. Les rendements de densité calculés montrent clairement que sur tous ces réseaux, il n'y a aucun gain à mettre plusieurs entreprises en concurrence *sur* l'ensemble des lignes de chaque réseau régional. Au contraire, il faut chercher à massifier au maximum le trafic.

Les rendements de taille sont plus faibles, plutôt croissants et concaves : ils varient de 1,216 (Poitou-Charentes 1994) à 0,988 (Rhône-Alpes 1998) pour une moyenne de 1,104 sur notre échantillon. Il apparaît toutefois que ces résultats ne sont pas significativement différents de 1. Au seuil de confiance de 95 %, moins de la moitié de nos observations (les régions à faible trafic) présentent encore des rendements d'échelle croissants (voir en annexe).

Ces résultats peuvent être interprétés ainsi : dans les régions à faible trafic où les rendements de taille sont significativement croissants, il n'y a pas intérêt à segmenter les réseaux régionaux, de ce point de vue [11]. En revanche, l'absence de rendements de taille croissants dans les autres régions, bien qu'étant compatible avec une sous-additivité de la fonction de coût, décrédibilise les arguments économiques en faveur d'un monopole d'envergure régionale. Il semble quasi certain que les réseaux de ces régions ne soient pas des monopoles naturels, au sens précisé par Gasmi, Laffont & Sharkey (2002) [26]. L'introduction de concurrents devrait largement contribuer à

[10] Cantos (2001) [10] obtient des rendements de densité de 1,42 pour la SNCF sur la période 1973-1990 et une moyenne européenne de 1,73. Il considère cependant le nombre de trains.kilomètres (quel que soit le trafic) comme variable de produit.

[11] Toutefois, du point de vue - plus large - de Gasmi, Laffont & Sharkey (2002) [26], il est très vraisemblable que l'optimum collectif consiste dans un premier temps à rendre ces marchés duopolistiques au terme d'un allotissement pertinent. En effet, compte tenu des difficultés qu'ont les nouveaux entrants à pénétrer sur les marchés ferroviaires, il paraît nécessaire dans un premier temps d'allotir les réseaux régionaux pour les rendre contestables.

réduire les coûts de l'opérateur historique. Il paraît donc *a priori* légitime de segmenter les réseaux régionaux actuels ; nous poursuivons donc cette analyse à l'échelle des sous-réseaux que nous avons constitués.

Résultats par sous-réseaux Notre échantillon de sous-réseaux peut être décrit statistiquement par les résultats du tableau 1.3 :

Variable	Moyenne	Écart-type	Minimum	Maximum
Coûts (k€ 2004)	13,3	14,5	0,6	86,8
Voyageurs.kilomètres (millions)	55	73	1,4	393
Longueur de lignes (km)	152	113	8	648

TAB. 1.3 – Statistiques descriptives des sous-réseaux (169 observations). Source : SNCF, RFF.

De même que dans le cas régional, nous conduisons nos estimations sous Limdep et les programmes et résultats figurent en annexe, page 212. À partir de la fonction translog complète (équation 1.1), nous estimons une fonction réduite, dans laquelle le coefficient δ est contraint, égal à zéro. La spécification retenue s'écrit donc :

$$\ln C = \alpha_0 + \beta_{VK} \ln VK + \beta_{LL} \ln LL + \gamma_{VK} (\ln VK)^2 + \gamma_{LL} (\ln LL)^2 \tag{1.3}$$

Il s'agit de la spécification qui minimise le critère d'information de Akaike et maximise le R^2, par rapport à la spécification complète. Un test du ratio de vraisemblance appliqué à cette restriction d'un degré de liberté, la justifie pleinement. En effet, le calcul donne :

$$\text{LR} = -2(12,9501 - 13,0352) = 0,1702 < 3,84 = \chi^2_{1;5\%}$$

Ainsi, pour un risque de 5 % de rejet à tort de l'hypothèse H_0, le χ^2 tabulé est bien supérieur à la valeur du ratio de vraisemblance ; nous retenons donc le modèle contraint.

À l'échelle des sous-réseaux, les tests du multiplicateur de Lagrange nous conduisent à rejeter l'hypothèse de présence d'effets spécifiques en dimension temporelle mais à conclure à l'existence d'effets spécifiques individuels. Les résultats fournis par Limdep vérifient : $\text{LM} = 259,40 \gg 3,84 = \chi^2_{1;5\%}$.

Le test d'Hausman (programme page 213) pour vérifier l'exogénéité des régresseurs est, cette fois, concluant. On calcule en effet $Q_H = 2,20 < 5,99 = \chi^2_{2;5\%}$, ce qui nous conduit à estimer un modèle à erreur composée avec l'estimateur des MCQG.

Les résultats de cette estimation de la fonction de coût (1.3) figurent dans le tableau 1.4 de la page suivante. Les résultats issus des MCO, très proches, sont présentés à titre de comparaison. Tous les coefficients sont significatifs à 5 %. Le signe négatif de β_{VK} n'a rien d'incohérent, dans la mesure où, sur l'intervalle de

fluctuation de nos variables, l'élasticité estimée coût-trafic demeure positive. Dans ce cas encore, les variables exogènes introduites expliquent la quasi-totalité de la variance du coût (94 %).

Observations : 169 Variable endogène : ln$C$				
	Modèle à erreur composée MCQG		MCO	
Paramètre	Coefficient	t-test	Coefficient	t-test
α_0	18,27	7,086	18,38	4,042
β_{VK}	-1,074	-3,465	-1,184	-2,068
β_{LL}	1,092	3,902	0,8894	3,925
γ_{VK}	0,04312	4,790	0,05202	3,156
γ_{LL}	-0,07802	-2,646	-0,07325	-2,995
R^2	0,941		0,940	

TAB. 1.4 – Estimations de la fonction de coût (1.3) des sous-réseaux par les MCQG et les MCO.

Les résultats présentés dans le tableau 1.4 permettent de calculer pour chacune de nos observations :

– les rendements de taille :

$$RTS = \frac{1}{-1,074 + 1,092 + 2(0,04312 \ln VK - 0,07802 \ln LL)}$$

– les rendements de densité :

$$RTD = \frac{1}{-1,074 + 2.0,04312 \ln VK}$$

Les résultats des calculs, présentés en annexe page 221, montrent que les rendements de densité sont toujours largement supérieurs à 1 (2,7 en moyenne), ce qui incite fortement à massifier les trafics sur les différents sous-réseaux. Ces valeurs illustrent l'importance des économies de densité à l'échelle d'une ou de quelques lignes.

Les rendements de taille varient, eux, de 1,77 (Paray-le-Monial 1995) à 1,02 (Lyon - Saint-Étienne 1993). Les rendements de taille étant significativement croissants sur les sous-réseaux susceptibles d'être encore partitionnés, il n'apparaît pas intéressant de segmenter davantage les réseaux régionaux. Au contraire, il peut être pertinent de regrouper certains de ces sous-réseaux, afin de bénéficier davantage des économies d'échelle.

Les nomogrammes suivants (figures 1.1 et 1.2) permettent de repérer sur une échelle large et sur une autre, plus détaillée, les rendements de taille pour les différents sous-réseaux. Chacun des sous-réseaux est caractérisé par sa longueur (en abscisse) et son trafic (en ordonnée) ; les courbes de niveaux indiquent la valeur des rendements

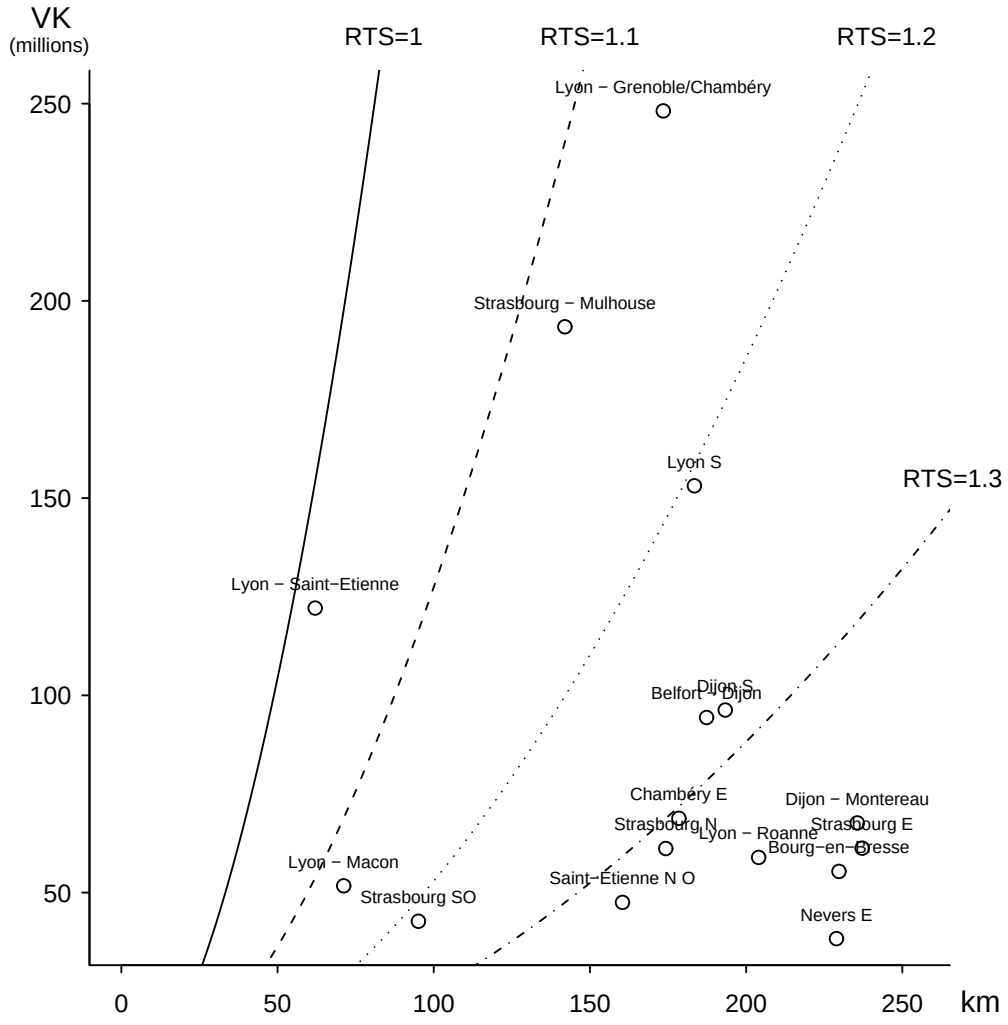

Fig. 1.1 – Nomogramme des rendements de taille pour les sous-réseaux : trafic en fonction de la longueur de lignes.

de taille. On retrouve sur la première figure la ligne Lyon - Saint-Étienne dont les rendements de taille sont presque constants (proches de 1) : un accroissement de la longueur du réseau et du trafic se répercutera proportionnellement sur les coûts, c'est-à-dire que les rendements d'échelle y sont déjà totalement consommés. À l'opposé, sur la seconde figure, le réseau de Paray-le-Monial présente de forts rendements de taille : une hausse de la longueur du réseau et du trafic augmenterait les coûts relativement faiblement, les rendements d'échelle jouant à plein.

Conclusion : Au delà des enseignements que l'on peut tirer de ces résultats, il convient de ne pas perdre de vue que ceux-ci demeurent conditionnés par le processus

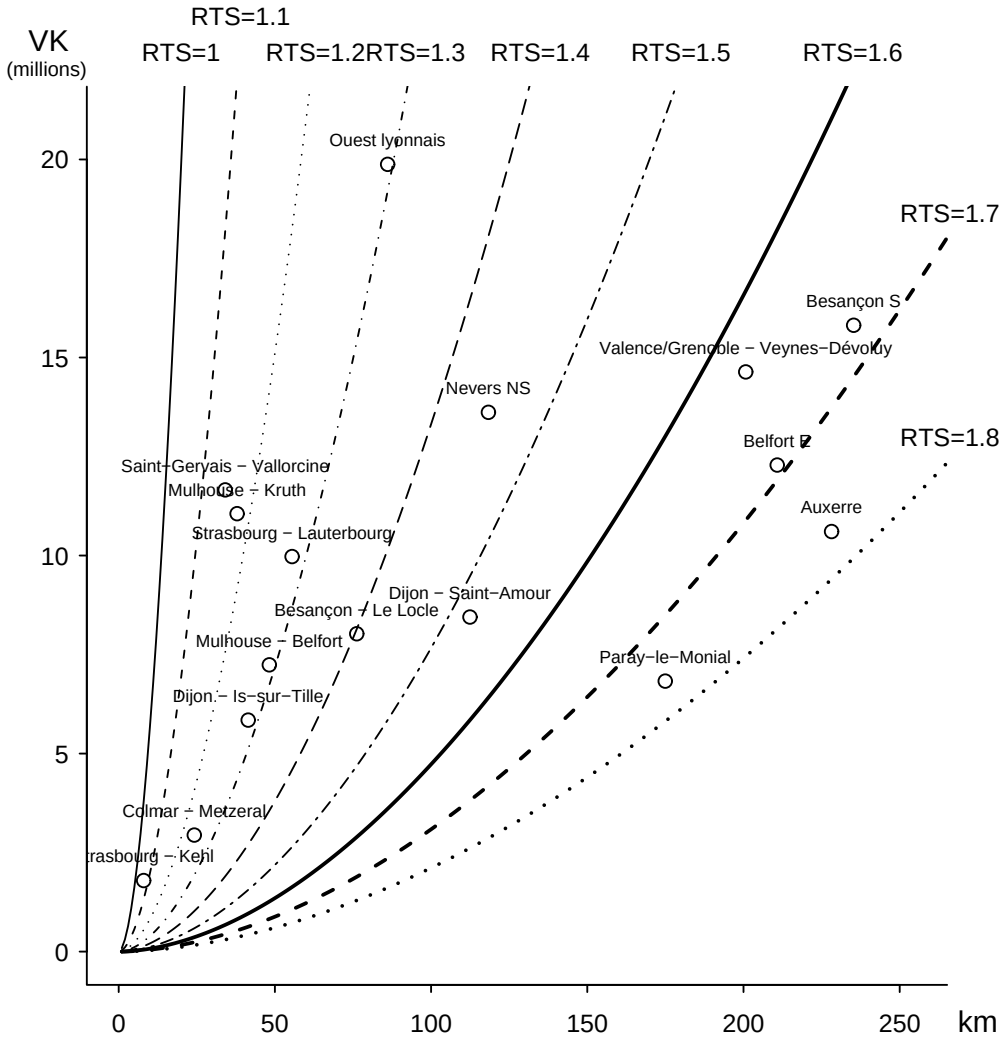

FIG. 1.2 – Nomogramme des rendements de taille pour les sous-réseaux. Détail des plus petits réseaux : trafic en fonction de la longueur de lignes.

de production de l'opérateur SNCF, en situation de monopole[12]. L'apparition de nouveaux entrants sur le marché, à des coûts inférieurs, va modifier la fonction de coût estimée et donc les rendements d'échelle que nous avons calculés.

Ces résultats sont également tributaires de la qualité des données comptables sur lesquelles ils s'appuient, de même que toute analyse de rendements d'échelle. En l'occurrence, si l'affectation des différentes charges par le logiciel comptable FC12K paraît acceptable pour les réseaux régionaux, celle qui porte à l'échelle de segments de lignes est plus discutable. En effet, il est permis de douter de l'attribution de certaines charges communes (accompagnement, manœuvre, structure) à telle ou telle portion de ligne. Il semble néanmoins que les imperfections associées à la répartition des charges communes par le FC12K sont moins problématiques que les biais introduits dans les estimations internationales par l'hétérogénéité des données.

Les résultats que nous avons obtenus plaident en faveur d'une certaine division des actuels réseaux régionaux. Les sous-réseaux que nous avons constitués pourraient parfaitement être ouverts à la concurrence dans le cadre d'enchères lancées par les régions, à l'image de la régulation adoptée par les *Länder* allemands. Mais si ces sous-réseaux s'avèrent pertinemment définis du point de vue de la structure des coûts, d'autres paramètres doivent être pris en compte lorsqu'on procède à l'allotissement. Il convient notamment d'anticiper les coûts de transaction associés à la régulation de ces marchés ; les procédures d'appels d'offre ont un coût et leur suivi contractuel également. Ainsi, il peut être collectivement préférable de réduire le nombre de sous-réseaux allotis. Il faut aussi tenir compte des externalités positives émises par ces réseaux qui n'ont peut-être pas été suffisamment prises en compte ici. En effet la présence de celles-ci, de même que des externalités négatives, rend les marchés défaillants, en l'absence d'intervention publique.

1.2 L'émission d'externalités

Il y a émission d'externalités lorsque la consommation ou la production d'un bien ou service par un agent, modifie la fonction d'utilité ou la fonction de production d'autres agents. Les externalités conduisent donc à des distorsions de consommations. En effet, la maximisation des utilités individuelles engendre un équilibre sous-optimal d'un point de vue collectif. La consommation est alors inférieure à l'optimum collectif en présence d'externalités positives de réseau négligées. Elle est supérieure à l'optimum collectif en présence d'externalités négatives non prises en compte.

1.2.1 Les externalités positives de réseau

Les externalités de réseau sont des externalités positives, liées à l'importance de l'offre, et donc à la consommation d'un bien ou service. Elles ont été définies théoriquement par Katz & Shapiro (1985) [32] qui, dans leur article fondateur, les distinguent selon qu'elles sont issues :

[12] Encore une fois, nous renvoyons le lecteur à l'article de Gasmi, Laffont et Sharkey.

- d'un effet physique direct du nombre de consommateurs sur la qualité du produit. Ainsi, l'accès à un réseau téléphonique est d'autant plus intéressant que le nombre d'abonnés potentiellement joignables est élevé. Economides (1996) [21] remarque que ces effets directs, essentiels, sont propres aux réseaux d'échange [13] ;
- d'effets indirects, au sens où il y a d'autant plus d'options et de services connexes proposés pour un produit que celui-ci est largement diffusé. Ces effets contribuent ainsi à renforcer la qualité du produit. Ainsi, sur les marchés informatiques, la plupart des logiciels développés sont d'abord compatibles avec le système d'exploitation dominant ;
- de la disponibilité et de la qualité du service après-vente (dans le cas d'un bien durable). L'appréciation du service après-vente augmente avec la taille du réseau et donc avec la diffusion du produit ;
- de divers autres effets comme l'accès à l'information sur le produit, le rôle de la part de marché en tant qu'indicateur de qualité, ou encore l'effet psychologique lié au fait que le produit soit à la mode.

Ces externalités aux origines diverses ont également une variété de destinations. Curien & Dupuy (1996) [16] distinguent les externalités intra-sectorielles des externalités extra-sectorielles, selon qu'elles bénéficient aux utilisateurs du réseau ou à des tiers. On ne parlera pas dans cette analyse d'« effets de club », terme employé uniquement dans la littérature francophone portant sur le sujet, d'autant que les différents auteurs n'en ont pas la même définition [14].

Afin de faire coïncider la somme des utilités individuelles avec l'utilité collective, il est nécessaire d'internaliser les externalités de réseau. Le prix doit ainsi être ajusté au coût marginal diminué de la variation de surplus collectif résultant de la consommation d'une unité supplémentaire. Page & Lopatka (2000) [41] soulignent que cette solution est plus aisément envisageable lorsque la propriété du réseau peut être confiée à une seule entreprise, en position de monopole naturel. Il arrive toutefois que le réseau pertinent puisse être celui de plusieurs entreprises du marché si leurs produits sont parfaitement substituables ou compatibles. Dans ce dernier cas de figure, l'internalisation est beaucoup plus complexe.

Ainsi, afin de développer leurs services, les entreprises en réseau ont largement eu recours aux mécanismes de péréquation. Il s'agissait d'introduire dans leur tarification des subventions croisées depuis les services rentables pour financer les extensions du réseau. Ce type de mécanisme permet de couvrir les coûts tout en incitant au développement du réseau, avec l'internalisation des externalités associées. Nous décrivons dans les paragaphes qui suivent le cas particulier des chemins de fer. Les réseaux d'infrastructure et de services qui les constituent émettent des externalités

[13] Littéralement, « two-way networks » ; il s'agit des réseaux au sein desquels l'échange entre deux antennes peut avoir lieu dans un sens ou dans l'autre, comme, par exemple, les réseaux routiers, ferroviaires ou téléphoniques, contrairement aux réseaux électriques, télévisuels *etc.*

[14] Curien (2000) [15] inclut les effets directs et indirects de réseau dans ce qu'il appelle « effets de club », tandis que Creti & Perrot (1997) [14] n'y incluent que les effets directs.

positives, caractérisées par des origines et des destinations diverses[15].

Les effets directs

Les réseaux de transport ferroviaire étant des réseaux d'échange, ils émettent d'importantes externalités sous forme d'effets directs. On peut distinguer ceux-ci, selon qu'ils sont intra-sectoriels ou extra-sectoriels :

- lorsque, du fait d'un potentiel de consommation suffisant, l'offre de service augmente au départ d'une ville, les voyageurs perçoivent un gain sous la forme d'une accessibilité plus large à de nouvelles destinations. Cette accessiblité se trouve notamment facilitée par l'universalité des systèmes tarifaires, informationnels et commerciaux (par rapport à un réseau partitionné). Elle est également facilitée par l'optimisation d'un graphe d'interconnexion opérée globalement - plutôt que par composantes séparées -, ce qui permet une certaine optimisation des dessertes et correspondances.
  En France, le développement des réseaux d'infrastructure, suivi de celui des réseaux de services, a fait appel à des mécanismes de péréquation qui ont permis une internalisation partielle des externalités positives de réseau. À la fin du xixe siècle, la construction des réseaux d'intérêt local a été en partie financée par les usagers des réseaux d'intérêt national. Plus récemment, certains services à grande vitesse (TGV province-province) sont financés par d'autres services TGV plus profitables. De même, au niveau de l'infrastructure, RFF finance une partie de la maintenance du réseau à partir de la rente qu'il dégage sur la ligne à grande vitesse Paris-Lyon.
- indépendamment de l'usage ou non des services ferroviaires, la proximité d'une gare au départ de laquelle l'offre de transport s'enrichit engendre des externalités positives répercutées sur le prix de l'immobilier, au bénéfice des propriétaires fonciers.

Les effets indirects

Les effets indirects liés à un accroissement de trafic sont nombreux et variés. Ainsi, plus une liaison ferroviaire est fréquentée, plus l'entreprise de transport proposera de fréquences quotidiennes et assurera la qualité du service. De même, plus une gare est fréquentée, plus les services associés (location de voiture, systèmes d'information aux voyageurs...) y seront développés.

Contrairement aux effets directs, qui induisent une sous-consommation du service, ces effets indirects ne contribuent pas à détourner l'équilibre de l'optimum collectif. Tel n'est pas le cas des externalités négatives : celles-ci contribuent toutes, quel que soit leur type, à une sur-consommation du service par rapport à l'optimum collectif.

[15] Voir Dupuy (1999) [20] pour une application à l'automobile (pages 33 à 42).

1.2.2 Les effets externes

Dans l'analyse des externalités négatives, le niveau absolu des émissions ferroviaires a moins d'intérêt que leur faible valeur relativement aux autres modes de transport. Cette faculté des systèmes de transport ferroviaire à mieux préserver l'environnement plaide en faveur de ce mode, collectivement moins coûteux. Nous présentons les différents effets externes émis par le transport ferroviaire en déclinant la typologie établie par Bonnafous (1992) [6], suivant la sphère impactée et le caractère (pécuniaire ou technologique) de l'externalité. Ce cadre d'analyse permet de discuter de façon cohérente les solutions proposées par l'économie publique aux effets externes, en terme de tarification (internalisation) et de réglementation.

Les externalités marshalliennes

Il s'agit des déséconomies externes (ou externalités pécuniaires) engendrées par un agent économique auprès d'un autre. Ces externalités sont relativement masquées en l'absence de concurrence, lorsqu'une entreprise intégrée gère elle-même l'ensemble des circulations. Néanmoins, l'usage d'une infrastructure ferroviaire partagé entre plusieurs types de circulations est de nature à engendrer des externalités marshalliennes, *a fortiori* lorsque plusieurs opérateurs ont accès à l'infrastructure :

- **Les perturbations du trafic** dues aux opérateurs (retards ou cas de force majeure) ou au gestionnaire de l'infrastructure (dysfonctionnement de l'infrastructure) sont ressenties par tous les usagers de l'infrastructure, même s'ils sont étrangers à l'incident. Cela induit des retards parfois importants qui induisent des surcoûts d'exploitation et, à terme, des pertes de revenu.
- **La gestion opérationnelle des circulations** en situation perturbée peut également engendrer des externalités au-delà de la perturbation initiale. Ainsi, les règles de priorité des circulations conduisent à retarder davantage certains convois au profit d'autres (*a priori* les plus sensibles aux retards).
- **L'occupation de l'infrastructure pour travaux** décidée par le gestionnaire d'infrastructure peut aussi entraîner des dés-économies externes aux opérateurs. En effet, la politique optimale d'entretien de l'infrastructure pour un gestionnaire d'infrastructure n'a pas de raison *a priori* de s'intercaler entre les sillons demandés par les opérateurs. La minimisation des coûts d'entretien de l'infrastructure ne correspond pas forcément à la minimisation de la consommation de capacités de l'infrastructure.
- **Les modifications de l'infrastructure** peuvent aussi engendrer des externalités marshalliennes auprès d'un opérateur. On peut imaginer que le gestionnaire d'infrastructure modifie le réseau afin qu'un opérateur bénéficie d'un accroissement de vitesse. Mais si l'ouvrage est mal dimensionné, il peut induire des surcoûts d'exploitation à d'autres opérateurs.

Ces externalités deviennent donc importantes lorsque l'ensemble des marchés n'est plus intégré au sein d'une même entreprise. Nous analysons dans Bouf, Crozet & Lévêque (2005) [7] comment la régulation ferroviaire en Grande-Bretagne et en France peut réduire la portée de ces externalités marshalliennes.

Les coûts externes collectifs

On entend par coûts externes collectifs, les externalités pécuniaires à la charge de la collectivité. Il s'agit ici du fait que l'activité ferroviaire réduit la qualité de biens et services collectifs, laissant à la collectivité la charge d'y remédier. Dans les activités ferroviaires marquées par une dynamique de (très) long terme (infrastructure, matériel roulant), ce type d'externalités se manifeste souvent avec retard. Là encore, l'établissement d'entreprises ferroviaires intégrées a longtemps dissimulé ces effets :

- **L'entretien insuffisant d'une infrastructure dégradée par le trafic** ne se traduit qu'à long terme par une baisse des performances de l'infrastructure (réduction des vitesses, incidents fréquents...). Ainsi, un gestionnaire d'infrastructure peut différer les dépenses de maintenance d'une infrastructure, jusqu'à réduire de façon conséquente la valeur de cet actif. La collectivité qui souhaite continuer à disposer de cette infrastructure est alors contrainte d'intervenir financièrement.
- **L'investissement insuffisant en matériel roulant** conduit au même effet. L'âge moyen d'un parc de matériel roulant devient un jour insupportable au point qu'il faut investir massivement, ce à quoi la collectivité soucieuse de maintenir le service de transport ferroviaire est contrainte.
- **L'endettement des entreprises ferroviaires** suit la même logique. Les dettes des entreprises ferroviaires (européennes, mais aussi japonaises) se sont développées bien au-delà de la capacité des marchés ferroviaires à les résorber, contraignant ainsi les États à en reprendre à leur charge tout ou partie.

Les effets externes inter-individuels

Ces effets externes qui aboutissent à une réduction des utilités individuelles sont les plus connus. Ils sont caractéristiques des activités de transport. Bien qu'ils soient essentiellement émis par le trafic routier, des effets externes inter-individuels sont également engendrés par le transport ferroviaire :

- **Les émissions de polluants atmosphériques** par les moteurs thermiques des engins à traction diesel peuvent être à l'origine d'importantes externalités. Les émissions d'un autorail en milieu rural sont relativement anecdotiques. En revanche, celles de locomotives diesel, chauffant avant le départ des trains dans une gare en milieu urbain dense, sont de nature à porter un préjudice grave aux individus alentour [16].
- **Les émissions de nuisances sonores** par les moteurs, le contact roue-rail et les écoulements aérodynamiques sont difficiles à monétariser [17]. Les différentes études empiriques portant sur ce sujet ont montré une tolérance particulière des individus, à intensité sonore égale, envers les émissions sonores d'origine

[16] Les valeurs tutélaires retenues en 2001 varient de 4 € pour 100 trains.kilomètres en milieu rural, jusqu'à 164 € pour 100 trains.kilomètres en milieu urbain, contre respectivement 0,1 € et 2,9 € pour 100 VP.kilomètres.

[17] Ceci, d'autant qu'il peut être nécessaire, dans une analyse coûts-avantages, de distinguer les émissions dues à l'infrastructure (le rail) de celles dues au transport (la roue)...

ferroviaire par rapport à celles d'origine routière ou aérienne. Aussi, le mode ferroviaire bénéficie-t-il de niveaux d'émission maximaux supérieurs de 3 dB à ceux du mode routier.

- **Les accidents** ont un coût pour les victimes. Si les modes de transport collectif présentent l'avantage d'être plus sûrs que les modes individuels, leur dimension collective marque beaucoup plus l'opinion. Dans ces modes, les valeurs tutélaires associées au mort (1,5 M€), blessé grave (225 k€) ou léger (33 k€) sont ainsi majorées de 50 % par rapport aux modes routiers individuels. Signalons que les statistiques nationales attribuent au mode ferroviaire les accidents intervenant lors de collisions aux passages à niveaux.
- **La congestion** est difficilement identifiable dans le transport ferroviaire. La forme la plus perceptible de congestion est dite « opérationnelle » ; elle correspond à la propagation amplifiée d'un retard sur les lignes les plus chargées. Mais il existe aussi une congestion « structurelle », lorsque la capacité de l'infrastructure ne permet pas d'établir un graphique de circulation intégrant toutes les demandes des opérateurs. La priorité alors donnée à certains trains sur d'autres ne résulte pas d'un calcul économique qui promet d'être complexe, tant les connaissances sur la congestion ferroviaire sont encore balbutiantes.
- **L'occupation de l'espace** n'est qu'imparfaitement prise en compte par le prix du foncier. Les effets de coupure, l'affectation des espaces publics et les effets sur les paysages sont à l'origine d'externalités dont il convient de tenir compte. Ces considérations sont prises en compte dans les projets d'infrastructure qui doivent tenir compte de l'insertion de celle-ci dans l'environnement.
- **Les nuisances dues à la production nucléaire d'électricité** sont bien réelles, qu'il s'agisse des nuisances directement émises par les centrales en fonctionnement (rejets d'effluents fluides, risque de fuite radioactive) ou du stockage des déchets. Les coûts correspondants sont censés être internalisés dans le prix du kilowatt.heure des analyses coûts-avantages des activités de transport.

Les effets externes sur la biosphère

Enfin, les effets externes les plus diffus atteignent l'environnement au sens extensif du terme, comme l'émission de CO_2 par les moteurs thermiques. Les valeurs tutélaires de 2001 valorisent à hauteur de 100 euros la tonne de carbone émise. Dans le cas où l'électricité fournie au réseau serait produite par des centrales thermiques, il conviendrait de tenir compte des émissions correspondantes. Mais la production de l'électricité alimentant le réseau ferroviaire français (en grande partie d'origine hydraulique et nucléaire) n'engendre pas d'émissions de gaz à effet de serre.

1.3 Les propriétés des biens collectifs

Dernière caractéristique des industries de réseau, la vérification des propriétés de biens collectifs. Les chemins de fer peuvent en effet être assimilés à un bien collectif mixte en raison des indivisibilités partielles précédemment explicitées. Se posent

alors des problèmes de non-révélation des dispositions à payer. Nous reprenons ici les termes de l'analyse développée par Lévêque (1998) [37], à savoir la non-rivalité et la non-excludabilité.

1.3.1 La non-rivalité

La propriété de non-rivalité correspond au fait qu'un bien puisse être consommé simultanément par plusieurs agents sans que la quantité consommée par l'un diminue les quantités encore disponibles pour les autres. Ce phénomène est fréquent dans les industries de réseau, tant qu'il n'y a pas de congestion. Il en résulte un coût marginal nul qui engendre une difficulté de tarification. En effet, l'instauration d'un péage économique pur, afin de couvrir les coûts, exclût les agents dont la capacité à payer est faible (alors qu'il ne coûte rien de les satisfaire). Cela conduit à un rationnement sous-optimal.

Dans les chemins de fer, la non-rivalité se retrouve d'abord au niveau de l'infrastructure. L'utilisation de celle-ci à travers la consommation de sillons est divisible. En revanche, la robustesse du graphique des circulations, (c'est-à-dire la qualité de service), est une caractéristique indivisible de l'infrastructure. Ainsi, si le coût marginal d'usage de l'infrastructure ferroviaire n'est pas nul, il est cependant très faible par rapport aux coûts fixes, en l'absence de congestion. Dès lors, l'allocation optimale des capacités d'infrastructure nécessite l'élaboration d'un mécanisme tarifaire très fin. Sous contrainte budgétaire, la tarification de l'infrastructure se complexifie radicalement (voir Baumstark & Bonnafous (1998) [3]). Du point de vue de l'allocation optimale des ressources, il ne s'agit plus de répartir des coûts, mais de favoriser la réalisation de surplus. Dans cette optique, une segmentation de la demande permet de procéder à une discrimination tarifaire, socialement efficace [18]. Une tarifiation de type Ramsey-Boiteux autorise ainsi les prix à s'écarter du coût marginal (afin d'atteindre l'équilibre budgétaire), d'autant plus que l'élasticité de la demande est faible. La non-rivalité apparaît également dans les services de transport ferroviaire, notamment lorsque la réservation n'est pas obligatoire. En effet, le coût marginal d'un passager supplémentaire peut être considéré comme étant nul. C'est la raison pour laquelle les entreprises ferroviaires développent des politiques tarifaires fondées sur des techniques de *yield management*. Celles-ci permettent de pratiquer une discrimination tarifaire afin de capter le maximum de surplus.

1.3.2 La non-excludabilité

La propriété de non-excludabilité désigne l'impossibilité d'écarter qui que ce soit de l'utilisation d'un service, y compris les individus qui ne participeraient pas à son financement : les passagers clandestins. Elle résulte de l'absence d'un dispositif technique ou juridique limitant l'accès au service. Il s'ensuit alors une désincitation à

[18] On doit à Maurice Allais la remise en question de la tarification des infrastructures au coût marginal d'usage. Ce système tarifaire est critiqué quant à sa difficile mise en œuvre, ses effets pervers et ses fondements théoriques, tandis qu'une différenciation des prix peut favoriser l'efficacité de l'économie. Pour plus de détails, se référer à Baumstark & Bonnafous (2000) [4].

produire pour l'agent privé qui anticipe ce manque de participation au financement.

Dans la pratique, les contrôles et les portails d'accès permettent de réduire la non-excludabilité dans les services de transport ferroviaire. Les passagers clandestins (ici, sens propre et figuré se confondent) restent relativement peu nombreux.
Il est toutefois possible de considérer l'infrastructure ferroviaire comme un bien partiellement non-excludable. Bien que les entreprises ferroviaires utilisant l'infrastructure soient peu nombreuses et parfaitement identifiées, le problème de la non-excludabilité peut se présenter du fait de la décision publique de ne pas exclure une entreprise de l'usage de l'infrastructure. Anticipant ce soutien public, une entreprise de transport qui n'aurait aucune incitation à révéler sa disposition à payer adopte alors un comportement de passager clandestin. Ce comportement peut notamment se révéler à travers la sous-estimation par l'entreprise ferroviaire de la rentabilité d'un service de transport à grande vitesse afin de minorer sa participation au financement de l'infrastructure. Ce comportement se traduit alors par un recours plus important aux financements publics et donc, sous contrainte budgétaire, à retarder la mise en œuvre de projets aux dépens de la collectivité.

Conclusion : une industrie de réseau à réguler

Ce chapitre nous a permis de justifier, à l'aide des concepts de l'économie publique traditionnelle, dans quelle mesure le secteur ferroviaire fait partie des grandes industries de réseau. Nous pouvons notamment en retenir les multiples incertitudes qui surviennent en terme de tarification de l'infrastructure, comme des services. La présence de monopoles exclut déjà toute pratique tarifaire concurrentielle. À cela s'ajoute les externalités (positives ou négatives) dont l'internalisation dans la tarification n'est - par nature - pas spontanée. Enfin, nous avons vu que les propriétés de non-rivalité et de non-excludabilité conduisaient également à des mécanismes tarifaires sous-optimaux.

Ainsi, l'existence de monopoles naturels, l'émission d'externalités et des propriétés des biens collectifs ne permettent pas à la main invisible du marché de réguler les chemins de fer sans une intervention publique. Au delà de ce premier diagnostic, le travail de l'économiste consiste à rechercher les modalités optimales de cette intervention publique. Il s'agit notamment de rechercher l'organisation économique du système (définition des marchés, des agents et du régime concurrentiel) qui favorisera le développement efficient du système économique.

Chapitre 2

Quelle réglementation pour les chemins de fer ?

À la suite des résultats tirés du chapitre précédent, nous étudions à présent les opportunités qui s'offrent à une intervention publique sur les marchés du système ferroviaire. Notre analyse procède chronologiquement, en commençant par tirer quelques leçons des expériences menées en matière de réglementation ferroviaire depuis l'apparition des chemins de fer. Nous détaillons ensuite les limites du monopole public intégré qui s'est imposé en Europe au cours du xxe siècle. Nous montrons notamment qu'en parallèle aux défaillances du marché, *market failures*, les défaillances de l'intervention publique, *public failures* ont une large responsabilité dans l'échec de ce mode d'organisation. Nous analysons enfin les formes de concurrence à l'aide desquelles il est envisageable de réglementer pertinemment les marchés ferroviaires. Nous montrons néanmoins que ces différentes formes de concurrence doivent être accompagnées d'une importante régulation *a priori* des marchés.

2.1 L'apport historique

L'enseignement historique est d'autant plus informatif qu'il permet une certaine prise de recul par rapport aux dernières décennies de monopole public intégré. Il a surtout le mérite d'apporter des résultats que l'on pourrait qualifier d'« expérimentaux », opportunité suffisamment rare en économie pour qu'on s'y intéresse. Nous étudions d'abord comment un certain nombre de marchés ferroviaires nationaux ont évolué, d'un grand nombre de petits concessionnaires, à des monopoles publics intégrés. De là, nous analysons plus particulièrement quelques expériences de concurrence intramodale. Enfin, nous rappelons que les crises financières ferroviaires qui ont induit des changements de réglementation résultaient, elles-mêmes, de lacunes réglementaires antérieures.

2.1.1 L'intégration verticale

Historiquement, on constate peu d'évolution dans la dimension verticale de la structure des compagnies de chemins de fer. Les entreprises européennes ont toujours intégré verticalement les trois activités de gestion de l'infrastructure, de transport de marchandises et de voyageurs. Certains réseaux français et britanniques ont cependant largement recouru à la sous-traitance sur des activités cruciales, au début de leur exploitation telles que la traction, l'entretien du matériel ou des voies (voir Ribeill (1993) [45]).

Le principal changement opéré dans la structure verticale des compagnies de chemins de fer concerne la conception du matériel roulant. En effet, cette activité qui fut autrefois la raison d'être de certaines compagnies, n'a été externalisée que très récemment, au terme d'une lente transition. Lors de la construction des réseaux, au xixe siècle, certains concepteurs de matériel roulant prenaient une concession afin de valider leur matériel [1]. Au xxe siècle, la conception du matériel roulant a d'abord lieu au sein même des grandes compagnies de chemin de fer. Celles-ci transmettent alors aux constructeurs des cahiers des charges extrêmement précis, leur laissant peu de marge de manœuvre dans la réalisation des matériels. Progressivement, ces cahiers des charges se sont assouplis jusqu'à se limiter à définir le produit en termes de performances, laissant aux industriels la charge de concevoir et proposer des projets adaptés [2]. Cette dé-intégration verticale, accompagnée de l'ouverture des marchés, permet de faire jouer la concurrence au moyen d'appels d'offres pour la fourniture du matériel roulant.

Enfin, plus récemment, certains opérateurs ferroviaires ont procédé à un second mouvement de dé-intégration verticale en externalisant les activités d'entretien et de maintenance du matériel roulant [3]. Celles-ci sont alors généralement confiées au constructeur, choisi suite à un appel d'offre plus incitatif, qui intègre donc la construction et la maintenance du matériel roulant.

2.1.2 L'intégration horizontale

Une concentration progressive

Nous analysons à présent les évolutions historiques de l'étendue spatiale des monopoles ferroviaires en Europe, en nous appuyant sur neuf pays :

- la France ;
- le Royaume-Uni ;

[1] Cela a par exemple été le cas de la ligne de Sceaux (tronçon sud du RER B) dont certaines courbes à faible rayon devaient permettre d'essayer un matériel roulant à essieux orientables.

[2] Le TGV français illustre bien cette évolution des rôles de chacun : si les premières rames PSE ont été conçues par la SNCF et réalisées par Alstom, les rames duplex ont été totalement élaborées par l'industriel.

[3] C'est le cas, par exemple, de la franchise britannique TransPennine dont l'exploitant a confié en 2003 à Siemens un marché de fourniture de matériel roulant couplée à sa maintenance ; l'industriel a pris à sa charge la construction de deux dépots de maintenance. De même, c'est Alstom qui assure la maintenance des Pendolino de Virgin sur la West Coast Main Line.

- l'Allemagne ;
- les Pays-Bas ;
- la Belgique ;
- l'Italie ;
- l'Espagne ;
- la Suisse ;
- la Suède.

Nous détaillons en annexe (page 226) les chronologies de l'organisation des marchés ferroviaires de ces pays. À quelques détails près, il est possible de résumer cette évolution par une concentration progressive des multiples concessionnaires privés opérant au milieu du xixe siècle, en monopoles nationaux publics, en charge des réseaux des différents États, au milieu du xxe siècle.

Dans la plupart des États, l'activité ferroviaire a d'abord été développée par des compagnies privées qui obtenaient une concession afin de construire puis d'exploiter une infrastructure ferroviaire. Parmi notre panel d'États, seule la Suède se distingue ; d'emblée, l'État suédois a décidé de confier l'essentiel de ses projets ferroviaires à une entreprise publique (SJ). Des opérateurs privés y ont toutefois développé peu à peu, avec l'aide de l'État, un réseau majoritairement à voie étroite, au côté de SJ.

D'une manière générale, il y a ensuite eu concentration de ces concessionnaires pour donner naissance à un nombre réduit de réseaux dans chaque État, suivant un rythme et des modalités propres à chacun. En France, les six grandes compagnies se partagent le réseau dès 1857 ; en Italie, quatre puis trois compagnies opèrent sur le même modèle à partir de 1865. En Suisse, cinq grands réseaux émergent dans les années 1880. En Espagne et surtout au Royaume-Uni, la concentration est plus progressive : les *big four* ne se partagent le réseau britannique qu'en 1923. Enfin, le mouvement de concentration s'achève dans chacun des neufs pays étudiés avec la nationalisation des réseaux, intégrés au sein d'un monopole national. Cette opération a lieu dans la première moitié du xxe siècle (de 1902 pour les CFF, à 1948 pour les BR). En Suisse et en Espagne où coexistent de nombreux réseaux à écartement différent, la nationalisation a surtout concerné le réseau principal.

Cette analyse historique des structures des marchés ferroviaires permet également de mettre en lumière quelques expériences de mise en concurrence intra-modale. Bien qu'elles n'aient pas fait preuve d'une grande stabilité, celles-ci méritent d'être étudiées plus en détail.

Quelques enseignements en terme de concurrence intra-modale

La concurrence intra-modale, c'est-à-dire entre opérateurs ferroviaires est appelée couramment concurrence *sur* le marché. Dans le cas très particulier du transport ferroviaire, celle-ci peut revêtir deux formes différentes :

- **La concurrence *entre* les lignes** correspond au cas où deux opérateurs se font concurrence sur une même origine-destination (OD), mais en empruntant chacun une infrastructure différente. Il s'agit évidemment d'une forme de

concurrence moins répandue dans la mesure où elle suppose une duplication de l'infrastructure. Mais on peut considérer qu'elle se manifeste toujours. Ainsi, l'acheminement de marchandises depuis les ports de la mer du Nord jusqu'en Italie peut se faire depuis la rive gauche ou depuis la rive droite du Rhin.

– **La concurrence *sur* les lignes** a lieu lorsque deux opérateurs se font concurrence sur une même infrastructure ferroviaire. Il s'agit de la forme la plus courante de concurrence sur le marché ; elle semble amenée à se développer avec la libéralisation des marchés ferroviaires. Pour la rendre loyale, il convient de réduire les barrières à l'entrée et les distorsions d'accès à l'infrastructure.

La concurrence entre les lignes a joué quelques temps, mais elle a engendré des effets pervers, comme la multiplication de lignes parallèles. En France, Caron (1997) [11] rapporte le contexte de concurrence que se livraient trois compagnies (Midi, PO et PLM) pour l'acheminement des vins depuis le sud du pays vers Paris. Ces trois compagnies conclurent un accord en 1863 afin de réglementer la concurrence entre elles. Cet accord stipulait que les acheminements en situation de concurrence seraient confiés à la compagnie dont les lignes offraient la plus courte distance. Cette compagnie devait alors percevoir un dédommagement raisonnable en cas de transport confié à un autre réseau, ce qui semble avoir été souvent le cas. À l'issue de cet accord, le PO en bénéficia d'abord avec l'ouverture de la ligne Périgueux - Agen. Caron rapporte qu'une guerre de réduction des distances s'engagea alors, qui explique la construction de nombreuses lignes à travers le Massif Central. En effet, les compagnies étaient incitées à construire des lignes afin d'assurer le trafic ou de percevoir le dédommagement grâce à une plus courte distance. La perversité de cet accord résultait du faible risque encouru. En effet ce mécanisme entrainait le versement d'un dédommagement quand bien même le profil et les aménagements de la ligne ne permettaient pas d'y faire passer un important trafic.

En Allemagne, la concurrence entre les lignes s'est développée progressivement à partir de la fin des années 1850, afin de réduire le pouvoir de marché des compagnies en situation de monopole (voir Fremdling (1997) [23]). Mais ce régime concurrentiel n'a fonctionné que le temps de la construction d'un réseau très dense. Lorsqu'à partir des années 1870, le maillage était tel que les compagnies n'étaient plus incitées à construire de nouvelles voies, des comportements d'entente se développèrent, poussant le gouvernement à recourir à la concurrence sur les lignes.

La concurrence sur les lignes a sans doute été moins efficace que l'autre forme de concurrence sur le marché, les insuffisances de la régulation au xixe siècle ne lui ayant jamais permis d'être effective. En effet, la concurrence sur les lignes n'a jamais été frontale, au sens où des compagnies de chemins de fer ne se faisaient jamais concurrence sur une même OD, en utilisant la même infrastructure. En revanche, plusieurs dispositions réglementaires ont préservé la possibilité d'une telle concurrence sur des infrastructures qui, bien que privées, devaient être ouvertes à d'autres transporteurs, tout comme les canaux et les routes. En pratique, l'accès des tiers à une infrastructure privée interdira le cabotage sur la section de ligne commune. Aussi, il s'agissait plus d'un droit de circulation sur la voie - moyennant péage - que d'une véritable

concurrence sur le marché.

Ribeill (1997) [46] rapporte ainsi que dès 1835, les cahiers des charges imposés par l'État français aux concessionnaires, comportent une clause de « libre parcours », suivant laquelle une nouvelle concession peut venir embrancher son réseau sur un autre réseau existant. Les conventions conclues entre concessionnaires excluent le cabotage et fixent les redevances d'usage de l'infrastructure. Celles-ci, bien que théoriquement égales au coût marginal d'usure de la voie, seront fixées de façon plutôt arbitraires et constitueront une rente appréciable pour la compagnie contrôlant la voie commune [4]. En fait, il a pu y avoir concurrence entre compagnies utilisant une infrastructure commune parce qu'au delà, les réseaux propres à chacune et leurs prolongements successifs desservaient des localités voisines. Toutefois, les conflits de toutes sortes se sont multipliés, tant et si bien que les lignes embranchées finiront absorbées par les compagnies contrôlant les lignes maîtresses.

En Grande-Bretagne, où des clauses similaires (qui avaient inspiré la France) existaient, la concurrence sur les lignes ne fut pas davantage observée. En effet, l'accès aux facilités essentielles (implantation dans les gares) n'était pas garanti, entraînant d'autres distorsions (non monétaires) de concurrence. En Allemagne et aux Pays-Bas, la possibilité de concurrence sur les lignes est apparue après la construction des réseaux, afin de limiter les comportements monopolistiques des compagnies en place. À la différence de la France et de la Grande-Bretagne, cette réglementation a échoué, face aux comportements d'entente adoptés par les compagnies. Aux Pays-Bas, lorsqu'il y a eu desserte d'une même OD par deux compagnies utilisant la même ligne, après que les prix ont baissé, les compagnies se sont entendues (avec l'État) pour mettre un terme à cette réglementation. En Allemagne, en revanche, la concurrence frontale a toujours été soigneusement évitée : certaines compagnies utilisaient l'infrastructure d'une autre, sur une partie du trajet, dans le seul but d'économiser des ressources et de gagner en temps.

Ainsi, les expériences les plus abouties en terme d'intensité concurrentielle ont engendré divers effets pervers, comme la fusion des compagnies, la formation de cartels, des comportements d'entente ou encore, une multiplication d'infrastructures parallèles. L'échec de ces expériences est venu renforcer le mouvement généralement observé de concentration que nous avons décrit. Cette évolution similaire des réseaux de chemins de fer de neuf pays européens (d'une multitude de concessionnaires privés à des monopoles nationaux publics) est d'autant plus marquante que les contextes nationaux, politiques et économiques étaient, à l'époque, relativement indépendants. Cela conduit à nous interroger sur les raisons qui ont conduit à la généralisation du modèle d'organisation ferroviaire reposant sur un monopole national public.

[4] Ribeill a calculé que la Compagnie de Saint-Germain percevait 40 % de ses recettes par le biais des péages versés par la Compagnie de Versailles Rive-Droite et par la Compagnie de Rouen pour l'utilisation commune des gares et de la voie entre Paris - Saint-Lazare et Asnières.

Des lacunes réglementaires aux crises financières

L'Histoire a surtout retenu les crises financières importantes (et, dans une moindre mesure, les éléments politiques) qui ont conduit les pouvoirs publics à revoir périodiquement la réglementation des chemins de fer. En effet, si les crises financières des compagnies de chemins de fer ont marqué l'opinion publique et conduit les gouvernements à agir, il serait erroné de croire qu'elles ont été exogènes. Il s'agissait en fait, chaque fois, d'un révélateur de l'instabilité de l'équilibre économique institué par la réglementation antérieure. Les principales lacunes des marchés, avant leurs nationalisations, étaient notamment dues aux difficultés qu'il y avait à réguler les tarifs, à développer les externalités de réseau ou encore à profiter des économies d'échelle :

- En France, la concentration en six grandes compagnies menée dans les années 1850 a été motivée par les difficultés qu'il y avait à réguler les nombreux concessionnaires antérieurs. Aussi bien le contrôle des tarifs, que la coordination des réseaux y étaient délicats. Par ailleurs, dans un contexte économique relativement incertain, l'État - comme les compagnies - souhaitait que les réseaux concédés aient une importance suffisante afin de résister aux retournements de conjoncture.
- La réforme de 1937 qui a abouti à la création de la SNCF est due, outre le marasme financier des grandes compagnies, à l'incapacité de l'État à péréquer les disparités structurelles qui affectaient les différents réseaux.
- La réforme italienne de 1885 est également marquée par des préoccupations économiques non financières. Il s'agissait alors de stabiliser les versements de l'État (en distiguant les dépenses de capital fixe et les dépenses d'exploitation), mais aussi de réduire l'hétérogénéité des tarifs.
- Le rachat progressif par l'État belge des réseaux à la fin du xixe siècle visait (au delà de son objectif politique initial) à réduire les difficultés d'exploitation que le morcellement des lignes avait engendrées. Cette réforme devait aussi homogénéiser les tarifications appliquées par les différents opérateurs.
- Enfin, la réforme prussienne des années 1860 qui a autorisé la concurrence entre les lignes de chemins de fer a été mise en œuvre afin d'éviter qu'apparaissent des monopoles régionaux, susceptibles d'abuser de leur pouvoir de marché.

À une époque historique où les entités nationales se renforçaient, le caractère stratégique des chemins de fer a progressivement conduit les États à s'assurer de leur bonne gestion, ce qui nécessitait que le nombre d'opérateurs soit faible. C'est finalement vers une organisation des chemins de fer en monopoles nationaux, qu'ont convergé les neuf pays étudiés. La nationalisation apparaît alors comme la solution pour financer les réseaux tout en garantissant que leur gestion sera guidée par l'intérêt collectif. L'envergure nationale de ces monopoles semble s'expliquer bien plus par leur appartenance au secteur public, que par la pertinence économique supposée d'une telle organisation (rendements d'échelle de toutes sortes, externalités de réseau...). En effet, dès lors que la collectivité prend à sa charge les compagnies de chemins de fer en les nationalisant, elle n'a pas intérêt, à l'époque, à conserver la gestion de plusieurs entités distinctes. Au contraire, les États qui peinaient à réguler leurs opérateurs

privés, profitent de la nationalisation pour charger l'entité publique créée, supposée bienveillante, de régler en interne les questions de tarification et péréquation. En revanche, dans l'hypothèse où les crises des systèmes ferroviaires auraient pu être résolues sans nationalisation, on imagine mal un regroupement de compagnies au sein d'un monopole privé d'envergure nationale que la collectivité n'aurait pas pu réguler efficacement, faute des moyens adéquats.

Cette organisation des chemins de fer en monopoles publics intégrés a vraisemblablement été une réponse pertinente aux problèmes de régulation, investissements, externalités *etc.*, tels qu'ils se manifestaient dans la première moitié du xxe siècle. Mais depuis, les changements profonds intervenus dans les technologies, la demande et l'analyse économique ne justifient plus de maintenir cette organisation qui a atteint ses limites au début des années 1980.

2.2 Les limites du monopole public intégré

Après une deuxième moitié du xxe siècle marquée dans les principaux États européens par une organisation homogène des marchés ferroviaires, les dernières années du siècle ont vu se tourner la page des monopoles ferroviaires publics intégrés. Encore une fois, ainsi que nous l'avons déjà noté à d'autres périodes de l'histoire des chemins de fer, les lacunes réglementaires ont débouché sur une crise financière contraignant les États à revoir la réglementation de leurs chemins de fer. Nous analysons dans cette section, dans le cas particulier de la France, comment la faiblesse de la régulation, couplée à l'absence de risque, a conduit l'entreprise ferroviaire à développer un fort excédent organisationnel, couplé à l'inefficience-X [5], paralysant son efficacité. Enfin, nous nous interrogeons sur les motifs de la bienveillance étatique à l'égard des chemins de fer, avant d'envisager les solutions proposées dans le cadre de la libéralisation.

2.2.1 La bienveillance des autorités de tutelle

Nous défendons ici la thèse selon laquelle la crise financière des chemins de fer a pour origine la bienveillance manifestée par l'autorité de tutelle envers le monopole ferroviaire, c'est-à-dire, la confiance dans la capacité de ce dernier à s'autoréguler [6].

[5] Leibenstein (1966) [36] a proposé le concept d'inefficience-X qu'il oppose à l'inefficience allocative. Celui-là recouvre l'inefficience due au manque de motivation des salariés, à l'inefficacité des unités de production, ou encore à la faible durée quotidienne de travail.

[6] Il s'agit donc de la thèse opposée à celle qui prétend, au contraire, que les difficultés financières de la compagnie de chemins de fer sont dues au surcroît d'exigences de l'autorité publique qui aurait contraint l'entreprise ferroviaire à se « surendetter », afin de financer des infrastructures et des services non rentables. Concernant les infrastructures, il est erroné de penser qu'en France la crise financière des chemins de fer qui (n')a (pas) été résolue par le transfert de dette à RFF est due au financement par la SNCF des lignes TGV (Sud-Est, Atlantique, Nord et Méditerranée). En effet, ces infrastructures à grande vitesse, socio-économiquement pertinentes, sont également financièrement amortissables, quand bien même leur rentabilité a pu, parfois, être inférieure à celle initialement envisagée. Mais, dans une logique comptable, on a considéré que les recettes issues des services TGV devaient d'abord être affectées à la péréquation des services déficitaires (gestion de l'infrastructure, transport de marchandises, transport régional), plutôt qu'à l'amortissement de

En effet, l'essentiel des difficultés financières des monopoles ferroviaires est dû au fait que les pouvoirs publics ont laissé ces difficultés s'accumuler sans se donner les moyens de discipliner le monopole. Il serait faux de mettre en cause l'asymétrie d'information (dont nous ne nions pas l'existence), quand il n'a même pas été tenté de mettre en œuvre des mécanismes incitatifs, susceptibles de s'y heurter. L'essentiel de la régulation économique des chemins de fer a été assuré par des contrats entre l'État et l'entreprise.

Les contrats État-SNCF

L'histoire de ces contrats est à l'image de celle de leurs résultats : chaotique (voir la figure 2.1 page suivante).

Le premier contrat « de programme » a été conclu pour les années 1969-73 et prolongé par avenant jusqu'en 1975 ; il posait le principe du *retour* progressif à l'équilibre budgétaire. L'insuccès de ce contrat peut être relativisé par le premier choc pétrolier. Le second contrat « d'entreprise » signé en 1979 prévoyait un *retour* à l'équilibre du compte d'exploitation à son terme, c'est-à-dire en 1982. La divergence croissante entre les recettes et les charges, à partir de 1981, en explique l'échec. Le troisième contrat « de plan » fixe, en 1985, le *retour* progressif à l'équilibre de l'exploitation pour 1989, objectif raisonnablement atteint. Un quatrième contrat de plan lui succède qui fixe l'objectif de maintenir un résultat net équilibré sur la période 1990-94. Il prévoit également le cantonnement de la dette cumulée des déficits passés dans un service annexe d'amortissement. Mais ce contrat marque en fait le retour aux déficits (8,9 Md€ 2004 cumulés entre 1992 et 1996, soit 20 % du chiffre d'affaire), engendrant un accroissement de la dette financière (+ 11,2 Md€ sur la même période). Enfin, le dernier projet de contrat a été emporté dans les mouvements sociaux de décembre 1995. Notons que le Décret 83-817 (Cahier des charges de la SNCF) prévoit toujours la conclusion de tels contrats État-SNCF...

La bienveillance des pouvoirs publics s'est manifestée dans ces contrats, tant du côté des prix pratiqués par l'entreprise ferroviaire, de ses coûts et de sa production, que du côté de l'exécution des contrats qui était très peu suivie et ne prévoyait aucune conséquence en cas de mauvaise performance.

La régulation des tarifs

Depuis la création de la SNCF en 1937, l'État a toujours eu la possibilité de s'opposer aux augmentations de tarifs proposées par l'entreprise, afin de contrôler son pouvoir de marché. Depuis 1979, les contrats limitent la hausse des tarifs à l'inflation. Si l'État a parfois différé certaines hausses, il l'a fait afin de préserver le pouvoir d'achat, plutôt que pour réduire le pouvoir de marché de l'entreprise. En effet, dans ces cas là, les contrats de plan prévoient une compensation publique, ce qui est arrivé fréquemment. Par ailleurs, l'État ne s'est jamais donné les moyens pour contrôler sérieusement les

l'infrastructure correspondante. Mais il s'agit d'une négation de la réalité économique, destinée à masquer une structure de financement malsaine.

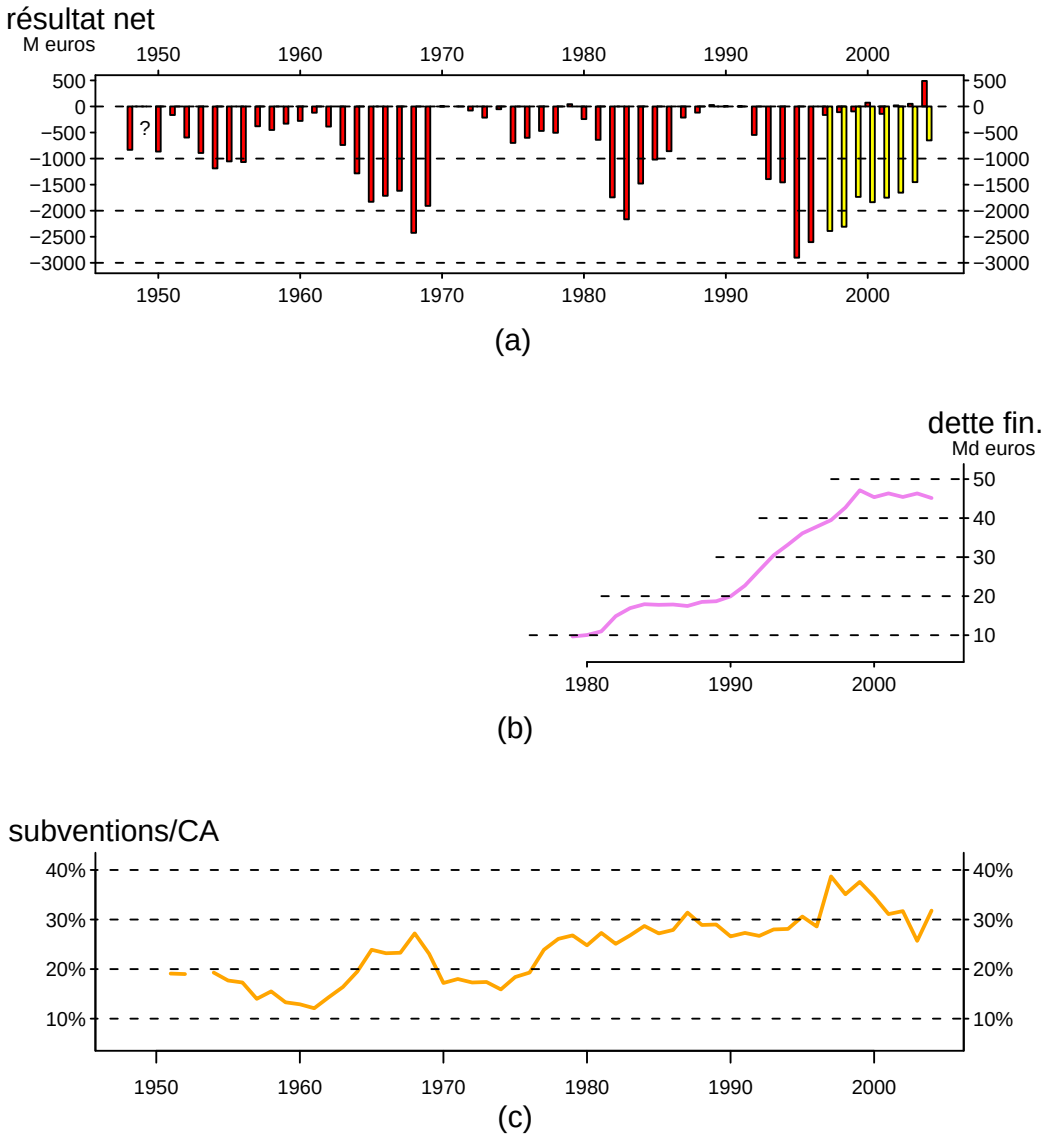

FIG. 2.1 – Évolution de paramètres financiers du système ferroviaire français :
(a) : Résultats nets SNCF ■ & RFF ■ (M€ 2004) - donnée manquante en 1949 ;
(b) : Dette financière SNCF + SAAD + RFF (Md€ 2004) ;
(c) : Part des concours publics (hors investissements, retraites et tarifs sociaux) rapportée au CA (hors doubles comptes).
Sources : SNCF, RFF et INSEE.

tarifs proposés, afin d'être en mesure de s'y opposer de façon justifiée. Le cas échéant, l'administration se serait alors heurtée à l'asymétrie d'information. Le postulat initial de bienveillance de l'entreprise publique conduisait, en effet, à réduire autant que possible le coût de la régulation, et donc à ne pas collecter l'information nécessaire à celle-ci.

La régulation des coûts

Du côté des coûts, la régulation a été indirecte dans la mesure où elle découle de l'objectif d'équilibre du budget de l'entreprise (moyennant un certain volume de subventions). Le contrat de 1979 prévoyait que les coûts ne devraient pas varier plus fortement que l'indice des prix ; celui de 1985 (qui a tenu ses objectifs) imposait même 2 % de gain de productivité annuel. Concernant plus particulièrement les charges d'infrastructure, l'évolution de la contribution publique correspondante est censée inciter à la productivité (diminution de 1 % par an dans le contrat de plan de 1985). Mais, ainsi que nous l'avons précisé, ces incitations pêchent par l'absence de pénalités crédibles en cas de non-réalisation des objectifs assignés.

La régulation de la production

Enfin, la régulation du monopole ferroviaire aurait également dû porter sur sa production, dans la mesure où celle-ci comporte des services déficitaires que l'entreprise ferroviaire n'a aucun intérêt à opérer. Pourtant, là encore, la régulation a été insuffisante. Comme pour les tarifs, les suppressions successives de dessertes non rentables ont été avalisées par les pouvoirs publics plutôt que de contraindre l'entreprise à réduire ses coûts afin de pouvoir perpétuer ces services.

2.2.2 L'absence de risque

Relativement préservée de toute contrainte imposée par la régulation sur ses prix, ses coûts et sa production, l'entreprise ferroviaire l'était aussi du côté statutaire puisqu'elle ne pouvait pas faire faillite. En effet, le statut d'Établissement Public à caractère Industriel et Commercial fait de l'entreprise une émanation de l'État qui ne peut être mise en faillite ; l'État est le responsable ultime de sa solvabilité et se substitue en dernier ressort à l'EPIC défaillant. Ce statut n'est évidemment pas incitatif dans la mesure où il permet à l'entreprise d'anticiper que ses dettes seront, tôt ou tard, reprises par l'État, dès lors qu'elles deviendront insupportables. Cette assurance rend l'entreprise totalement neutre au risque. C'est bien là l'ultime source de désincitation à la performance des monopoles publics : elle rend toute pression financière relativement infructueuse puisque toute sanction financière se traduira *in fine* dans les comptes de la collectivité. Ainsi, le système est devenu extrêmement pervers : recettes et charges évoluent de façon totalement déconnectée au niveau global de l'entreprise et, *a fortiori*, pour chaque service. Les écarts négatifs entre les deux n'appellent pas une réduction des coûts mais une augmentation du déficit et,

partant, une hausse des contributions publiques (qu'elle soit immédiate ou différée par endettement).

Dans la logique de bienveillance développée par les pouvoirs publics, l'absence d'incitations financières à la performance pouvait être compensée par le contrôle dont ils disposent sur le(s) dirigeants au sein du conseil d'administation de l'entreprise. Il est probable qu'à ce niveau l'asymétrie d'information ait permis à l'entreprise d'alléger ce contrôle. En effet, de nombreux éléments plus ou moins exogènes peuvent être mobilisés, afin de justifier ou atténuer tout résultat inférieur aux prévisions et engagements. À cela s'ajoute les changements répétés des responsables de la régulation des chemins de fer, ce qui ne favorise pas la continuité d'action face à l'entreprise ferroviaire.

2.2.3 L'excédent organisationnel

La bienveillance manifestée par les pouvoirs publics a permis aux entreprises ferroviaires de mener la « vie tranquille » dont Hicks affirmait qu'elle constituait le meilleur des profits pour un monopole. Cette vie tranquille s'est traduite sous forme d'excédent organisationnel (ou *slack*), c'est-à-dire un « ensemble de dépenses effectuées par l'organisation en excédent de ce qui est fondamentalement nécessaire à l'activité de l'organisation » (Cyert & March (1963) [17]) : salaires, effectifs, consommation... trop élevés[7]. Hart (1983) [29] montre que le monopole a naturellement tendance à développer un *slack* élevé, contrairement aux entreprises opérant dans un environnement concurrentiel, « source de discipline » (voir encadré ci-après).

Le modèle de Hart (1983) :

Considérant des entreprises évoluant sur un marché concurrenciel, Hart suppose qu'il existe une composante commune aux coûts des différentes entreprises. Lorsqu'une évolution (technologique, managériale,...) intervient et permet de diminuer la composante commune des coûts, la concurrence contraint les entreprises à répercuter cette baisse sur leurs prix. Cela leur permet d'accroître (ou au moins maintenir) leur part de marché. Hart prouve qu'une entreprise opérant sur un marché non concurrentiel va, au contraire, dépenser cette source de revenu supplémentaire afin d'augmenter son slack. Ce transfert de dépenses au bénéfice du slack s'opère aisément en l'absence de régulation, mais également en cas de régulation avec asymétrie d'information.

Ce mécanisme théorique décrit parfaitement ce qu'on a pu observer en pratique dans les chemins de fer. En effet, de nombreuses opportunités ont permis aux entreprises ferroviaires d'augmenter leur productivité du travail (mesurée par le ratio (tonnes.kilomètres + voyageurs.kilomètres)/personnel), ainsi que l'illustre la figure

[7] Pour plus de détails concernant l'excédent organisationnel, voir Michard (2000) [39].

2.2 (ci-dessous). La plupart des compagnies de chemins de fer ont doublé leur productivité en quinze ans. On peut notamment citer :

- la contraction des réseaux aux lignes les plus denses ;
- le développement des automatismes et de l'informatique ;
- la réduction des manœuvres grâce aux automoteurs et rames-blocs ;
- la réduction du nombre d'agents de traction de deux à un seul ;
- la substitution de services à grande vitesse à des services classiques ;
- *etc.*

Pourtant, ces gains de productivité ont été largement réinvestis dans le *slack* des entreprises, notamment dans la masse salariale. En effet, comme l'illustre la figure 2.3 dans le cas de la France, la masse salariale a diminué moins rapidement que les effectifs. Cela a entrainé une hausse des charges moyennes de personnel par agent (en euros constants), d'autant plus importante que les difficultés financières de l'entreprise augmentaient : + 1,5 % par an, en moyenne, entre 1991 et 2001. De manière plus fine, Garcia (2001) [25] montre, dans son analyse des comptes de surplus sur la période 1990-1999, que le personnel a systématiquement bénéficié d'une part de surplus de productivité totale. À l'inverse, les clients n'en ont profité qu'à travers un effet de rattrapage, à l'issue d'années difficiles, en 1994 (après la récession de 1993) et en 1996 (après les conflits sociaux de 1995).

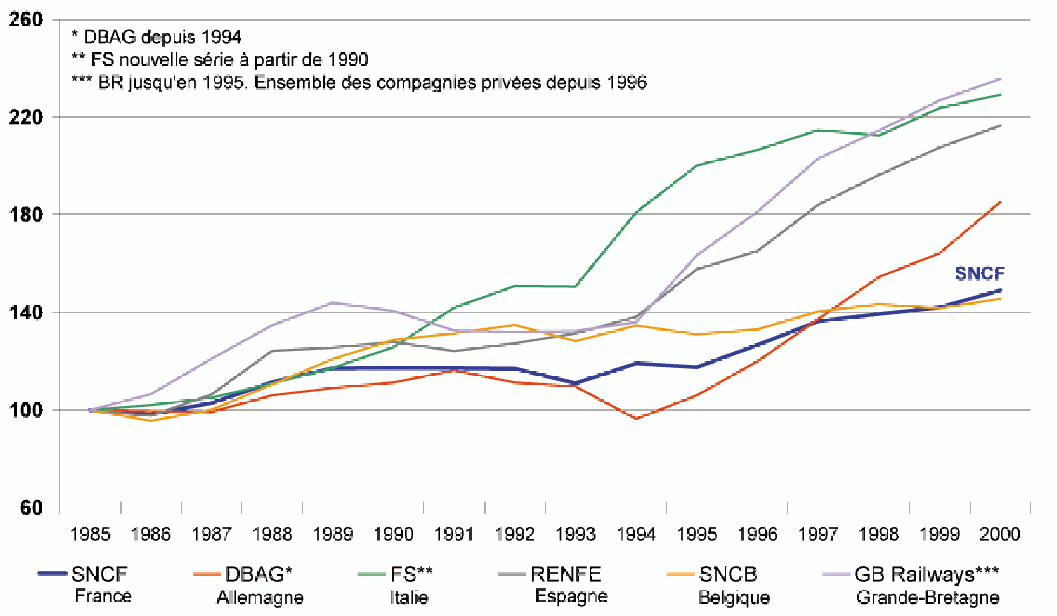

FIG. 2.2 – Évolution de la production par agent en Europe
milliers d'unités.km par agent et par an (base 100 en 1985).
Source : SNCF-UIC.

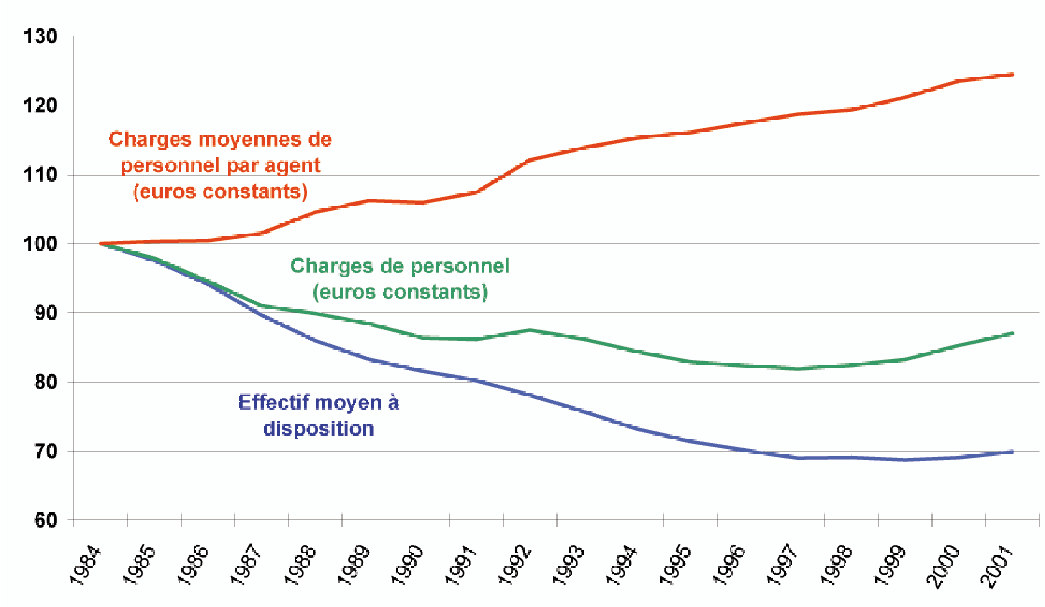

FIG. 2.3 – Évolution des charges de personnel et des effectifs à la SNCF (base 100 en 1984). Source : SNCF.

Conclusion : une tutelle « capturée » ?

Nous avons mis en évidence la bienveillance dont les pouvoirs publics ont fait preuve à l'égard de l'entreprise ferroviaire, ce qui a conduit notamment au développement d'un fort excédent organisationnel. Mais il convient de s'interroger quant à la motivation d'un tel comportement. Notamment, cette bienveillance est-elle simplement le résultat de la conception idéaliste du système étatique français, dénoncée par Jean-Jacques Laffont, ou bien, est-elle due à la « capture » de l'autorité publique par un groupe d'intérêt ?

À la suite des travaux de Bernstein, ce que l'on appelle couramment théorie de la capture correspond à la théorie de la régulation économique de l'École de Chicago. Elle fut initiée par Stigler (1971) [50], développée par Posner (1974) [44] et modélisée quantitativement par Peltzman (1976) [43] ; sur ces fondements, plusieurs aspects ont ensuite été approfondis. Cette théorie repose sur l'idée centrale suivant laquelle la régulation d'une industrie est instituée et opérée essentiellement pour le bénéfice (privé) de cette dernière. Cette conception de la régulation découle du contexte américain de l'époque : les grandes industries sont trustées par des oligopoles, et plus ou moins efficacement régulées par des agences fédérales ou étatiques. Constatant que ces mêmes oligopoles sont les principaux pourvoyeurs de fonds des campagnes électorales, Stigler et ses élèves remettent en cause la fonction d'utilité traditionnellement assignée aux régulateurs, à savoir, la maximisation du surplus collectif. Ils considèrent au contraire que la régulation économique est un produit que le gouvernement peut mobiliser au bénéfice d'individus ou de groupes particuliers,

et dont l'allocation est soumise aux lois de l'offre et de la demande. Analysant les déterminants de l'influence politique en faveur de la régulation, Posner propose trois systèmes, coexistant dans les pays démocratiques :

- **Le système entrepreunarial** qui correspond à la vente d'une régulation favorable aux industries ayant une capacié à payer suffisante. Typiquement, il s'agit des cartels américains dont une partie de la rente monopolistique qu'ils se partagent leur permet de financer des décisions qui leurs sont favorables.
- **Le système coercitif** selon lequel la bienveillance « est accordée aux groupes capables de menacer crédiblement de représailles avec violence (ou désordre, ou grèves, ou protestations) si la société ne leur fournit pas un traitement favorable. »
- **Le système démocratique** qui reconnaît l'influence que peuvent avoir les citoyens lors des élections.

Qu'en est-il dans nos chemins de fer ? Compte tenu de la (in)capacité à payer du système ferroviaire, le système entrepreunarial n'est guère adéquat pour justifier une quelconque capture. Cela pourrait néanmoins être le cas si la Société Nationale fournissait sans contrepartie des titres de voyages aux différents responsables de l'État (parlementaires, tutelles). On peut en revanche considérer qu'en proposant des postes de direction, en guise de point de chute, aux cabinets ministériels, la SNCF s'attire les faveurs des pouvoirs publics.
Mais c'est surtout la menace de conflits sociaux et leur perception par l'opinion publique qui semblent avoir contraint les autorités à concéder aux chemins de fer une régulation extrêmement favorable. Cette prise en compte de la menace sociale est manifeste pour qui s'intéresse de près aux actualités ferroviaires. Cette coercition peut être exercée directement par les groupes de pression. Mais elle peut aussi être relayée et exploitée par la direction de l'entreprise, dans une logique entrepreunariale, complice de la logique coercitive directe. Enfin, à ces déterminants vient s'ajouter la domination intellectuelle de l'opérateur historique qui réduit les capacités de la tutelle à développer une régulation émancipée.

Bien sûr, la capture n'explique pas tout. Aux débuts de la première moitié du xxe siècle, les coûts de transaction et l'insuffisance de l'analyse économique pouvaient expliquer la régulation lâche, appliquée par les pouvoirs publics aux entreprises ferroviaires. Mais lorsqu'à partir des années 1990, il a fallu accroître la dette pour subvenir à l'existence de ces entreprises, qui n'étaient manifestement pas bienveillantes, la capture explique bien la relative inertie des pouvoirs publics. L'incapacité des entreprises ferroviaires à mettre en œuvre les changements qui s'imposaient à elles, et l'existence bien réelle d'une contrainte budgétaire, ignorée le plus longtemps possible, ont engendré des crises financières dans les années 1990. Celles-ci ont trouvé un cadre de résolution dans les directives européennes qui tournent la page de l'ère du monopole public, intégré, au profit de la libéralisation des marchés. Prenant acte de la faiblesse des pouvoirs publics qui n'ont pas su réguler leurs monopoles ferroviaires, les institutions européennes ont proposé un nouveau cadre réglementaire qui prévoit le développement des marchés à la place des organisations monopolistiques,

et contraint les États à réguler ces marchés.

2.3 La libéralisation des marchés ferroviaires

Afin de discipliner les acteurs des marchés ferroviaires, il est donc apparu nécessaire, dans le contexte général de libéralisation des marchés, de remettre en cause l'ancienne réglementation des chemins de fer. L'introduction de la concurrence intra-modale paraît offrir un certain nombre de garanties, que les travers du monopole public intégré et les *public failures* avaient cessé d'apporter. En retirant aux chemins de fer le régime d'exception dont ils ont joui trop longtemps, l'introduction de régimes concurrentiels rend crédible la menace de faillite et, par là même, empêche les entreprises d'être durablement déficitaires. Mais la mise en œuvre de cette libéralisation ne conduit pourtant pas au retrait de la puissance publique. Au contraire, ses interventions n'en sont que plus nécessaires, dans le cadre de la nouvelle réglementation qui se met en place. Mais celles-ci sont à présent encadrées elles-mêmes par l'objectif de promotion d'une concurrence équitable. Ainsi, il a été procédé à la séparation de l'infrastructure et des services de transports, afin de mieux garantir l'équité d'accès au réseau.

Les avancées majeures de l'analyse économique ont ainsi permis de dégager plusieurs formes de concurrence susceptibles d'être mises en œuvre sur les marchés particuliers des industries de réseau. Outre la concurrence intermodale, il s'agit de la concurrence *pour* le marché, de la concurrence *sur* le marché, et de la concurrence potentielle. L'examen rapide[8] de chacune d'elle nous offre d'analyser les principales limites qui résulteraient de leur application aux chemins de fer. Tout au long de cette section, notre analyse s'appuie sur Vickers & Yarrow (1988) [53] et sur Armstrong, Cowan & Vickers (1994) [1] ; nous verrons que toutes ces formes de concurrence impliquent une importante régulation, parallèlement à leur mise en œuvre.

2.3.1 La concurrence intermodale

La concurrence intermodale intervient sur des marchés dès que plusieurs produits offerts aux consommateurs sont substituables les uns aux autres, mais élaborés suivant un mode technologique différent. L'élasticité croisée entre les modes est alors d'autant plus élevée que ceux-ci sont substituables. Les marchés des transports sont les plus concernés par cette forme de concurrence qui oppose les chemins de fer aux modes routiers, fluvial et aérien.

La concurrence intermodale a d'abord favorisé le développement des chemins de fer au détriment des routes et canaux au xixe siècle. Elle a ensuite confronté, sur les courtes distances, les chemins de fer au renouveau des modes routiers à partir des années 1920-1930, avec la diffusion du moteur à explosion. Enfin, dès les années 1970, les chemins de fer ont dû faire face, sur les plus longues distances, à l'essor du transport aérien (intracontinental). Les figures suivantes (2.4 et 2.5) illustrent l'évolution

[8] Les développements qui suivent ne sont en effet que des bases de réflexion dont l'analyse approfondie de leur application aux chemins de fer requerrerait d'y consacrer toute une partie...

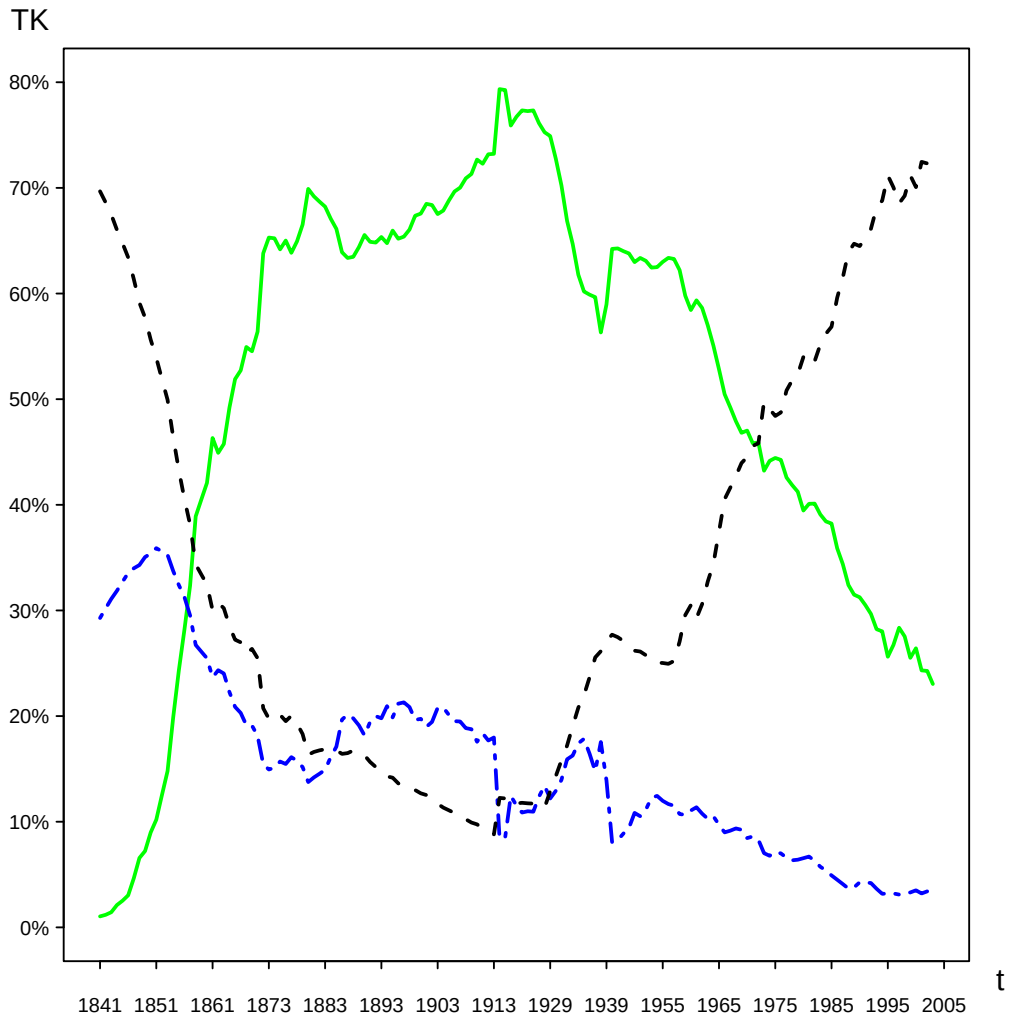

FIG. 2.4 – Évolution de la répartition modale des tonnes.kilomètres (trafic int.) ; : fer ; : route ; : voie d'eau.
Sources : Toutain (1967) [51], Sitram.

sur une longue période des parts modales entre les principaux modes de transport suivant qu'il s'agit de transport intérieur de marchandises (tonnes.kilomètres) ou de voyageurs (voyageurs.kilomètres).

Principe d'application aux chemins de fer

Cette forme de concurrence est marquée par trois caractéristiques :

- La concurrence intervient, au niveau de l'offre, entre différentes technologies : les différents modes de transport se distinguent notamment par leur flexibilité, plus ou moins importante. Celle-ci varie selon qu'il s'agit de transport

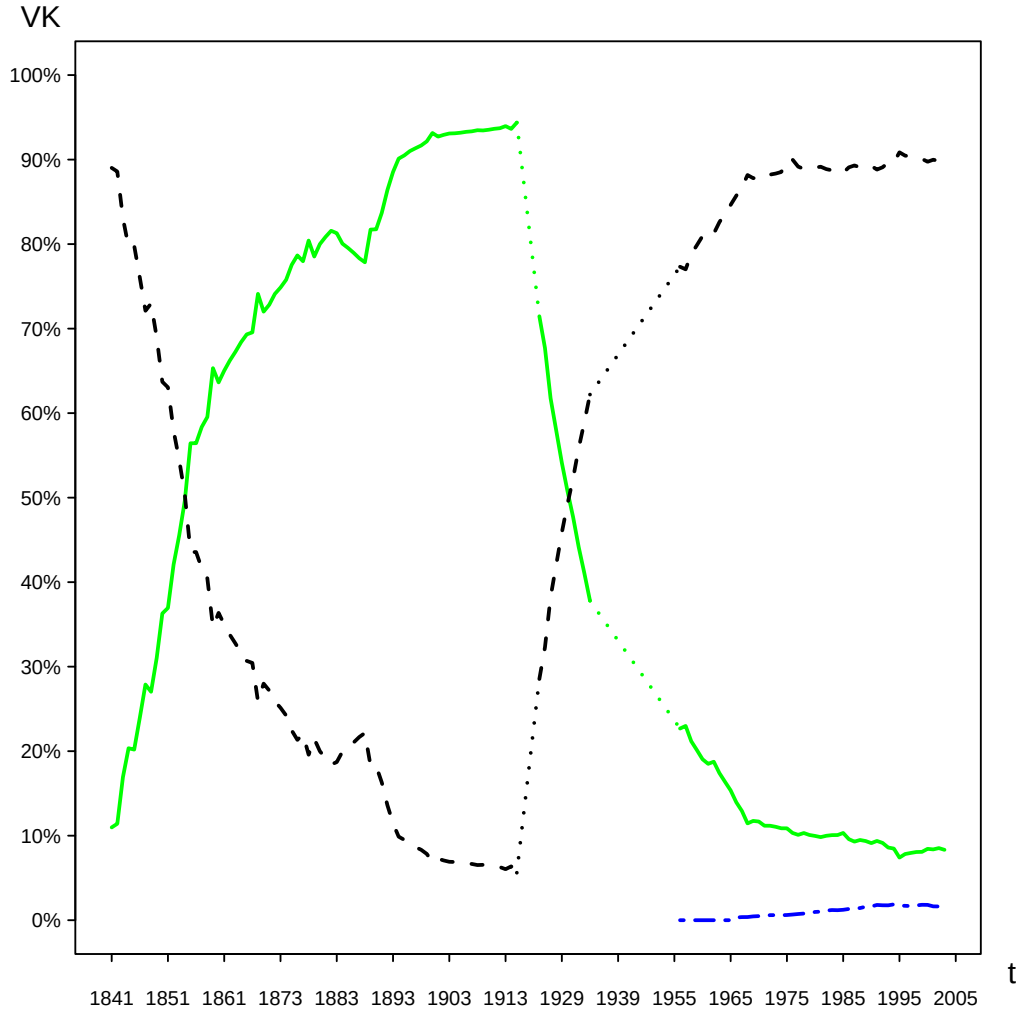

FIG. 2.5 – Évolution de la répartition modale des voyageurs.kilomètres (trafic int.) ; ——— : fer ; – – – : route ; –·–·– : aérien ; ····· : prolongement par continuité. Sources : Toutain (1967) [51], SNCF, INSEE.

individuel ou en compte propre, ou selon qu'il faille passer par un opérateur. Elle dépend également du degré d'universalité de l'infrastructure mise en jeu (contrairement aux gares, ports et aéroports, il y a des routes accessibles de partout). Seabright (2003) [48] rappelle que l'existence de rendements d'échelle croissants dans la production de l'un des modes de transport engendre une distorsion de concurrence.

- Les produits en concurrence, au niveau de la demande, sont plus ou moins différenciés : les transports proposés le sont par leur durée et, dans le cas du transport de voyageurs (mais dans une moindre mesure), par le confort qu'ils offrent. Ces deux paramètres, que sont le temps de transport et le confort, sont valorisés différemment par les consommateurs.
- Les modes concurrents sont régulés suivant des règles différentes. Or, pour que l'équilibre concurrentiel entre les modes soit optimal, il convient d'homogénéiser la régulation entre les différents modes. Il faut notamment éviter que le choix en faveur d'un mode résulte de sa sous-tarification en termes de coûts externes, mais aussi internes.

  Ainsi, en France, les pouvoirs publics ont imposé des contraintes réglementaires au transport routier de marchandises jusqu'à sa libéralisation, au nom de l'égalité des conditions de concurrence entre le rail et la route. Par ailleurs, il a été tenu compte du fait que les camions ne contribuaient pas aux coûts d'infrastructure avant la construction des autoroutes à péage : les gouvernements ont octroyé à la SNCF une contribution aux charges d'infrastructure (CCI), à présent perçue par RFF.

Des effets limités dans les transports

Pour être effective en tant que force régulatrice sur les marchés, la concurrence intermodale doit donc intervenir entre des modes de transports substituables les uns aux autres, et ce, en dépit des différences technologiques caractérisant leur production, de la différenciation verticale des produits et de l'asymétrie de la régulation appliquée à chacun d'eux. Autant de conditions théoriques dont la probabilité d'observation est faible, ce qui explique les faibles effets de la concurrence intermodale constatés sur le chemin de fer. En effet, même en supposant l'harmonisation de la régulation de ces modes, l'écart entre les offres de transport (en termes de flexibilié, durée, confort) est tel que pour chaque consommateur, le recours à tel ou tel mode s'impose rapidement.

Lorsque la durée de transport ferroviaire en TGV avoisine les trois heures, le train fait à peu près jeu égal avec l'aérien, compte tenu des difficultés d'accès aux aéroports depuis les centres-villes. Étant donnée la distribution des valeurs du temps, l'arbitrage prix-temps, pour une telle durée de transport, assure une réelle concurrence intermodale [9]. Au delà des trois heures, l'avion prend l'avantage, en deçà, c'est

[9] Précisons que cet équilibre concurrentiel autour des trois heures de transport n'a rien d'exogène. En France, la contribution du voyageur ferroviaire aux charges d'infrastructure est calculée de façon à assurer un trafic ferroviaire significatif, compte tenu de la concurrence aérienne. Le reste du financement de l'infrastructure est ensuite assuré par des subventions.

le train. Ainsi, sur l'OD Paris-Lyon, le choix en faveur du TGV n'en est pas véritablement un, compte tenu des avantages du trains sur cette OD (rapidité et accessibilité). En revanche pour les voyages en famille, la voiture particulière s'impose d'elle même la plupart du temps. Dans les transports urbains et périurbains, le choix entre mode collectif ou mode particulier est également rapide à faire, compte tenu des caractéristiques de durée et d'accessibilité de l'OD. Reconnaissons toutefois qu'il existe sur certaines OD une intense concurrence intermodale ; citons par exemple :

- l'Eurostar et l'avion sur Paris-Londres. Mais ce dernier devrait véritablement décrocher avec l'ouverture du deuxième tronçon de la ligne nouvelle britannique qui réduira à 2 h 15 le temps de trajet ;
- les navettes Eurotunnel et les ferries pour le transbordement de camions à travers la Manche ;
- le TGV et les compagnies aériennes sur des relations de plus de trois heures comme Paris-Genève ou Paris et le Midi.

Cette concurrence intermodale bénéficie généralement au consommateur à travers la guerre des prix que se livrent les concurrents. Mais Seabright rappelle que la faiblesse des élasticités croisées entre les modes, réduit les effets à court terme d'une concurrence intermodale sur les prix : « les stratégies de bas prix devront être maintenues de manière visible sur un nombre d'années significatif avant qu'un volume consistant de trafic puisse être gagné sur les autres modes ».

Mais lorsque des bas tarifs sont proposés par une compagnie ferroviaire publique, il convient de veiller à ce que ce ne soient pas les deniers publics qui financent cette guerre des prix ; en d'autres termes, à la baisse des prix doit correspondre une diminution des coûts. Cette évidence n'est cependant pas aisée à mettre en œuvre lorsque l'organisation intégrée des ressources humaines traite de façon homogène l'ensemble des trafics. Cela ne permet pas d'ajuster ponctuellement certaines consommations de facteurs de production. Dans ce cas, la solution consiste à pratiquer une péréquation, quitte à faire supporter l'effort de réduction des coûts par l'ensemble de l'entreprise. C'est ainsi que l'alignement des prix du transport ferroviaire de marchandises sur ceux de la route a conduit l'activité fret à être durablement déficitaire (sans pour autant maintenir ses parts de marché). Mais dans le contexte d'ouverture des marchés, les opportunités de péréquation se réduisent dans la mesure où celle-ci engendre des distorsions de concurrence. On observe alors un effort d'adaptation de la production locale afin de réduire les coûts sur les produits concurrencés [10].

Malheureusement, la concurrence intermodale n'a pas toujours eu les effets bénéfiques que l'on aurait pu attendre d'une force régulatrice des marchés. Dans bien des cas, les entreprises ferroviaires n'ont pas tenté de mettre en œuvre des efforts de productivité, afin de conserver leurs trafics, leurs ressources et leurs emplois. Trop souvent, les trafics déficitaires ont été un temps maintenus avant d'être transférés sur la route, afin de soulager le déficit. La concurrence intermodale a été appréhendée avec fatalisme, mais aussi avec opportunisme. Elle a ainsi permis, et permet toujours

[10] Il en va ainsi de la mise en service de trains économiques IdTGV ou encore de la réorganisation de la distribution de titres de transport dans les gares du TGV Est.

en 2005, de justifier auprès des autorités, des ressources pour maintenir des services en activité. Aussi, l'existence d'alternatives modales apparaît davantage comme un argument en faveur d'une réduction des services de transport ferroviaire que comme une incitation concurrentielle à la discipline sur les marchés. Elle ne permet pas de contraindre une entreprise ferroviaire monopolistique à ne pas abuser de son pouvoir de marché. Elle ne dispense donc pas les pouvoirs publics de réguler les entreprises ferroviaires.

2.3.2 La concurrence pour le marché

Modèles et limites théoriques

C'est à Demsetz (1968) [18] que l'on doit l'idée de mettre aux enchères, sous forme de franchises, les activités monopolisées par des entreprises. La mise en concurrence se fait par enchères *pour* pouvoir opérer sur le marché ; la franchise est alors accordée pour une période finie à l'enchérisseur qui offre le produit au meilleur prix. Selon Demsetz, « si le nombre d'enchérisseurs est important ou si, pour d'autres raisons, ils ne peuvent entrer en collusion, le prix du contrat peut être très proche du coût de production unitaire. ». Dans le sillage des autres membres de l'École de Chicago qui dénonçaient les travers de la régulation « capturée », Demsetz propose dans son article intitulé "Why regulate utilities?" la suppression de celle-ci, grâce à son mécanisme de franchise [11]. Le mécanisme permet que les prix soient fixés de façon concurrentielle et non par un régulateur. Demsetz pose donc deux hypothèses alternatives au bon fonctionnement du mécanisme :

- **La participation de nombreux enchérisseurs** assure qu'aucun des offreurs n'a individuellement de pouvoir ou d'influence sur le marché par les décisions qu'il prend. (On reconnaît l'hypothèse d'atomicité des marchés concurrentiels.) Cette hypothèse assure l'absence d'entente sur les prix.
- **L'impossibilité de collusion** permet d'assurer également l'absence d'entente sur les prix, dans le cas où le nombre d'offreurs est peu élevé.

L'idée de Demsetz a été développée formellement par Laffont & Tirole (1987) [34] et par Riordan & Sappington (1987) [47] qui ont analysé le processus d'enchères optimal lorsque les offreurs disposent d'une information privée sur les coûts, dont le régulateur n'a pas connaissance. Ils montrent qu'il est optimal pour le régulateur de sélectionner le titulaire du marché en proposant aux offreurs d'enchérir sur un menu de contrats, comme dans le cas de la régulation du monopole. Le mécanisme de concurrence pour le marché s'est avéré efficace pour des activités peu exigeantes en coûts irrécouvrables (*sunk costs*), comme le transport par bus ou par car. Il a également été largement adopté par le secteur privé à travers la pratique des appels d'offre. En revanche, il semble plus délicat à mettre en œuvre dans les industries de réseau qui ont une consommation intensive de capital spécifique (transport ferrovaire, distribution d'eau, énergie,*etc.*).

[11] Demsetz reconnaît tout de même qu'on ne peut se passer de régulation lorsque les investissements ont une durée d'amortissement élevée.

En effet, dans ce cas, plusieurs auteurs ont montré qu'outre l'hypothèse de non-collusion, d'importantes limites s'opposaient au bon fonctionnement du mécanisme. Celles-ci ont notamment trait aux avantages conférés au titulaire d'une franchise, lors du renouvellement, par rapport aux autres offreurs :

- **L'avantage informationnel au sortant** recouvre deux sources de distorsion de concurrence. La première concerne l'expérience acquise par le sortant, en termes de délais d'apprentissage technique et organisationnel. Elle lui a permis de réduire ses coûts par une meilleure connaissance du mode de production optimal. Cette expérience l'autorise donc à s'engager dans les enchères avec une meilleure connaissance des techniques de production optimales que ses concurrents.
  La deuxième source de distorsion de concurrence résulte de l'asymétrie d'information entre le sortant et ses concurrents concernant la demande et les coûts de production. Les nouveaux entrants sont pénalisés par l'incertitude qui les conduit à soumettre des offres très prudentes (au risque de perdre le marché), plutôt que d'être victime de la « malédiction du vainqueur » (c'est-à-dire remporter l'enchère suite à une surévaluation de la valeur du marché).
- **La transmission des actifs** soulève l'épineuse question des incitations à l'investissement dans un tel contexte réglementaire. Il s'agit à la fois des actifs humains spécifiques développés qui ne sont pas toujours transposables d'une entreprise à une autre, mais aussi, et surtout, du capital physique résiduel à la fin de la franchise. Les conditions de transmission de ce dernier sont en effet délicates à régler, notamment en raison de l'incertitude sur l'évaluation de son montant. Si ces actifs ne sont pas transférables, l'opérateur en place n'est incité à investir que s'il est certain d'amortir ses actifs sur la durée de la franchise. Dans le cas contraire, il requiert une assurance contre les pertes qui résulteraient d'un non-renouvellement de son contrat.

Laffont & Tirole (1988) [35] ont proposé que le régulateur biaise la concurrence en faveur du sortant ou de ses concurrrents en fonction des caractéristiques des investissements (importance, transférabilité). Une autre solution, développée notamment dans les transports consiste à réduire autant que possible les investissements à la charge des opérateurs ; ceux-là sont alors pris en charge la plupart du temps par la collectivité délégatrice.

Ces limites ne s'opposent pas à la mise en œuvre de la concurrence pour le marché. Elles soulignent néanmoins la nécessité d'une régulation spécifique des marchés ainsi réglementés, contrairement à ce que Demsetz espérait. En effet, ces limites commandent de spécifier des contrats de franchise précisant essentiellement le régime des investissements. Au delà de la question des investissements, les franchises apparaissent significativement plus complexes que ce que laisse imaginer la proposition de Demsetz.

Il est notamment apparu nécessaire de réguler explicitement l'offre, les prix, les recettes et les transferts entre le régulateur et le délégataire. Idéalement, un contrat *complet* spécifie les termes régissant le marché pour toute la durée de la franchise

et dans toutes les circonstances susceptibles d'advenir. Évidemment, en présence d'incertitude, l'écriture d'un tel contrat est impossible. Théoriquement, un contrat inconditionnel détaille exhaustivement les conditions d'exécution du contrat, en fonction de l'évolution de la demande, de l'inflation... Mais ces contrats inconditionnels manquent de flexibilité, ce qui s'avère pénalisant lorsque, sur des franchises de longue durée, l'offre requiert d'être adaptée aux évolutions imprévisibles de la technologie et de la demande. Ainsi, les contrats de franchise sont nécessairement incomplets. Cela rend d'autant plus nécessaire une régulation continue de la franchise, afin de pouvoir renégocier les termes du contrat, le cas échéant. Estache, Guasch & Trujillo (2003) [22] et Guasch, Laffont & Straub (2003) [28] ont analysé les mécanismes de renégociation des contrats de concessions, dans le contexte financier particulier de l'Amérique latine [12].

Ainsi, la pertinence de la concurrence pour le marché est réduite lorsque l'industrie est caractérisée par une grande incertitude sur la technologie et la demande. Analysons à présent ce qu'il en est concernant les chemins de fer.

Une forme de concurrence adaptée aux chemins de fer

La concurrence pour le marché a été mise en œuvre par plusieurs États européens dans le cadre de la libéralisation des chemins de fer :

- en Grande-Bretagne, la plupart des trafics de voyageurs sont assurés dans le cadre de contrats de franchises ;
- dans de nombreux *Länder* allemands et en Suède, ce sont les services régionaux qui sont concernés par cette forme de concurrence.

L'hypothèse initiale posée par Demsetz concernant l'absence de collusion sur ces marchés semble relativement acceptable dans la mesure où la libéralisation a permis à de nombreux opérateurs de s'implanter sur les marchés. En Grande-Bretagne, la désintégration, suivie de la privatisation de l'opérateur historique, a donné naissance à de nombreuses entreprises ferroviaires. Ailleurs, des entreprises locales et des groupes étrangers se sont développés aux côtés des opérateurs historiques. Aussi, dans le contexte d'ouverture des marchés ferroviaires, les acteurs cherchent à conquérir des marchés, ce qui assure une intensité suffisante à la concurrence lors des appels d'offre.

En revanche, l'application de la concurrence pour le marché dans les chemins de fer se heurte à l'avantage informationnel en faveur des opérateurs historiques qui occupent le marché depuis des dizaines d'années. Cette limite n'a toutefois pas empêché les nouveaux entrants de remporter de nombreux appels d'offre (un sur deux en Allemagne) dans la mesure où ils sont nettement plus productifs que les opérateurs historiques. En Grande-Bretagne, les efforts en faveur d'une régulation favorisant la

[12] Après avoir modélisé l'impact sur la probabilité de renégociation des contraintes institutionnelles, des chocs exogènes, du financement externe et de la régulation, Guasch, Laffont et Straub ont testé empiriquement leurs résultats. Il ressort de leur analyse que la qualité des institutions (autorité de la loi, absence de corruption) réduit les cas de renégociation, de même que les chocs exogènes positifs ou encore, la présence d'un régulateur. En revanche, les cycles politiques favorisent les renégociations (post-électorales), tandis que l'impact des investissements est plus ambigu.

concurrence, réduisent l'avantage informationnel du titulaire de la franchise lors de son renouvellement. Mais, en France où l'existence d'opérateurs alternatifs aux côtés de l'opérateur historique est très relative, l'avantage informationnel en faveur de l'opérateur historique risque de pénaliser durement les nouveaux entrants [13].

Enfin, les chemins de fer étant une industrie nécessitant beaucoup de capitaux et des compétences humaines spécifiques, la question de la transmission des actifs y joue un rôle prépondérant lors des mises aux enchères. Si la séparation de l'infrastructure et des services de transport permet déjà d'en atténuer les conséquences, les questions d'investissement en matériel roulant et de transmission du personnel demeurent très importantes. En Grande-Bretagne, la réglementation confie à d'autres acteurs que les franchisés la tâche d'investir dans le matériel roulant. Les *Rolling Stock Compagnies*, essentiellement détenues par des banques, achètent, puis louent aux franchisés, du matériel roulant [14]. Dans Gauthier-Lescop & Lévêque (2005) [27], nous avons analysé comment ces questions de transmissions des actifs étaient résolues dans divers pays européens, dont l'Allemagne, avant de formuler quelques propositions concernant la France.

Quel bilan de la concurrence pour le marché dans les chemins de fer britanniques ?

La réforme ferroviaire britannique de 1993 a introduit la concurrence pour le marché sur 25 franchises, disposition qui a fait ses preuves malgré certaines limites. Si elle a permis une très forte hausse du trafic, stimulé par le développement de nouveaux services, la concurrence pour le marché a échoué à réduire le coût des services de transport ferroviaire pour la collectivité. Les subventions versées aux opérateurs ont, certes, légèrement diminué après la privatisation de British Rail. Mais cette baisse n'a pas atteint les espérances initiales, qui prévoyaient que la plupart des opérateurs dégagent suffisamment de bénéfices pour en reverser une partie à l'État. De ce fait, en partie, beaucoup d'opérateurs ont dû renégocier leurs contrats de franchise, insuffisamment flexibles, compte tenu des turbulences des premières années de la réforme.

En 2004, le gouvernement a entrepris une actualisation de la réforme afin de tirer les leçons des expériences passées. Il s'est notamment engagé à ne plus renégocier les contrats, dont la durée s'est stabilisée à 7 ou 8 ans, tandis que le nombre de franchises a été réduit par concentration spatiale.

L'intensité concurrentielle lors des appels d'offre semble actuellement satisfaisante avec l'arrivée de filiales des opérateurs historiques du continent. Le marché

[13] Voir à ce sujet Gauthier-Lescop & Lévêque (2005) [27] pour une analyse détaillée.

[14] Signalons néanmoins que cette location du matériel roulant, bien que garantissant l'efficacité des mécanismes d'enchères, s'opère sur un marché qui n'est ni concurrentiel, ni régulé. Aussi le cartel tacite entre les trois RoSCos assure à chacune d'elle une confortable rente, indirectement financée par la collectivité qui subventionne les franchisés.

reste toutefois assez concentré : trois groupes (National Express, Virgin et First) se partagent plus de la moitié du trafic. Les avantages au sortant ne paraissent pas biaiser les appels d'offre au détriment des nouveaux candidats, le gouvernement tenant compte des performances passées du sortant. Ainsi, le mécanisme fonctionne suffisamment bien pour avoir été conforté politiquement en 2004, bien que certaines dispositions inhérentes à sa mise en œuvre, comme l'existence des *Rolling Stock Companies*, doivent encore être améliorées. Il n'en demeure pas moins que pour parvenir à ce résultat, les Britanniques ont dû élaborer une importante régulation [15].

Bien que nous discutions essentiellement de la réglementation des services de transport ferroviaire, les infrastructures peuvent également être concernées par cette forme de concurrence. En Grande-Bretagne, les concessions accordées par le gestionnaire d'infrastructure, Railtrack, pour l'entretien de celle-ci ont été un échec. Mais la concurrence pour le marché est tout à fait adaptée lorsqu'il s'agit de construction d'infrastructure. Ainsi, en France, les différents lots de maîtrise d'œuvre, génie civil et génie ferroviaire, associés aux investissements en infrastructure, sont attribués par appels d'offre lancés par RFF, maître d'ouvrage. Le gestionnaire d'infrastructure a d'ailleurs été mis en concurrence (pour la ligne nouvelle franco-espagnole entre Perpignan et Figueras), avant même que l'entreprise ferroviaire historique ne le soit. Il est vrai que pour de tels marchés, dont les enchères ne sont pas répétées, les questions d'avantage informationnel au sortant et de transmission des actifs ne limitent pas le mécanisme.

La concurrence pour le marché semble donc être une forme de réglementation bien adaptée aux chemins de fer. D'une certaine façon, c'est elle qui s'applique sur les marchés de transport de marchandises lorsque les chargeurs concluent un contrat d'acheminement avec l'opérateur qui s'engage à leur fournir le meilleur service.

Comme le prouve l'ouverture du réseau allemand à la concurrence, la concurrence pour le marché est très efficace lorsqu'il s'agit du transport régional. Il devrait en être de même en France. Sur la base des sous-réseaux régionaux identifiés dans la partie précédente (qui constituent des monopoles naturels) et en tenant compte des externalités de réseau, les services de transport régional peuvent être mis aux enchères.

En revanche, nous ne pensons pas que les services de transport longue distance doivent être soumis à cette réglementation. En effet, l'importance de ces services en France, conduirait à allotir au préalable le réseau (à la manière britannique), en raison de l'avantage considérable acquis par l'opérateur historique. Or un tel allotissement nous semble actuellement peu pertinent, compte tenu des externalités positives qui résultent de ce réseau de services. Aussi, dans la suite de notre étude des formes de concurrence ferroviaire, nous nous intéressons plus particulièrement à la réglementation des services de transport à longue distance.

[15] Voir à ce sujet Bouf, Crozet & Lévêque (2005) [7].

2.3.3 La concurrence sur le marché

À défaut de concurrence *pour* le marché qui exige un allotissement des réseaux, nous analysons à présent dans quelle mesure la concurrence *sur* le marché est adaptée à la réglementation des services ferroviaires longue distance. Il s'agit tout simplement d'ouvrir l'accès aux infrastructures ferroviaires à d'autres opérateurs habilités au trafic de voyageurs. Cette forme concurrentielle, couramment appelée « *open access* », consiste donc à réduire au maximum la réglementation, afin de permettre à la main invisible de réguler le marché.

Limites théoriques à une application aux chemins de fer

L'objection théorique essentielle qui peut être avancée à la mise en œuvre de cette forme de concurrence sur les marchés ferroviaires est liée à l'existence d'économies de densité comme celles que nous avons évaluées au chapitre précédent. De ce fait, sur une ligne ou un réseau d'infrastructure donné, le coût marginal lié au trafic est nettement inférieur au coût moyen. Aussi, en régime concurrentiel, les incitations à diminuer les prix sont très fortes à court terme, afin de capter le trafic, mais insoutenables à moyen terme, puisque les prix concurrentiels ne permettent pas de rémunérer le capital. Ainsi, Seabright (2003) [48] remarque qu'« un régime stable de concurrence sur le marché risque souvent de ne pas être viable : soit il ne peut être mis en œuvre, soit il peut l'être et conduit à une telle concurrence sur les prix, qu'à moins que les concurrents aient précisément la même structure de coût, l'un d'eux sera contraint de quitter le marché ». Ainsi, à supposer qu'un nouvel entrant vienne concurrencer un opérateur historique (en dépit des coûts d'entrée sur le marché et de l'asymétrie entre eux deux), la guerre des prix, un temps profitable aux consommateurs, aboutira rapidement à une nouvelle monopolisation du marché.

Pour se prémunir d'une telle issue, des concurrents dont les structures de coût sont semblables peuvent tenter d'atténuer l'impact de la concurrence en différenciant leurs produits. Cette différenciation, verticale, consiste à proposer des produits de qualités différentes, à des prix distincts. Dans les chemins de fer, cette différenciation apparaît principalement sur deux paramètres :

- le temps de transport qui s'effectue à plus ou moins grande vitesse ;
- le confort du voyage selon, par exemple, que l'on est en première ou en seconde classe.

Cette différenciation permet aux entreprises de maximiser leurs profits en segmentant le marché selon la capacité à payer des consommateurs. Gabszewicz, Shaked, Sutton *et al.* (1986) [24] ont déterminé la répartition optimale entre les produits d'un monopole qui segmente son marché. Wauthy (1996) [55] caractérise dans un modèle de duopole les choix de différenciation, lorsque les entreprises décident simultanément de la qualité de leur produit, puis se font concurrence sur les prix. Enfin, de Villemeur, Ivaldi & Pouyet (2003) [54] analysent, dans le cadre de l'ouverture des marchés ferroviaires, les diverses opportunités de différenciation verticale, l'impact qui en résulte et les stratégies à mettre en œuvre. En supposant que les coûts d'en-

trée sur le marché sont nuls, et suivant l'hypothèse retenue pour la distribution des consommateurs suivant leur capacité à payer, leurs résultats sont les suivants :

- **Lorsque la population est uniformément distribuée,** l'écart de capacité à payer entre les extrêmes de la population doit être suffisamment élevé pour qu'il puisse y avoir concurrence. Il faut que la capacité à payer la plus grande soit le double de la plus faible, c'est-à-dire que la population présente une hétérogénéité importante, pour qu'il y ait suffisamment de demande sur le marché en faveur de produits différenciés.
- **Lorsque la population valorise peu la qualité,** c'est-à-dire, lorsque sa distribution est concentrée autour d'une faible valorisation de la qualité (hypothèse plus pertinente [16]), il ne peut y avoir de concurrence. En effet, dans ce cas, l'élasticité de la demande est alors trop forte pour permettre à un concurrent d'entrer sur le marché.

Ainsi, les opportunités de développement de la concurrence sur les marchés ferroviaires s'avèrent extrêment limitées d'un point de vue théorique. En pratique, les rares cas de concurrence sur le marché ne démentent pas cette tendance. En effet, aux limites théoriques étudiées ci-dessus s'ajoutent d'autres éléments qui réduisent encore la pertinence de la concurrence sur le marché :

- **Les coûts de changement,** *switching costs*, modélisent la relative inertie des consommateurs ayant la possibilité de se détourner de l'opérateur en place. Ces coûts limitent sensiblement la concurrence, ainsi qu'il est possible de l'observer dans les autres industries de réseaux déjà libéralisées. Seabright remarque que ces coûts sont très faibles pour les voyageurs occasionnels qui font jouer la concurrence, tandis qu'ils sont très élevés pour les voyageurs fréquents (commercialement plus intéressants), fidélisés par l'opérateur historique.
- **L'asymétrie d'offre** entre l'opérateur historique et les nouveaux entrants accentue le biais favorable au premier. En effet, si l'opérateur historique propose des services de transports optimisés sur un réseau entier avec de nombreuses offres commerciales connexes, le nouvel entrant ne peut d'abord proposer qu'une relation point à point. Il apparaît alors d'autant plus désavantagé que le réseau est monocentré, comme en France. De plus, afin de pouvoir concurrencer l'opérateur historique, le nouvel entrant doit pouvoir offrir une fréquence de services semblable. Cela nécessite une forte réserve de capacités d'infrastructure, et un lourd investissement initial.
- **Les coûts d'entrée** sont d'autant plus pénalisants qu'ils sont irrécouvrables.
- **L'écrémage du marché,** si la concurrence peut s'y développer, rappelle qu'il n'y a pas de raison, *a priori*, pour que les nouveaux entrants concurrencent l'opérateur historique sur tous les segments du marché. Au contraire, ceux-là cherchent d'abord à maximiser leur profit, plus qu'à discipliner l'opérateur historique. La différenciation des produits aidant, il y a tout lieu de penser que l'ouverture des marchés conduise à un développement inégal de la concurrence. Précisons enfin que les différents marchés ont déjà été écrémés par... le ges-

[16] Il y a effectivement plus de demande pour voyager en seconde classe qu'en première.

tionnaire d'infrastructure dont la tarification repose en grande partie sur la capacité à payer des utilisateurs. Sur les liaisons à grande vitesse, près du tiers du chiffre d'affaire de la SNCF est prélevé par le gestionnaire d'infrastructures (et plus de 50 % sur Eurostar).

Deux expériences de concurrence sur les marchés ferroviaires

À notre connaissance, peu d'expériences ont récemment illustré les effets de la concurrence sur les marchés ferroviaires. Le premier cas de concurrence oppose en Allemagne l'opérateur historique DB à Connex, sur trois lignes interrégionales (non subventionnées), en étoile autour de Berlin : Gera-Berlin-Rostock, Zittau/Dresde-Berlin-Stralsund et Cologne-Berlin-Rostock. Il s'agit ici de concurrence *entre* les lignes dans la mesure où le nouvel entrant remet en service des lignes secondaires abandonnées par l'opérateur historique. Les deux compagnies se font néanmoins concurrence sur certaines OD. Du fait de la forme de concurrence sur le marché choisie, le nouvel entrant différencie fortement ses services, en optant pour une desserte fine des territoires, et des temps de parcours plus longs que son concurrent. Si les deux premières relations (qui s'articulent avec les dessertes régionales subventionnées, remportées par Connex lors d'appels d'offre) perdurent depuis 2002, la liaison Cologne-Berlin n'a pas tenu plus de 100 jours, faute de demande suffisante.

Le second cas de concurrence sur le marché oppose, depuis 2003, Thalys et l'ICE allemand sur la relation entre Bruxelles et Cologne. Il s'agit ici de concurrence *sur* les lignes puisque les trains utilisent la même infrastructure, le premier en prolongation de certains services Paris-Bruxelles, le second en prolongation de la relation Francfort-Cologne. Bien que la concurrence semble frontale, il convient de la relativiser pour des raisons de coopération et différenciation. En effet, bien que l'opérateur allemand ait imposé sa présence sur le marché, les deux entreprises se sont entendues sur le partage des capacités horaires, de façon à se répartir les neufs allers-retours offerts entre les deux villes. Il s'agit en fait beaucoup plus d'une coopération que d'une concurrence, ainsi que le prouve l'organisation cadencée des trains[17]. Sur ce marché, la différenciation intervient sur les offres commerciales, chacune étant adaptée à la clientèle nationale d'origine : réservation obligatoire pour Thalys, libre accès aux ICE. De ce fait, et en raison de la faible intensité de cette concurrence, chaque clientèle nationale reste attachée à son opérateur : Thalys pour les Belges, ICE pour les Allemands, chacune des offres ayant une notoriété assez faible de l'autre côté de la frontière ; cela peut être perçu comme une manifestation des *switching costs*.

Ainsi, la concurrence sur le marché ne semble pas en mesure, tant en théorie qu'en pratique, de s'appliquer efficacement aux chemins de fer. De nombreuses limites intrinsèques au secteur, ainsi que l'inertie des consommateurs, rendent peu crédible une remise en cause du pouvoir de marché de l'opérateur historique. Une

[17] Pour la journée du mardi 8 mars 2005, les départs de Bruxelles desservant Cologne étaient les suivants : 6h25 (ICE), 8h28, 10h28, 12h28 (Thalys), 13h28 (ICE), 14h28 (Thalys), 16h28 (ICE), 17h28, 19h28 (Thalys). Précisons que le site Internet de DB indique ces neuf circulations, tandis que celui de Thalys n'indique que ses six TGV...

solution consisterait à ce que les pouvoirs publics mettent en œuvre une politique pro-concurrentielle, en faveur des nouveaux entrants ; là encore, cela implique une régulation très forte des marchés. Il y a donc fort à craindre qu'en l'absence de telles politiques, la libéralisation des marchés ferroviaires sous forme d'*open access* soit inefficace. Elle risque même d'engendrer des effets pervers, dans la mesure où l'ouverture d'un marché ne garantit pas qu'il y aura effectivement concurrence sur celui-ci. Aussi, comme le soulignent Armstrong, Cowan & Vickers (1994) [1], le risque est grand que « le laisser-faire conduise non pas à la concurrence, mais au monopole non régulé. » En effet, le libre accès au marché, n'autorise plus à réguler les tarifs, la consistance du réseau de dessertes, *etc.* Cette perspective de *dérégulation* enthousiasme d'ailleurs les différentes compagnies ferroviaires européennes qui s'accordent plus ou moins ouvertement sur la politique d'*open access*. Les entreprises ferroviaires historiques ont bien compris [18] qu'elles y avaient tout intérêt. Non parce qu'elles s'estiment chacune supérieure aux autres, mais parce qu'elles savent que leur pouvoir de marché sera non seulement préservé, mais encore renforcé par le retrait de la régulation économique qui pèse encore sur elles. Cette conclusion peu encourageante, nous invite à poursuivre notre examen des formes de concurrences susceptibles de s'appliquer aux marchés ferroviaires. En l'absence de concurrence directe sur le marché, examinons dans quelle mesure la concurrence potentielle est susceptible de discipliner l'opérateur en place.

2.3.4 La concurrence potentielle

À défaut de concurrence effective *sur* le marché, Baumol, Panzar & Willig (1982) [2] ont montré que le même résultat pouvait être atteint, en terme de concurrence potentielle. Ils montrent que celle-ci peut être appliquée aux marchés « contestables », c'est-à-dire caractérisés par une liberté d'entrée et de sortie.

Principe, hypothèses et limites théoriques

Ces auteurs ont montré que sur de tels marchés, les prix ne peuvent s'éloigner du coût minimum. En effet, une [19] entreprise dont les prix seraient au delà du coût minimum serait sous la menace de l'incursion (« *hit and run entry* ») d'un concurrent potentiel. Celui-ci pourrait rapidement pénétrer le marché, y réaliser des profits en s'attirant la demande, avant de se retirer lorsque les prix de l'entreprise en place diminuent. L'éventualité même d'une telle incursion doit suffire à inciter l'entreprise en place, à agir comme dans un environnement concurrentiel. Cet intéressant principe théorique mobilise cependant quelques hypothèses conséquentes :

- **La libre entrée-sortie du marché.** Cette hypothèse comprend, d'une part, l'absence de barrières à l'entrée, et d'autre part, l'absence de coûts irrécouvra-

[18] L'une d'elles ayant d'ailleurs commandé à l'IDEI les rapports de Seabright et de Villemeur que nous avons cités, et qui sont sans ambiguité sur le sujet.

[19] Nous considérons ici un marché initialement occupé par une seule entreprise, compte tenu de l'objet qui nous intéresse, mais la théorie des marchés contestables s'applique également à des marchés occupés par plusieurs entreprises.

bles (*sunk costs*), c'est-à-dire l'absence de barrières à la sortie. Si l'entrant peut avoir à supporter des coûts fixes avant de mener son incursion, ceux-ci doivent pouvoir être recouvrés lors de sa sortie du marché (diminués de la part correspondant à la dépréciation du capital engagé). Les barrières à l'entrée, quant à elles, peuvent prendre de multiples apparences. Les barrières d'ordre réglementaire, administratif et technique peuvent être supposées levées par une politique pro-concurrentielle. Mais les barrières à l'entrée peuvent aussi se manifester de façon stratégique, sous la forme d'une prolifération des produits, telle que l'entreprise en place occupe la totalité de l'espace de vente. Elles peuvent également apparaître à travers les choix d'investisement de l'entreprise en place. Celle-ci peut en effet dissuader l'entrée en augmentant sa capacité de production, de façon à rendre crédible l'idée d'une concurrence féroce en cas d'incursion sur le marché. Enfin, de nombreuses autres barrières à l'entrée, basées sur les effets de réputation de l'entreprise en place, sont également possibles.

- **Des temps de réactions variables entre les différents acteurs.** L'entrant doit pouvoir pénétrer le marché et s'attirer la majorité des consommateurs avant que l'entreprise en place ait eu le temps de réagir, alors même qu'il débute une activité pour laquelle sa notoriété est nulle. Cette hypothèse est notamment remise en cause par l'inertie des consommateurs, modélisée par les *switching costs* dont nous avons déjà discuté.
- **L'absence de coût d'apprentissage.** Cette hypothèse découle de celle suivant laquelle les concurrents potentiels pourraient produire rapidement avec la même structure de coût que l'entreprise en place sur le marché.

À ces hypothèses peu réalistes, s'ajoutent un certain nombre de limites, notamment la non-soutenabilité du monopole naturel, qui peut conduire à l'absence d'équilibre sur un marché contestable. Enfin, dans l'hypothèse où la concurrence potentielle peut être appliquée à certains marchés, ceux-ci doivent par ailleurs être régulés. Baumol, Panzar et Willig ne préconisent d'ailleurs pas la disparition de la régulation : « La théorie du marché contestable n'a pas pour objectif de soutenir les thèses de ceux qui pensent que le pur laissez-faire résout automatiquement tous les problèmes économiques, et que toute activité anti-trust ou régulatrice constitue une source coûteuse d'inefficience économique. » Ainsi, pour que la menace d'incursion sur le marché soit crédible, il convient d'aider les nouveaux entrants à pénétrer le marché. Cette assistance peut prendre diverses formes : de l'aide financière directe, à la réduction des coûts de changement, en passant par des restrictions sur la tarification de l'entreprise en place, *etc.* Voyons à présent comment ces divers résultats théoriques s'appliquent aux chemins de fer.

Des marchés ferroviaires peu contestables

La théorie des marchés contestables a été développée, afin de légitimer l'existence de quasi-monopoles sur des marchés concurrentiels. Elle présente donc l'intérêt d'être compatible avec la présence d'économies d'échelle. Mais les limites théoriques que

nous avons relevées tendent à s'opposer à l'application de cette forme de concurrence aux chemins de fer.

L'hypothèse d'absence de barrière à l'entrée est encore relativement acceptable. Une réglementation très importante peut garantir l'accès aux facilités essentielles. Elle ne pourra toutefois garantir que l'opérateur historique n'adoptera pas un comportement stratégique propre à dissuader l'entrée sur le marché [20].

En revanche, l'hypothèse d'absence de coûts irrécouvrables est peu acceptable malgré l'interopérabilité croissante des matériels roulants. Ces coûts peuvent être induits par l'obtention de la licence et des droits d'accès à l'infrastructure, le recrutement et la formation de personnels pour chaque ligne, la publicité ou encore, l'établissement d'un système de distribution de titres de transport...

Mais c'est surtout l'hypothèse relative au déroulement chronologique de l'incursion qui paraît improbable sur les marchés ferroviaires. La récente ouverture des réseaux à la concurrence dans le transport de marchandises montre qu'il faut plus d'un an entre l'intention de pénétrer le marché et l'entrée effective sur ce dernier.

Enfin, compte tenu de l'importance des coûts fixes à mobiliser pour entrer sur le marché, aucune entreprise ne peut crédiblement menacer un opérateur historique, sur une part significative de ses trafics, même en cas d'abus généralisé. Aussi, la concurrence potentielle qui est peut-être adaptée aux marchés du transport aérien, s'avère imparfaite pour discipliner des entreprises ferroviaires.

Conclusion : une régulation indispensable

Ce chapitre nous a permis de comprendre comment diverses réglementations ont pu, peuvent ou pourront être mises en œuvre plus ou moins adéquatement, compte tenu des caractéristiques des chemins de fer. Après avoir rappelé les différentes réglementations appliquées jusqu'à la concentration des réseaux de chemins de fer en monopole publics intégrés, nous avons justifié l'actuelle libéralisation par les limites de ces monopoles insuffisamment régulés. Considérant les différents marchés ferroviaires, nous avons étudié dans quelle mesure l'analyse économique propose des formes de concurrence susceptibles de discipliner ces marchés. Il en est ressorti que la concurrence intermodale n'agit que de façon limitée, mais que la concurrence pour le marché est très pertinente dès qu'il y a monopole naturel. Ainsi, elle est tout à fait adaptée pour réglementer les marchés de transport régional de voyageurs, ou encore, certaines activités liées à l'infrastructure. À plus grande échelle, elle s'avère peu compatible avec les effets de réseau qui commandent de ne pas allotir les services longue distance. Nous avons cependant noté que cette forme de concurrence ne dispensait pas, au contraire, d'une régulation conséquente des marchés associés. Nous avons ensuite expliqué pourquoi les formes de concurrence reposant sur le libre accès au marché n'étaient pas à même de réduire le pouvoir de marché des opérateurs historiques.

[20] En ce sens, il n'est pas inutile de s'interroger sur les motivations des opérateurs historiques à investir considérablement dans le matériel roulant à grande vitesse.

Quelle que soit la réglementation retenue pour ces marchés, monopole ou une certaine forme de concurrence, il paraît nécessaire de développer une régulation importante, afin de discipliner l'opérateur en position de monopole, et de favoriser le jeu efficace de la concurrence. Nous concluons cette première partie en nous concentrant sur cette première espèce de régulation : comment discipliner une entreprise en situation de monopole.

Conclusion : Comment réguler de façon incitative ?

Cette première partie de notre travail nous permet de conclure les résultats suivants :

- D'un point de vue normatif, il faut réguler les activités ferroviaires (en raison des défaillances des marchés concernés).
- Par le passé, la régulation des chemins de fer s'est avérée difficile, notamment à cause de la capture du gouvernement.
- Bien qu'elle atténue le pouvoir de marché des monopoles, l'actuelle libéralisation ne dispense pas, au contraire, de réguler les marchés.

La nécessité d'améliorer la régulation des chemins de fer est donc établie. Il reste à préciser la forme du mécanisme susceptible d'inciter efficacement à la réduction des coûts.

Traditionnellement, l'économie de la réglementation était structurée autour des paradigmes de la régulation *cost plus* et de la tarification Ramsey-Boiteux. Laffont (1995) [33] rappelle que la tarification Ramsey-Boiteux requiert une telle quantité d'information sur les coûts et la demande, qu'il est très difficile de la mettre en œuvre. La tarification *cost plus* prévoit que les prix sont alignés sur les coûts moyens, *plus* une marge pour rémunérer le capital. À l'inverse de la tarification Ramsey-Boiteux, celle-là s'avère plus réaliste. Mais son coût élevé d'application (en raison des faibles incitations qu'elle fournit) est un inconvénient majeur.

Plus récemment, des mécanismes de type *price cap* sont apparus. Ils consistent à fixer un prix plafond à l'entreprise régulée, sans que ses coûts soient contrôlés ultérieurement. L'entreprise bénéficie alors des diminutions de coût qui résultent de ses efforts. Ce mécanisme incitatif favorise donc la mise en œuvre par l'entreprise du niveau d'effort optimal. Mais en intégrant les contraintes informationnelles de la nouvelle économie de la réglementation, il apparaît que ce résultat a également un coût. Celui-ci se manifeste sous la forme d'un renchérissement du prix plafond, permettant à l'entreprise de dégager une rente.

Au début des années 1980, la nouvelle économie de la réglementation a introduit l'asymétrie d'information entre le régulateur et les entreprises. Il s'agit simplement de tenir compte du fait que l'entreprise intervenant sur le marché dispose d'un avantage informationnel par rapport au régulateur. Cette asymétrie d'information à laquelle le régulateur est confronté intervient suivant les deux problèmes classiques de la théorie

des incitations :

- **L'aléa moral** se manifeste lorsque l'incertitude du régulateur porte sur des variables *endogènes*. L'idée sous-jacente est que l'entreprise peut agir sur son niveau d'effort de façon opportuniste, sans être détectée par le régulateur. Ce phénomène est susceptible de se manifester dès qu'un aléa survient postérieurement à la conclusion du contrat d'agence.
- **La sélection adverse** apparaît lorsque l'entreprise dispose d'informations sur des variables *exogènes* que le régulateur n'observe qu'imparfaitement. Notamment, l'entreprise connaît sa productivité et les technologies qu'elle utilise, mieux que le régulateur.

Ces contraintes informationnelles permettent à l'entreprise d'obtenir des rentes de la part du régulateur. Dans le cas d'un mécanisme incitatif *price cap*, par exemple, le régulateur qui ignore la productivité de l'entreprise, peine à fixer correctement le prix plafond. L'asymétrie d'information qui s'oppose à la mise en œuvre de systèmes incitatifs doit donc être réduite.

Dans la prochaine partie, nous montrons comment des mécanismes de concurrence par comparaison permettent au régulateur de réduire ce désavantage informationnel. Plus généralement, nous analysons comment la comparaison de signaux émis par des entreprises similaires peut améliorer la régulation des monopoles.

Deuxième partie

Analyse théorique de la concurrence par comparaison

Introduction : Genèse de la concurrence par comparaison

> « Finalement, introduire une corrélation entre les caractéristiques des firmes ouvre le domaine encore largement négligé de la « *yardstick competition* » ou concurrence par comparaison. »
>
> Jean-Jacques Laffont (1995),
> « La nouvelle économie de la réglementation dix ans après »,
> *Revue d'Économie Industrielle*, n° spécial.

Les mécanismes comparatifs ont d'abord été étudiés dans le cadre de la théorie des incitations, en présence d'aléa moral. Progressivement, l'idée est apparue de transposer ces réflexions, du cadre d'agence standard à la régulation des monopoles. Simultanément, les États-Unis mettaient en œuvre un mécanisme comparatif incitatif, afin de réguler les soins hospitaliers des personnes âgées. La proposition de « concurrence par comparaison » de Shleifer, en 1985, résulte de cette conjonction. Mais c'est seulement à partir des années 1990 que les théoriciens de la nouvelle économie de la réglementation ont véritablement impulsé l'analyse de cette forme de régulation.

La nouvelle économie de la réglementation se démarque alors de l'économie publique traditionnelle, dans la façon d'appréhender la régulation. Cette innovation théorique s'est manifestée à travers :

- la formulation des problèmes de régulation sous la forme d'une relation principal-agent ;
- l'importance accordée à l'asymétrie d'information ;
- la prise en considération des incitations du régulateur, potentiellement vénal et opportuniste.

De ce fait, et bien qu'elle ait peu contribué aux débuts de la concurrence par comparaison, la nouvelle économie de la réglementation a largement favorisé le développement de cette forme de régulation. Cela se conçoit aisément, dans la mesure où les comparaisons de signaux émis par des entreprises similaires sont susceptibles de réduire l'insuffisance informationnelle du régulateur.

Les travaux de recherche ont alors consisté à développer des modèles articulés autour d'une double logique. D'une part, dans une logique positive, ils fournissent un fondement normatif aux formes de régulation comparative. D'autre part, dans

une démarche purement normative, ils intègrent différents types de contraintes théoriques.

Dans cette partie, nous nous concentrons sur l'analyse théorique des modèles de concurrence par comparaison. Notre travail met l'accent sur les différents modèles de la littérature. Nous détaillons particulièrement ceux qui sont les plus susceptibles de contribuer significativement à une mise en œuvre ultérieure. Cette démarche nous permettra de dégager les principales contributions à la théorie de la concurrence par comparaison. Elle facilitera aussi la compréhension de l'analyse, en précisant à quel auteur correspond tel cadre de raisonnement et les hypothèses associées.

Dans un premier chapitre, nous analysons l'intérêt des mécanismes comparatifs en fonction des contraintes informationnelles. L'asymétrie d'information étant le principal obstacle à une régulation incitative, il importe d'abord de confirmer le bénéfice informationnel que le régulateur peut tirer des comparaisons. Dans un second chapitre, nous élargissons le champ d'analyse. Nous étudions la robustesse des mécanismes comparatifs en intégrant des contraintes supplémentaires, en plus des contraintes informationnelles. Il s'agit alors de vérifier la pertinence de la concurrence par comparaison dans un cadre de régulation plus réaliste.

Chapitre 3

Analyse unidimensionnelle : les contraintes informationnelles

Nous consacrons ce premier chapitre sur la concurrence par comparaison à l'analyse des différents mécanismes proposés selon les contraintes informationnelles. L'objectif de ce chapitre consiste à illustrer par quels mécanismes un usage pertinent des comparaisons entre plusieurs agents permet d'améliorer la régulation, par rapport au cas individuel. Nous verrons notamment comment la concurrence par comparaison est susceptible de réduire l'asymétrie d'information, dont on a vu qu'elle s'oppose à une régulation efficace, dans le cadre individuel classique. Nous avons choisi de rendre compte de ces travaux en suivant l'ordre chronologique de leur diffusion. Une telle présentation permet de resituer les modèles de façon cohérente, dans le cadre plus général de la modélisation de la régulation du moment. Elle offre également l'opportunité de mieux articuler les différents travaux entre eux, et de repérer les influences des uns sur les autres. Enfin, elle rend compte des progrès de l'analyse économique sur une thématique bien précise, entre 1980 et 2003.

Dans une première section, nous examinons l'apport informationnel des comparaisons dans une configuration d'aléa moral. Ces premiers travaux, qui datent du début des années 1980, marquent l'émergence des pratiques comparatives dans un autre cadre que celui de la régulation. Nous examinerons ensuite un mécanisme comparatif élémentaire sous la forme des tournois de rang. Dans la section 3.3, nous aborderons le modèle originel de Shleifer (1985), qui a été le premier à théoriser les mécanismes de concurrence par comparaison, dans son article « *A theory of yardstick competition* ». Ce modèle très simple nous incite à étudier l'impact de l'hétérogénéité entre les entreprises comparées, et les moyens de l'atténuer. Dans la section suivante, nous analyserons les modèles de la littérature en présence de sélection adverse. Nous montrons comment les comparaisons permettent d'améliorer les solutions mono-agent sous cette forme d'asymétrie d'information. Enfin, nous achevons cette analyse par les modèles les plus récents dans une configuration de double asymétrie informationnelle (aléa moral *et* sélection adverse).

3.1 L'émergence des pratiques comparatives dans un cadre d'aléa moral

La réflexion concernant l'apport informationnel des comparaisons, dans un cadre d'aléa moral, a été développée au début des années 1980, dans le prolongement des travaux fondateurs de Mirrlees [1]. Les travaux dont nous rapportons les résultats dans cette section, ont été menés entre 1980 et 1984, avant que Shleifer ne publie sa « *theory of yardstick competition* ». Il s'agit de Baiman & Demski (1980) [60], d'Holmström (1982) [71] et de Mookherdjee (1984) [81]. À la différence de ce dernier, les autres auteurs inscrivent leur analyse dans le cadre de la relation principal-agent, appliquée aux relations managériales au sein des entreprises [2].

Nous commençons cette section par la présentation du cadre de référence de l'aléa moral appliqué aux contrats individuels, auquel nous comparerons les résultats obtenus en ayant recours aux comparaisons. Nous analyserons ensuite le modèle d'Holmström (1982) [71]. Ce modèle nous semble être le plus général parmi ceux qui traitent de l'usage comparatif de l'information dans le cadre d'aléa moral ; il s'agit en effet du seul modèle multi-agent qui ne se restreigne pas aux tournois de rang [3].

3.1.1 Les modèles d'aléa moral

Afin d'introduire le modèle d'Holmström, nous rappelons succinctement [4] la problématique générale de l'aléa moral, puis la référence des contrats individuels de la relation d'agence classique. Nous présentons ensuite le principe général sur lequel reposent les modèles comparatifs en aléa moral.

La problématique de l'aléa moral

L'aléa moral se manifeste dès lors que le principal est confronté à une incertitude sur le comportement de l'agent dans le jeu qui les lie tous les deux. Ce jeu peut être décrit de la façon suivante :

1. le principal propose à l'agent de produire une quantité (monétaire) de produit x en échange d'un salaire w ;
2. la fonction d'utilité du principal correspond au bénéfice qu'il retire de l'opération $B(x-w)$; elle est supposée concave ;
3. l'agent décide d'une action a (son niveau d'effort), afin de réaliser sa tâche ; cette action n'est pas observable sans coût par le principal [5] ;

[1] Mirrlees a formalisé le contexte d'aléa moral en 1975 dans « The Theory of Moral Hazard and Unobservable Behavior, Part 1 ».

[2] C'est la raison pour laquelle nous employons dans cette section les mêmes termes génériques (de la relation d'agence) que ceux de la littérature étudiée.

[3] Mookherjee (1984) [81] qui établit un modèle du même type qu'Holmström ne considère qu'un contexte bi-agent. Le cas particulier des tournois de rang sera traité dans la section suivante.

[4] Pour une analyse plus détaillée, voir Laffont & Martimort (2002) [74].

[5] Ce niveau d'effort correspond à la réduction de l'excédent organisationnel de l'entreprise, voir page 53.

4. la fonction d'utilité de l'agent est supposée additive et séparable [6] en une utilité au salaire $u(w)$, concave ; et une désutilité à l'effort $v(a)$, convexe : $U(w, a) = u(w) - v(a)$;
5. la production x observée par le principal et l'agent résulte de l'action décidée par l'agent et d'un aléa θ (perturbation stochastique) [7].

L'asymétrie d'information intervient ici au niveau de l'action décidée par l'agent, que le principal ne peut observer directement sans coût. En effet, le principal n'observe que le résultat de cette action (la production de l'agent), lequel dépend de l'action choisie par l'agent, mais aussi de l'aléa, à l'origine de l'incertitude.

De ce fait, le contrat qui lie le principal à l'agent ne peut spécifier le niveau d'effort requis, puisque celui-ci n'est pas vérifiable *ex post*. De là, l'agent est à même d'adopter un comportement opportuniste, en minimisant son niveau d'effort. Cela s'oppose évidemment à l'intérêt du principal en faveur d'une solution plus productive. Il faut donc fournir des incitations à l'agent, afin de le contraindre à choisir un niveau d'effort suffisant pour assurer, en espérance, une production satisfaisante pour le principal. Ce faisant, la mise en œuvre de ces incitations soulève la question du partage du risque induit par l'aléa, entre le principal et l'agent (il s'agit du risque que la production soit faible en dépit d'un effort élevé, du fait d'un aléa défavorable).

Dans ce cadre d'aléa moral, le principal est classiquement suposé neutre au risque ; l'agent en revanche peut être averse au risque. Dans les paragraphes qui suivent, nous rappelons les différentes allocations théoriques, dans le cadre des contrats individuels entre un principal et plusieurs agents identiques, selon que les agents sont neutres au risque ou non.

La référence des contrats individuels en aléa moral

Lorsque l'information est parfaite, c'est-à-dire que le principal peut observer l'action des agents, il existe une solution Pareto-optimale, quelle que soit l'aversion au risque des agents. Cette solution peut être atteinte, en offrant aux agents une rémunération ajustée sur le niveau d'effort requis par le principal :

- Si les agents sont neutres au risque, il suffit [8] de leur offrir un salaire fonction de la production. Cette rémunération laisse l'incertitude sur la production finale à la charge des agents (qui y sont indifférents).
- Si les agents sont averses au risque, le principal leur offre alors un salaire fixe, qui est la somme monétarisée de leur utilité de réserve et de leur désutilité à l'effort. Cette rémunération assure alors les agents contre le risque.

[6] Cette hypothèse classique des modèles principal-agent est justifiée par les opportunités de calcul relativement simple qu'elle offre. Économiquement, cette hypothèse mathématique est tout à fait acceptable, bien qu'elle implique une désutilité à l'effort indépendante du niveau de revenu, ce qui n'est pas systématiquement vérifié.

[7] Nous verrons que certains auteurs supposent que l'agent observe un signal sur l'aléa au moment de décider de son action ; cette configuration du jeu correspond à de la sélection adverse *ex ante*.

[8] Il n'y a pas unicité.

Lorsque l'information est imparfaite, le principal qui n'observe plus l'action de l'agent doit inciter celui-ci, pour l'empêcher d'adopter un comportement opportuniste. Si les agents sont neutres au risque, on retrouve la solution de premier rang précédente, qui incite les agents à l'effort, et leur fait supporter le risque auquel ils sont indifférents. En revanche, si les agents sont averses au risque, le principal doit arbitrer entre les incitations à l'effort, et l'assurance contre le risque. Les premières font supporter le risque aux agents, ce dont le principal doit les dédommager par un transfert d'autant plus élevé que les agents sont averses au risque. Inversement, l'assurance des agents contre le risque par une rémunération forfaitaire, supprime toute incitation à l'effort. Dès lors, le principal rémunère les agents avec une part fixe d'assurance contre le risque, et une part variable d'incitation à l'effort. Mais cette solution de second rang n'est pas aussi efficiente que les précédentes.

Ainsi, en présence d'agents averses au risque, le contexte d'aléa moral ne permet plus d'atteindre l'optimum de premier rang. L'objectif de cette section consiste donc à établir les conditions dans lesquelles un mécanisme comparatif est susceptible d'être aussi efficient que les contrats individuels (lorsque les agents sont neutres au risque), ou même plus efficient, lorsque ces contrats n'offrent qu'une solution de second rang (agents averses au risque).

L'apport du contexte multi-agent

Dans un contexte multi-agent, les auteurs dont nous étudions les modèles ont montré qu'un certain usage des comparaisons permet de réduire l'aléa, et donc le risque qu'il engendre, par rapport au cas des contrats individuels. L'idée sous-jacente consiste à interpréter les productions des agents comme des signaux renseignant (imparfaitement) sur les actions qu'ils ont choisies. L'intérêt des comparaisons par rapport aux contrats individuels, est qu'elles permettent justement d'extraire de ces signaux une information plus riche.

La prise en compte de l'information comparative dans le cadre d'aléa moral, mobilise deux hypothèses, en plus de celles du contexte individuel :

- les agents sont identiques *ex ante*, de sorte que les équilibres sont supposés symétriques [9] ;
- les perturbations stochastiques, qui affectent l'effort des différents agents, sont corrélées [10].

L'introduction de cette dernière hypothèse est assez intuitive, à partir du moment où le principal cherche à induire un effet compétitif, en comparant les productions de ses agents. En effet, la production étant une variable aléatoire, on pressent bien que la comparaison de deux productions non corrélées conduit à augmenter le niveau de bruit, et donc le risque. Or, c'est précisément l'inverse qui nous intéresse : la réduction de l'incertitude, et donc la réduction de l'assurance contre le risque à

[9] Le concept d'équilibre retenu est donc celui de Nash : chaque agent maximise son utilité en réponse au même comportement supposé des autres agents.

[10] Dans les modèles que nous analysons, Mookherjee (1984) [81] précise d'ailleurs que cette corrélation doit être linéaire.

fournir. En termes mathématiques, cette hypothèse de corrélation est traduite par la décomposition de l'aléa θ_i, rencontré par l'agent i, en deux éléments :

- une composante commune, η ;
- une autre composante individuelle, ϵ_i.

On a donc $\theta_i = \theta_i(\eta, \epsilon_i)$ [11]. Ainsi, dans un contexte d'aléa moral multi-agent, l'intérêt majeur des comparaisons consiste à réduire le risque, et à fournir de meilleures incitations, grâce à l'exploitation de l'information véhiculée par la production des autres agents. Nous consacrons la fin de cette section à illustrer comment le mécanisme proposé par Holmström, permet d'atteindre cet objectif.

3.1.2 Le modèle d'Holmström (1982)

Holmström (1982) [71] est le seul auteur à développer un modèle multi-agent, hors du cas particulier des tournois de rang. Mookherjee (1984) [81] apporte toutefois quelques compléments, dans le cadre de son modèle général bi-agent. Holmström montre d'abord qu'en présence d'agents neutres au risque, les mécanismes de rémunération multi-agents permettent d'atteindre l'optimum de premier rang, comme les contrats individuels [12]. Dans ce cas, le mécanisme d'évaluation relative peut néanmoins être préférable, en raison des coûts de transaction réduits de sa mise en œuvre. Il est satisfaisant de vérifier que le recours aux comparaisons, dans ce cas particulier d'agents neutres au risque, permet toujours d'atteindre la solution de premier rang. Toutefois, dans le cas d'agents averses au risque, les mécanismes comparatifs pourraient s'avérer significativement intéressants s'ils permettent d'améliorer la solution de second rang des contrats individuels. C'est la raison pour laquelle nous nous concentrons sur ce cas d'agents averses au risque.

Résultats généraux pour des agents averses au risque

Les hypothèses essentielles sont celles que nous avons décrites dans les paragraphes précédents ; l'agent décide notamment de son niveau d'action sans avoir connaissance de l'aléa. De façon générale, Holmström suppose d'abord que le principal n'observe qu'un signal y de la production x des agents.

Holmström élabore ses résultats en s'appuyant sur la notion de *statistique suffisante* [13], qui établit les conditions dans lesquelles il est possible de filtrer des données, sans perte d'information. Holmström montre alors que les schémas de rémunération multi-agents s'appuyant sur une statistique suffisante des observations y sont dominants au sens de Pareto sur tous les autres (théorèmes 5 et 6). Il en résulte donc,

[11] Nous reprenons dans toute cette section ces notations qui sont celles d'Holmström (1982) [71].

[12] Les auteurs qui ont travaillé sur les tournois de rang parviennent également à la même conclusion.

[13] Une statistique $T_i(y)$ est dite suffisante par rapport à y si et seulement si, pour chaque observation y, on peut factoriser la densité conditionnelle de la façon suivante : $g(y|a) = p_i(T_i(y)|a).h_i(y|a_{-i})$, pour chaque valeur de a, où $h_i(y|a_{-i})$ ne dépend pas de a_i. Dans ce cas, le rapport de vraisemblance obtenu en faisant varier a_i dépend des observations y uniquement à travers la statistique suffisante : $g(y|a_i^\alpha, a_{-i})/g(y|a_i^\beta, a_{-i}) = p_i(T_i(y)|a_i^\alpha, a_{-i})/p_i(T_i(y)|a_i^\beta, a_{-i})$.

compte tenu de la définition d'une statistique suffisante, que « l'aléa n'apporte rien », dans la mesure où il est possible d'isoler toute l'information pertinente sur l'agent i.

Holmström considère ensuite que la production de chaque agent est observable. Dans le cas où les productions des différents agents sont indépendantes les unes des autres, il montre alors (théorème 7) que l'optimum est atteint par un ensemble de contrats individuels. Dans ce cas d'absence de corrélation, il n'y a aucun intérêt à inciter les agents à se faire concurrence, les comparaisons ne faisant qu'ajouter de l'incertitude [14]. Moyennant l'hypothèse de monotonie de la production par rapport à l'aléa [15], Holmström démontre également la réciproque : si les contrats individuels sont optimaux, alors il n'y a pas de corrélation entre les productions des agents. Cette condition nécessaire d'optimalité est discutée par Mookherjee qui souligne le recours nécessaire à une hypothèse supplémentaire [16], sans laquelle il existe des cas où les contrats individuels peuvent être optimaux sans qu'il y ait indépendance des productions. Pour autant, Mookherjee reconnaît que les cas de violation de cette condition nécessaire sont très particuliers.

Le cas particulier d'une fonction de production linéaire

À la suite de ces démonstrations générales, Holmström illustre et approfondit ses résultats par l'étude de deux cas particuliers de fonctions de production, avec un aléa $\theta_i = \eta + \epsilon_i$:

$$\begin{aligned} x_i(a_i, \theta_i) &= a_i + \eta + \epsilon_i \\ x_i(a_i, \theta_i) &= a_i(\eta + \epsilon_i) \end{aligned}$$

Nous présentons ici les résultats associés à la première spécification de la fonction de production qui est linéaire en action et en aléa. Holmström suppose que η et les ϵ_i sont indépendants et normalement distribués. Il démontre alors qu'il existe une statistique suffisante comparative, permettant de définir la rémunération optimale des agents. En l'occurence, cette statistique suffisante pour chaque agent i fait intervenir la production de l'agent, x_i, et la moyenne des productions, pondérée par la précision relative des aléas individuels $\bar{x} = \sum_i x_i \tau_i / \bar{\tau}$ (où τ_i est la précision - l'inverse de la variance - de ϵ_i). Cette mesure agrégée de la production, sous forme de moyenne pondérée des productions individuelles, capture ainsi toute l'information pertinente sur l'aléa commun η. L'idée sous-jacente consiste à considérer que « si ϵ_j a une grande précision (faible variance), alors x_j informe plus précisément sur η et doit donc être davantage pondéré dans la moyenne. Cela revient à dire que les x_j fortement corrélés à x_i devraient être des indicateurs plus significatifs dans l'évaluation de la performance de l'agent i. Inversement, lorsque $\tau_j \to 0$ x_j informe très peu sur η à cause du bruit de ϵ_j, et donc ne devrait pas trop être pris en compte.

[14] Les auteurs qui ont travaillé sur les tournois de rang montrent également que l'absence de corrélation (d'aléa commun) favorise les contrats individuels par rapport aux schémas multi-agents.

[15] C'est-à-dire, $\partial x_i / \partial \theta_i$ monotone.

[16] Il est nécessaire que la fonction de production conditionnelle $f_i(x_i|a_i)$ soit inversible en l'aléa.

Holmström analyse ensuite les résultats asymptotiques lorsque le nombre d'agents augmente. Il montre (théorème 9) que la solution optimale du programme multi-agent converge alors vers le vecteur des solutions optimales des programmes individuels, en absence d'aléa commun. Considérant, à titre illustratif, les cas particuliers des fonctions de production que nous avons exposées, Holmström prouve que la valeur de η peut être inférée *ex post*, à partir des observations indépendantes que constituent les x_i. Cela permet donc de réduire le risque θ_i de $\eta + \epsilon_i$ à ϵ_i [17], ce qu'illustrent les graphiques de la figure 3.1, pour différentes valeurs du rapport var(ϵ)/var(η).

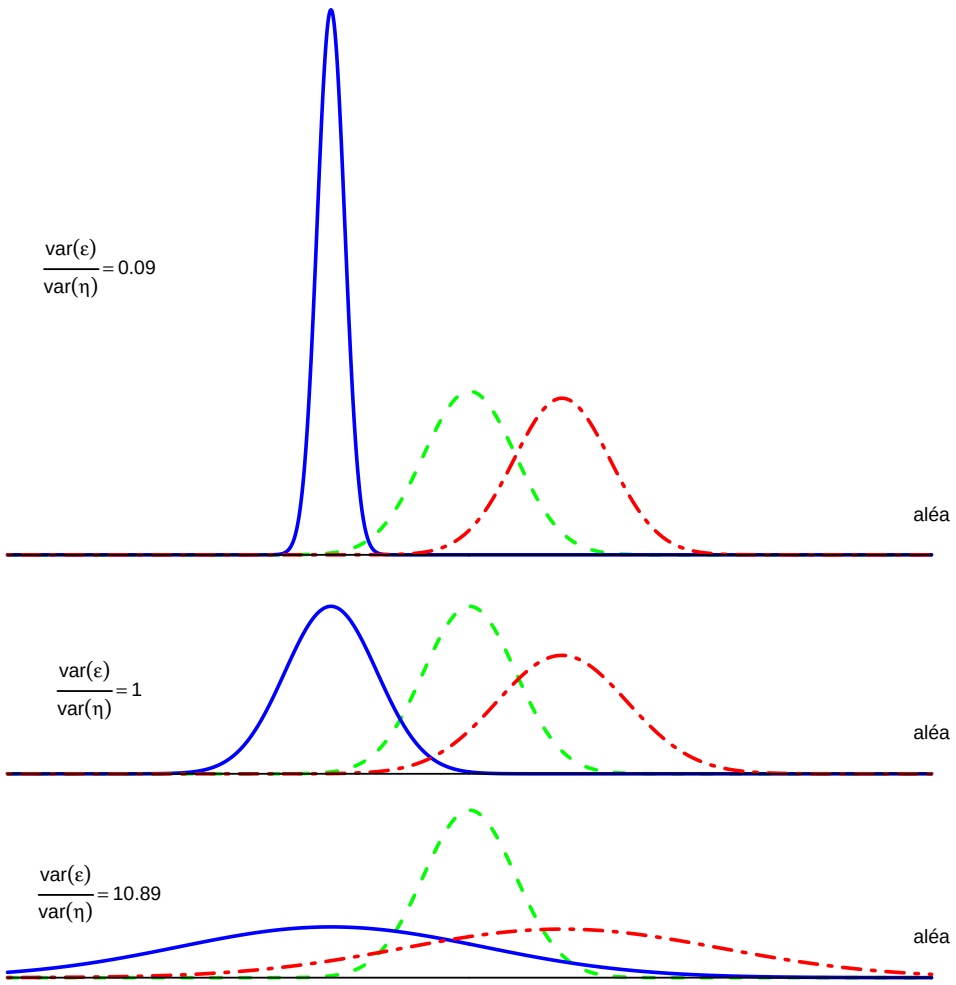

FIG. 3.1 – Distribution de l'aléa θ et de ses composantes η et ϵ : variation de l'apport des comparaisons en fonction du rapport var(ϵ)/var(η) - cas où var(η) est constante ; —— : ϵ ; - - - : η ; -·-·- : $\theta = \eta + \epsilon$.

[17] Soulignons que le risque véhiculé par les ϵ_i peut être réduit par ailleurs, grâce à des traitements économétriques adaptés, comme les frontières stochastiques, que nous détaillons dans le dernier chapitre, page 172.

Ce rapport de variances, qui traduit l'ampleur de l'aléa individuel relativement à celle de l'aléa commun, permet de caractériser l'apport des comparaisons, en présence de nombreux agents. Celui-ci est évalué par la réduction de l'incertitude, c'est-à-dire de la variance de l'aléa, lorsqu'on passe de θ à ϵ (sur la figure 3.1, cela correspond au passage de la courbe en pointillés mixtes à celle en trait plein). Mais ce rapport de variances renseigne aussi sur la corrélation des agents. En effet, le coefficient de corrélation des aléas auxquels deux agents sont confrontés [18] $R = \text{cov}(\theta_i,\theta_j)/(\text{var}(\theta_i)\text{var}(\theta_j))$ peut se mettre sous la forme :

$$R = \frac{1}{1 + \frac{\text{var}(\epsilon_i)+\text{var}(\epsilon_j)}{\text{var}(\eta)} + \frac{\text{var}(\epsilon_i)\text{var}(\epsilon_j)}{\text{var}^2(\eta)}}$$

Pour éclaircir l'analyse, supposons $\text{var}(\epsilon_i) \sim \text{var}(\epsilon_j) \sim \text{var}(\epsilon)$ et considérons les trois cas suivants illustrés par la figure 3.1 :

- $\mathbf{var(\epsilon) \ll var(\eta)}$ **:** on a alors $R \sim 1 - 2\,\text{var}(\epsilon)/\text{var}(\eta) \to 1$: la corrélation des agents est quasi-parfaite. En annulant *ex post* la part commune de l'aléa, le mécanisme comparatif réduit la quasi-totalité de l'incertitude et permet de tendre vers la solution de premier rang.
- $\mathbf{var(\epsilon) \sim var(\eta)}$ **:** on a alors $R \simeq 0,25$. En annulant *ex post* la part commune de l'aléa, le mécanisme comparatif réduit l'incertitude de moitié.
- $\mathbf{var(\epsilon) \gg var(\eta)}$ **:** on a alors $R \sim (\text{var}(\eta)/\text{var}(\epsilon))^2 \to 0$: les agents sont très peu corrélés. Dans ce cas, l'annulation *ex post* de l'aléa commun ne modifie pas sensiblement la solution des contrats individuels.

Ainsi, il apparaît que les comparaisons des productions des agents améliorent la solution des contrats individuels, grâce à un effet d'assurance sur le risque. Toutefois, l'ampleur de cette amélioration dépend encore de la corrélation entre les agents, c'est-à-dire des rapports $\text{var}(\epsilon_i)/\text{var}(\eta)$. La solution de premier rang n'est atteinte que pour $\text{var}(\epsilon_i) = 0, \forall i$. Dans son approche similaire, Mookherjee montre également que la solution du programme bi-agent converge vers la solution de premier rang, lorsque la corrélation des aléas tend vers 1 (en valeur absolue).

Enfin, Mookherjee pointe les deux limites principales des analyses multi-agents en aléa moral :

- l'éventuelle multiplicité des équilibres ;
- la menace de la collusion.

La première limite vient de ce que les différents modèles ignorent la possibilité qu'un équilibre différent puisse être choisi, lorsqu'il existe plusieurs équilibres. La seconde limite, que nous retrouverons dans la quasi-totalité des modèles de ce chapitre, sera discutée dans le chapitre suivant [19].

Cet examen approfondi du modèle d'Holmström permet de comprendre le mécanisme général d'apport informationnel dans un cadre d'aléa moral multi-agent. Pour

[18] Ce coefficient de corrélation peut être évalué empiriquement par le coefficient de corrélation des productions de ces mêmes agents, x_i et x_j, puisque les agents sont supposés identiques, et leurs actions non aléatoires.

[19] Bien que Holmström imagine les tournois de rang relativement robustes à la collusion, dans la mesure où il s'agit de jeux à somme nulle, Mookherjee prouve qu'il n'en est rien.

une structure particulière de la fonction de production, nous avons vu comment l'information comparative permet de réduire le risque supporté par les agents et, de là, d'améliorer l'efficacité de la solution. Dans la section suivante, nous montrons dans quelle mesure le cas particulier des tournois de rang permet de retrouver des résultats similaires.

3.2 Le cas particulier des tournois de rang

Les tournois de rang consistent à utiliser l'information sur la production observée des agents, pour ordonner ces derniers du plus productif au moins productif, et de les rémunérer de façon décroissante. Les tournois de rang, abordés par Holmström (1982) [71] et Mookherjee (1984) [81], ont été également étudiés plus en détail par Lazear & Rosen (1981) [77], Nalebuff & Stiglitz (1983) [83] et Green & Stokey (1983) [70]. Holmström montre qu'en général les tournois de rang ne reposent pas sur une statistique suffisante. Aussi, s'il est possible d'évaluer les productions des agents de manière cardinale, ce type de mécanisme est sous-optimal, en vertu de son théorème 6.

3.2.1 Principe général

Nous détaillons plus particulièrement les résultats obtenus par Nalebuff et Stiglitz. Notons que ces auteurs émettent l'idée de passer de comparaisons *au sein* des organisations à des comparaisons *entre* les organisations [20]. Précisons que le travail de Nalebuff et Stiglitz se situe dans un cadre d'aléa moral, doublé de sélection adverse. En effet, bien que l'effort demeure inobservable par le principal, la composante commune de l'aléa est connue avant que les agents ne choisissent leur action. Toutefois, cela ne modifie pas l'analyse qui peut être faite de leur mécanisme de tournoi [21].

Nalebuff et Stiglitz ont étudié l'impact d'un certain nombre de paramètres sur l'efficacité des tournois de rang, avant de comparer leur efficacité à celle d'autres mécanismes de rémunération. Ils envisagent un tournoi bi-agent à somme nulle. L'agent le plus productif gagne une rémunération Y_{W} égale au salaire moyen $\bar{Y}$ augmenté

[20] Nalebuff & Stiglitz (1983) [83] remarquent dans leur conclusion que « ces considérations (c'est-à-dire, l'environnement de travail, l'homogénéité des groupes) sont probablement moindres dans l'analyse de la concurrence entre les firmes par rapport à la concurrence au sein d'une firme. » Dans Nalebuff & Stiglitz (1983) [82], les mêmes auteurs proposent d'ailleurs explicitement de considérer un cadre de concurrence entre entreprises. Ils prennent l'exemple du cas où un gouvernement souhaite développer une bombe. Ils remarquent que les mécanismes de *price cap*, bien qu'incitatifs, deviennent très onéreux en raison du risque élevé supporté par l'agent. Les auteurs proposent alors de recourir aux comparaisons avec un marché duopolistique, sous une forme de *dual sourcing*.

[21] En effet, il s'agit d'un modèle de sélection adverse *ex ante* (voir Laffont & Martimort (2002) [74]) dont la problématique diffère du cas classique de sélection adverse (*ex post*). Il ne s'agit pas d'inciter l'agent à révéler sa caractéristique, puisqu'il l'ignore encore lors de la signature du contrat. Le principal doit en fait offrir une prime d'assurance contre le risque, au cas où l'agent averse au risque se découvre inefficient. Green & Stokey (1983) [70] considèrent, eux, le cas où les agents perçoivent chacun un signal différent de η avant d'engager leur action ; ils montrent (lemme 1) que le tournoi de rang optimal est indépendant de ce signal observé.

d'un *bonus* $X/2$. L'autre agent perçoit une rémunération Y_{L} égale à $\bar{Y}$ diminué d'un *malus* $X/2$. Ce mécanisme de tournoi incite les agents à fournir un effort, afin de maximiser leur probabilité d'emporter Y_{W}. Les agents étant identiques, l'équilibre est symétrique, et chaque agent a finalement une probabilité de 0,5 de remporter le *bonus* grâce à un aléa favorable.

On retrouve alors la discussion des paragraphes précédents, au sujet de l'apport des comparaisons en fonction du rapport $\mathrm{var}(\epsilon_i)/\mathrm{var}(\eta)$. Nalebuff et Stiglitz comparent la performance des tournois de rang aux contrats individuels. Leurs résultats sont semblables à ceux des autres auteurs, et cohérents avec ceux d'Holmström. Comme attendu, il apparaît que lorsque la composante commune de l'aléa est nulle, les contrats individuels dominent les tournois de rang (voir Lazear & Rosen (1981) [77] et Green & Stokey (1983) [70]). Inversement, la présence d'une composante commune favorise la solution des tournois de rang ; Nalebuff et Stiglitz montrent notamment que, pour des valeurs suffisamment grandes de $\mathrm{var}(\eta)$, les tournois de rang sont toujours dominants. Ils étudient ensuite quelques propriétés des tournois de rang.

3.2.2 Quelques résultats supplémentaires

Nalebuff et Stiglitz montrent d'abord que l'efficacité du tournoi est améliorée lorsque l'écart de prix entre les rémunérations des agents, X, augmente : l'intensité concurrentielle du tournoi s'accroît et les agents fournissent davantage d'effort. Ce faisant, la production totale augmente et le principal peut proposer des prix Y_{W} et Y_{L} relativement supérieurs à ceux du tournoi où X est faible.

Les auteurs examinent ensuite les effets de $\mathrm{var}(\eta)$, sans pouvoir les distinguer clairement. Dans le cas particulier où la fonction de désutilité à l'effort est quadratique ($v(a) = a^2$), ils montrent que le bien-être croît avec $\mathrm{var}(\eta)$.

Enfin, Nalebuff et Stiglitz proposent d'imposer un gap entre les niveaux de production permettant de déterminer le vainqueur du tournoi. Cette condition supplémentaire permet de réduire le risque supporté par les agents, en diminuant l'influence de l'aléa au moment de départager les agents : l'agent le plus productif remportera Y_{W} à condition d'avoir fait significativement mieux que son concurrent. Dans ce cas, cette performance doit alors relativement plus au niveau d'effort choisi qu'à l'aléa, par rapport à un tournoi sans gap. La présence d'un gap renforce donc les incitations du mécanisme.

Nalebuff et Stiglitz analysent également le problème de non-convexité, susceptible d'apparaître dans certaines configurations des tournois de rang. Il s'agit du fait qu'un agent peut être tenté de dévier de l'équilibre symétrique. Ce comportement apparaît si l'agent anticipe qu'en ne fournissant aucun effort, il s'épargne toute sa désutilité à l'effort, tout en percevant une rémunération acceptable. Ce phénomène peut survenir lorsque $\mathrm{var}(\epsilon)$ et l'écart de prix X sont faibles. En effet, dans ce cas, s'il veut avoir une chance de gagner un peu plus que Y_{L}, l'agent doit fournir un effort

concurrentiel, significatif. Une solution consiste alors à accroître l'écart de prix entre les agents, X.
Le problème de non-convexité peut aussi se manifester dans un tournoi multi-agent, dans la mesure où Nalebuff et Stiglitz proposent de rémunérer le vainqueur $\bar{Y}+(n-1)X$, et tous les autres agents $\bar{Y}-X$. Dans ce cas, la probabilité de gagner le premier prix étant faible $(1/n)$, les agents pourront préférer ne pas travailler et recevoir $\bar{Y}-X$. La solution consiste alors à imposer une pénalité à l'agent le moins productif, afin de dissuader un tel comportement déviant.

Au delà de ces résultats essentiels, Nalebuff et Stiglitz notent que les comparaisons « ont de plus l'avantage d'ajuster automatiquement les incitations aux changements de l'environnement économique ». Ils définissent cette propriété sous le terme de « flexibilité intégrée ». Cet apport supplémentaire de la régulation par comparaison n'est pas négligeable compte tenu de l'importance des coûts de transaction correspondant à une renégociation de contrat. Les auteurs soulignent que ces contrats, notamment les tournois de rang, sont moins coûteux à mettre en œuvre. D'une manière générale, la réduction des coûts de transaction est due aux moindres exigences informationnelles des contrats multi-agents, dès que les agents ne diffèrent que de l'aléa individuel qu'ils affrontent.

À l'issue de cette présentation de mécanismes comparatifs particuliers, nous analysons dans les deux sections suivantes l'apport des comparaisons dans d'autres configurations informationnelles. Dans un premier temps, le modèle de Shleifer nous amène à considérer des agents identiques, hypothèse que nous discuterons avant de la relâcher dans la section suivante portant sur les modèles en sélection adverse.

3.3 Le modèle originel de Shleifer et le problème de l'hétérogénéité

Dans cette section, nous étudions conjointement le modèle de Shleifer et la question du traitement de l'hétérogénéité dans les comparaisons. En effet, les hypothèses extrêmes de son modèle, en termes d'identité, conduisent Shleifer à discuter la question de l'hétérogénéité des entreprises comparées. Nous enchaînons donc à la suite de l'examen du modèle, l'analyse de l'hétérogénéité.

3.3.1 Le modèle de Shleifer (1985)

Shleifer n'est pas le premier, en 1985, à étudier l'apport informationnel des comparaisons de performance dans un contexte multi-agent. Pour autant, il semble justifié de lui attribuer l'origine des mécanismes de concurrence par comparaison dans la mesure où :

- il est le premier à appliquer un mécanisme comparatif à la régulation d'entreprises similaires ;
- il est à l'origine du terme « *yardstick competition* » qui a été ensuite repris par toutes les publications traitant de ce sujet.

Shleifer constate qu'en présence d'asymétrie d'information, les deux modèles dominants de régulation, *cost plus* et *price cap*, ne permettent pas d'atteindre la solution de premier rang du fait d'un manque, respectivement, d'incitations et d'information. En se fondant sur l'observation de mécanismes régulateurs qui ont recours aux comparaisons [22], Shleifer élabore un modèle très simple de concurrence par comparaison.

Shleifer considère $N \geqslant 2$ entreprises opérant sur des marchés géographiquement distincts. Il postule les hypothèses suivantes :

- les entreprises sont identiques ;
- il n'y a pas d'incertitude ;
- les entreprises font face à la même courbe de demande [23], dans un environnement identique (nous verrons comment relâcher cette hypothèse d'identité à la section suivante) ;
- chaque entreprise a un coût marginal initial c_0 qu'elle peut réduire à c_i en investissant $R(c_i)$ [24] ;
- le régulateur fixe les prix p_i et les transferts T_i en maximisant le surplus collectif.

Le déroulement du jeu est le suivant :

1. le régulateur annonce sa règle de fixation des prix et transferts ;
2. les entreprises investissent pour réduire leur coût ;
3. le régulateur observe alors les niveaux des c_i et $R(c_i)$ ce qui lui permet de déterminer les p_i et T_i suivant la règle annoncée ;
4. les entreprises produisent de façon à satisfaire la demande [25], vendent au prix fixé à l'étape précédente, et perçoivent le transfert annoncé.

Afin de fixer des prix et transferts incitatifs, en fonction des performances de chaque entreprise, le régulateur compare chaque entreprise à son *étalon*, entreprise jumelle de référence, défini par les performances suivantes :

$$\begin{aligned} \bar{c}_i &= \frac{1}{N-1}\sum_{j\neq i} c_j \\ \bar{R}_i &= \frac{1}{N-1}\sum_{j\neq i} R(c_j) \end{aligned}$$

[22] Le système américain d'assurance médicale pour les personnes âgées et handicapées, Medicare, a notamment inspiré Shleifer qui en expose le fonctionnement dans la dernière section de son article. Shleifer note également le recours au *dual sourcing* (duopole) du Département de la défense américain.

[23] Shleifer précise que cette hypothèse peut être relâchée, à condition de supposer que le régulateur connaît les courbes de demande de chaque marché.

[24] Cet investissement est assimilable à la désutilité à l'effort, $\psi(e)$, des modèles classiques de régulation.

[25] Elles ne peuvent pas y déroger en quittant le marché si le prix fixé par le régulateur à l'étape 3 ne leur convient pas.

Shleifer montre alors (proposition 1) que le régulateur peut atteindre l'optimum de premier rang en fixant :

- $\boldsymbol{p_i = \bar{c}_i}$ **:** le prix de vente de l'entreprise i est égal au coût marginal moyen des autres entreprises ;
- $\boldsymbol{T_i = \bar{R}_i}$ **:** le transfert forfaitaire versé à l'entreprise i est égal à la moyenne des investissements effectués par les autres entreprises pour réduire leur coût.

Notons que cette règle de tarification rompt la dépendance entre les prix et les coûts de chaque entreprise. Le prix étant désormais exogène, le monopole devient *price taker*, comme dans le cas de la concurrence pure et parfaite.

Shleifer démontre que l'optimum de premier rang est l'unique équilibre de Nash de ce jeu. C'est-à-dire que pour chaque entreprise, investir afin d'égaliser son coût marginal à l'effort, $R'(c_i)$, avec son gain marginal à l'effort, $q(\bar{c}_i)$, est la meilleure réponse aux stratégies des autres entreprises. En effet, le profit de i s'écrit : $\Pi_i = (p_i - c_i)q(p_i) + T_i - R(c_i)$, dont la maximisation fournit la condition de premier ordre $-R'(c_i) = q(\bar{c}_i)$ qui caractérise l'optimum de premier rang. On a donc $\forall i$, $c_i = c^*$. Une entreprise qui dévierait de cet équilibre, dans un sens ou dans l'autre, serait donc pénalisée, soit par un prix de vente insuffisant pour compenser ses coûts, soit par un transfert insuffisant pour compenser son investissement.

Mais ce résultat est conditionné par l'engagement crédible du régulateur à appliquer cette règle de tarification. Dans la mesure où il suppose que les entreprises ne peuvent se soustraire au jeu, une fois les prix et transferts fixés, le mécanisme repose sur la crédibilité du régulateur à s'engager à appliquer la tarification annoncée. Quand bien même les entreprises auraient fait un choix d'investissement inefficace, aux conséquences financières importantes, Shleifer précise que le régulateur ne doit pas céder aux menaces de faillite. Ainsi, le mécanisme induit un comportement concurrentiel entre les entreprises, chacune étant incitée à faire mieux que son étalon, c'est-à-dire que toutes les autres associées.

Shleifer propose donc un jeu relativement simple, permettant de réguler les entreprises à partir de l'observation *ex post* de leur performance. Le régulateur n'a qu'à tirer parti de ces informations, afin de mettre en œuvre le mécanisme. L'efficience du mécanisme est assurée par le fait que l'équilibre de Nash, qui oriente les entreprises, est justement l'optimum de premier rang. Ainsi, les entreprises sont *naturellement* incitées à se soumettre à cette forme de concurrence virtuelle.

Toutefois, ce résultat n'est atteint qu'au prix d'hypothèses très fortes. Nous discutons dans la section suivante la question de l'hétérogénéité des environnements. Shleifer montre que celle-ci ne réduit pas nécessairement l'efficacité du mécanisme, dans la mesure où elle peut être corrigée. Le relâchement de l'hypothèse d'absence d'identité des agents nous conduira, dans la section suivante, à envisager le problème de sélection adverse. Enfin, nous analyserons plus précisément au chapitre suivant la menace de la collusion et les questions d'engagement du régulateur, que Shleifer souligne au titre de limites théoriques du mécanisme.

3.3.2 Corriger l'hétérogénéité

L'hétérogénéité est susceptible d'apparaître dans les mécanismes comparatifs de diverses façons, avec des conséquences variables et donc, des solutions à apporter différentes. Il est possible de distinguer d'abord l'hétérogénéité *interne* de l'hétérogénéité *externe*. La première est propre aux agents, il s'agit du fait que certains sont plus productifs que d'autres. Cette hétérogénéité engendre un problème de sélection adverse dans la mesure où elle permet aux agents de cacher leur identité au principal. Nous traiterons donc de cette hétérogénéité interne dans la prochaine section, consacrée à la sélection adverse.

L'hétérogénéité externe qui affecte les coûts des entreprises, est susceptible de biaiser les étalons auxquels celles-ci sont comparées. Ce biais menace alors de distordre le mécanisme comparatif, et donc de réduire l'efficacité de la concurrence par comparaison. Il apparaît donc nécessaire de corriger cette source de distorsion qu'est l'hétérogénéité externe. Dans cette section, nous analysons donc les effets de cette hétérogénéité et les moyens de la corriger. Pour ce faire, nous distinguons encore l'hétérogénéité exogène au cadre de régulation de l'hétérogénéité endogène à ce cadre.

L'hétérogénéité exogène

L'hétérogénéité exogène au cadre de régulation est liée à l'environnement, variable d'une entreprise à l'autre. Elle est susceptible d'influer sur la performance des entreprises, selon que l'environnement est plus ou moins favorable à l'activité de celles-ci. En discutant du relâchement de son hypothèse d'identité des entreprises, Shleifer remarque que les distorsions induites par l'hétérogénéité peuvent être évitées. Il suffit que le régulateur observe parfaitement les différentes sources d'hétérogénéité et l'ampleur de leur impact sur les entreprises. Moyennant cette hypothèse, le régulateur peut alors corriger parfaitement l'hétérogénéité.

Pour ce faire, les informations comptables observées par le régulateur doivent donc être régressées sur les facteurs d'hétérogénéité par un traitement économétrique, avant d'être introduites dans le mécanisme d'étalonnage. Cette démarche nécessite d'estimer une fonction de coût incluant tous les facteurs d'hétérogénéité. Cela justifie que la concurrence par comparaison ne soit appliquée qu'à des entreprises très similaires, dans la mesure où :

- il n'est pas évident d'inclure tous les facteurs d'hétérogénéité, certains pouvant être difficilement mesurables ou quantifiables ;
- la forme mathématique retenue pour la fonction de coût influe sur les résultats de la correction.

Ainsi l'introduction d'une fonction de coût, dont la spécification et la forme sont nécessairement imparfaites, permet de réduire en partie l'impact de l'hétérogénéité. Bien que ce procédé ne conduise plus tout à fait à l'optimum de premier rang, il ne doit pas conduire pour autant à renoncer à la concurrence par comparaison ; Shleifer note ainsi que « La performance encourageante de la concurrence par comparaison

appliquée au système de remboursement par groupe de diagnostic de Medicare, en dépit des problèmes de diversité des entreprises, suggère que ces problèmes ne sont pas insurmontables. »

L'hétérogénéité endogène ou les « effets de dépendance spatiale »

L'hétérogénéité endogène au cadre de régulation est susceptible d'apparaître lorsque la régulation est décentralisée auprès de plusieurs principaux locaux (les P_i), contractant chacun avec un agent local A_i. Supposons qu'un principal P_j souhaite mettre en œuvre un mécanisme de concurrence par comparaison, en mesurant les performances de son agent A_j à celles d'autres agents $A_{i \neq j}$. Dans ce cas, le principal P_j risque de se heurter à un problème d'incomplétude informationnelle qui menace de distordre le mécanisme comparatif. En effet, il est peu probable que ce principal dispose de la même information concernant son agent A_j et les autres agents $A_{i \neq j}$.

Dans ce cadre, Bivand & Szymanski (1997) [62] montrent que si des principaux $P_{k \neq j}$ poursuivent une politique particulière inobservable, l'externalité informationnelle positive qu'ils sont supposés dégager devient négative et distord alors le contrat entre P_j et A_j. Les auteurs considèrent notamment le cas où un principal P_k, capturé, subventionne sa régie A_k de façon excessive, pour des raisons électoralistes. Bivand et Szymanski montrent que ce surcoût se propage alors avec la concurrence par comparaison, engendrant ce qu'ils appellent des « effets de dépendance spatiale. »

L'article de Bivand et Szymanski est intéressant, dans la mesure où les auteurs testent empiriquement leur modèle sur des données relatives aux marchés britanniques de collecte des déchets. Ils comparent les données de ces marchés avant et après 1988, année à partir de laquelle les collectivités locales ont été contraintes d'attribuer leurs marchés par appels d'offre. Bivand et Szymanski montrent que la part de franchises opérées par des régies est passée de 90 % à 70 %, et que le coût moyen a diminué de 20 %. Ils montrent également par un traitement économétrique approprié que les prix d'après 1988 dépendent significativement plus du taux de salaire moyen local. Inversement, l'autocorrélation spatiale diminue nettement après 1988. Ils interprètent ce phénomène par une réduction des effets de dépendance spatiale. Après 1988, les comparaisons menées par les régulateurs locaux avec les collectivités voisines portent davantage sur des prix exogènes, que sur des prix fixés par des objectifs politiques inobservables.

Les auteurs concluent que dans un tel cadre institutionnel, il est nécessaire que des règles communes encadrent les politiques particulières des régulateurs locaux, afin de garantir l'efficience de la concurrence par comparaison. Ainsi, l'intervention d'un « régulateur des régulateurs », ou la mise en place d'une agence commune, regroupant les régulateurs locaux, permet d'amoindrir le risque d'effets de dépendance spatiale.

Nous achevons ici l'analyse théorique de l'hétérogénéité que nous avons qualifiée d'*externe*. Nous rediscuterons, de façon approfondie, ces questions d'hétérogénéité

externe dans la dernière partie, lorsqu'il s'agira de mettre en œuvre le mécanisme comparatif que nous appliquerons aux chemins de fer. Nous passons à présent à l'analyse de l'hétérogénéité propre aux agents, c'est-à-dire celle d'un cadre de sélection adverse.

3.4 L'apport des comparaisons dans un cadre de sélection adverse

La réflexion concernant l'apport informationnel des comparaisons dans un cadre de sélection adverse a été développée progressivement dans plusieurs directions. On doit l'essentiel de ces résultats aux travaux d'Auriol et Laffont, initiés dès le début des années 1990.

Comme pour l'analyse en aléa moral, nous commençons cette section par la présentation du cadre de référence de sélection adverse, appliqué aux contrats individuels. Nous analysons ensuite les principaux modèles de la littérature : celui de Crémer et McLean, puis celui d'Auriol et Laffont.

3.4.1 Les modèles de sélection adverse

Afin d'introduire les principaux modèles comparatifs, nous rappelons succinctement [26] la problématique générale de sélection adverse, puis la référence des contrats individuels de la régulation classique. Enfin, nous présentons les principes généraux sur lesquels reposent les modèles comparatifs en sélection adverse.

La problématique de la sélection adverse

Le phénomène de sélection adverse se manifeste dès lors que le principal est confronté à une incertitude sur l'identité de l'agent. Appliqué à la régulation, le jeu qui lie le principal et l'agent peut être décrit de la façon suivante :

1. le régulateur propose à l'entreprise de produire une quantité de produit q en échange d'un transfert monétaire t et de la vente au prix unitaire p ;
2. le paiement du transfert t est affecté par le coût social des fonds publics, $\lambda > 0$, pour tenir compte de la désutilité des contribuables face à la taxation ;
3. l'entreprise est caractérisée par son niveau de productivité $\beta \in [\underline{\beta}; \overline{\beta}]$, que le régulateur n'observe pas [27] ; sa fonction de coût s'écrit : $C = \beta q + K$;
4. l'entreprise maximise son profit $\pi(q, \beta) = t + p(q)q - C(q, \beta)$;
5. le régulateur maximise un surplus collectif.

L'asymétrie d'information intervient ici au niveau de l'identité, la productivité β (*exogène*) de l'entreprise, que le régulateur ne peut observer sans coût [28]. Le régula-

[26] Pour plus de détail, voir l'analyse approfondie de Laffont & Martimort (2002) [74].

[27] Ce niveau de productivité peut être considéré comme étant le résultat d'un investissement antérieur. Nous envisagerons explicitement ce cas dans la section 4.3 du chapitre suivant.

[28] On suppose généralement qu'un audit plus ou moins coûteux permet au régulateur d'acquérir plus ou moins d'information sur β.

teur ne parvient donc pas à ajuster le niveau de son transfert au coût de l'entreprise. De ce fait, une entreprise productive ($\beta \gtrsim \underline{\beta}$) a tout intérêt à se faire passer pour peu efficiente ($\beta \lesssim \overline{\beta}$), afin de bénéficier d'un transfert élevé par rapport à ses coûts. Le problème consiste donc à déterminer les incitations à fournir, afin que l'entreprise productive révèle son efficience. Pour autant, le régulateur ne peut exclure l'éventualité que l'entreprise soit inefficiente ; il ne peut donc pas formuler un contrat unique destiné aux seules entreprises efficientes. La solution à ce dilemme consiste à proposer à l'entreprise, non pas un contrat, mais un menu de contrats auto-sélectif, c'est-à-dire tel que l'entreprise soit incitée à choisir le contrat correspondant à sa productivité. Nous rappelons dans les paragraphes qui suivent les différentes allocations théoriques, dans le cadre des contrats individuels entre un régulateur et plusieurs entreprises.

La référence des contrats individuels en sélection adverse

Lorsque l'information est complète, c'est-à-dire que le régulateur peut observer l'identité des entreprises, Baron & Myerson (1982) [61] rappellent que la solution dépend du coût social des fonds publics. Si $\lambda = 0$, il est possible d'atteindre l'optimum de premier rang, avec une tarification au coût marginal et un transfert remboursant les coûts fixes K. En revanche, pour $\lambda > 0$, on obtient une solution de second rang, avec une tarification de Ramsey-Boiteux.

Lorsque l'information est incomplète, le régulateur n'observe plus le niveau de productivité des entreprises. Il doit alors les inciter à révéler leur efficience avec un menu de contrats approprié. Baron et Myerson montrent qu'à l'optimum, l'offre de contrats varie entre les deux *extrema* que nous détaillons :

- Les entreprises efficientes ($\beta = \underline{\beta}$) dévient de la solution de premier rang. Ce résultat sous-optimal est dû aux incitations à leur verser pour qu'elles acceptent de révéler leur productivité. Le transfert qui leur est alloué comporte alors deux termes, l'un correspondant au remboursement des coûts non couverts par les prix, l'autre étant une *rente informationnelle* cédée à l'entreprise.
- Les entreprises inefficientes ($\beta = \overline{\beta}$) dévient, quant à elles, de la solution de premier rang à cause du niveau d'effort sous-optimal qu'il leur est demandé de fournir. En effet, afin de limiter le coût de la rente informationnelle précédente, le régulateur a intérêt à distordre la production en réduisant la quantité optimale de production des entreprises inefficientes. Ce faisant, le transfert à leur verser est plus faible qu'à l'optimum, ce qui incite moins les entreprises efficientes à se faire passer pour inefficientes. Cela permet donc de réduire la rente informationnelle à verser à ces dernières.

Ainsi, dans un contexte de sélection adverse, il n'est plus possible d'atteindre l'optimum de premier rang. Dans la suite de cette section, nous détaillons les principes par lesquels des mécanismes comparatifs permettent de réduire l'asymétrie informationnelle.

L'apport du contexte multi-agent

Dans un contexte multi-agent, un recours adéquat aux comparaisons permet de diminuer l'incertitude du régulateur quant à l'identité des entreprises. L'idée des modèles que nous présentons consiste à tirer parti des annonces que font les entreprises de leur caractéristique. Ces annonces dégagent des externalités informationnelles que le régulateur peut exploiter, afin d'inciter les entreprises à révéler leur identité.

L'introduction de mécanismes comparatifs dans le cadre de sélection adverse nécessite une hypothèse supplémentaire par rapport au contexte avec un seul agent : les caractéristiques des entreprises doivent être corrélées. Comme dans le cadre d'aléa moral, cette hypothèse est assez intuitive. La comparaison des annonces des entreprises ne peut être informative qu'à la condition que les paramètres qu'elles annoncent ne soient pas indépendants les uns des autres.

Deux principaux types de modèles ont analysé l'apport informationnel des comparaisons dans un cadre de sélection adverse multi-agent. D'une part, certains modèles inspirés de Crémer & McLean (1988) [66] montrent que le régulateur peut mettre en œuvre une solution de premier rang, moyennant une hypothèse forte de neutralité au risque. D'autre part, on trouve dans la littérature les modèles développés sur la base d'Auriol & Laffont (1992) [56], que nous détaillons dans la dernière partie de cette section.

3.4.2 Le modèle de Crémer & McLean (1988)

Crémer & McLean (1988) [66] ont montré que la corrélation des informations cachées des agents permet au principal d'extraire la totalité de leurs rentes. Ce modèle a été développé dans le cadre d'un mécanisme d'enchères privées, où le vendeur ignore la valorisation que les enchérisseurs ont de son bien. Contrairement aux modèles d'enchères classiques, Crémer et McLean montrent qu'un vendeur peut atteindre une solution de premier rang. Leur modèle ne concerne donc pas directement la régulation des entreprises. McAfee & Reny (1992) [79] ont repris ce modèle discret en l'étendant au cas continu. Ils font remarquer qu'il peut être transposé à une relation d'agence entre un régulateur et des entreprises. Ce principe, communément admis, à été formalisé par Pouyet (2002) [85] qui analyse le rôle de la corrélation dans les incitations à la collusion[29]. Nous présentons donc ici le cadre général du modèle bi-agent de Pouyet, inspiré par les résultats de Crémer et McLean.

Pouyet considère deux entreprises neutres au risque, dont la caractéristique de productivité peut prendre deux valeurs discrètes : $\beta \in \{\underline{\beta}; \overline{\beta}\}$. Le régulateur confie la production à l'entreprise la moins-disante, ou, en cas d'annonces identiques, partage la production entre les deux entreprises. Pouyet applique le processus de Crémer et McLean, qui suppose que les entreprises soient corrélées. Le régulateur parvient alors à contraindre les entreprises à révéler leur caractéristique, sans leur abandonner une

[29] Nous présentons les résultats de Pouyet concernant la collusion dans le chapitre suivant.

rente informationnelle. Cela lui permet de mettre en œuvre la solution de premier rang.

Mais le mécanisme révélateur qui permet d'atteindre ce résultat doit être discuté. En effet, à l'issue du jeu, le profit $\pi_{\beta_i;\beta_j}$ de l'entreprise i, de type β_i, face à l'entreprise j de type β_j, vérifie selon les cas :

$$\begin{aligned}
\pi_{\underline{\beta};\underline{\beta}} &\leqslant 0 \\
\pi_{\underline{\beta};\overline{\beta}} &> 0 \\
\pi_{\overline{\beta};\underline{\beta}} &< 0 \\
\pi_{\overline{\beta};\overline{\beta}} &\geqslant 0
\end{aligned}$$

En espérance, quelle que soit la valeur de β_i, l'entreprise i escompte un profit nul. Mais selon la valeur de β_j, ce profit sera finalement négatif (si les deux entreprises sont efficientes, ou si seule j l'est) ou positif. Le mécanisme repose donc largement sur l'hypothèse de neutralité au risque des entreprises. Celles-ci ne doivent pas craindre de participer à la *loterie* que le principal leur propose, bien qu'elles n'espèrent rien en tirer.

Précisons qu'à l'*extremum* où la correlation des entreprises est quasi-nulle, les profits tendent vers l'infini (en valeur absolue). En revanche, en cas de corrélation parfaite, les profits $\pi_{\underline{\beta};\underline{\beta}}$ et $\pi_{\overline{\beta};\overline{\beta}}$ (les deux seules issues possibles du jeu) sont nuls.

Ainsi, ce mécanisme, *a priori* séduisant, doit beaucoup aux fortes hypothèses sur lesquelles il repose. La connaissance commune de la distribution de probabilité des caractéristiques est sans doute une hypothèse excessive ; mais elle paraît être un mal nécessaire de nos modèles. Le recours à l'hypothèse de neutralité au risque est, en revanche, plus contestable. Crémer et McLean remarquent d'ailleurs en conclusion, le décalage entre les résultats théoriques puissants qu'ils établissent, et l'absence de mise en œuvre pratique de leur mécanisme. Ils attribuent justement cette observation à leurs hypothèses, sans doute trop fortes.

Enfin, ce mécanisme se rapproche davantage d'un modèle de concurrence *pour le marché* que de concurrence *par comparaison*. Et ce, bien qu'il mobilise explicitement la corrélation des entreprises. Ceci est assez logique dans la mesure où le modèle de régulation de Pouyet dérive de celui d'enchères de Crémer et McLean. Mais n'est-ce pas le propre de la concurrence pour le marché, que de comparer les annonces d'entreprises, *a priori* corrélées, puisque susceptibles de produire sur le même marché ? Aussi nous nous attardons davantage dans la suite de cette section sur le modèle d'Auriol et Laffont qui traite plus spécifiquement de concurrence par comparaison.

3.4.3 Le modèle d'Auriol & Laffont (1992)

Nous détaillons à présent le modèle d'Auriol & Laffont (1992) [56], à travers le principe de modélisation retenu, les résultats obtenus et l'extension proposée par Auriol

(1993) [57]. Ce modèle a été développé afin d'endogénéiser la structure monopolistique ou duopolistique d'un marché. Auriol et Laffont ont élaboré leur modèle en s'inspirant des pratiques de *dual sourcing*, déjà mentionnées par Shleifer. Leur réflexion est également motivée par le développement de marchés duopolistiques, là où d'importants coûts fixes laissaient présager une structure monopolistique. À partir d'un modèle général, les auteurs évaluent la pertinence de marchés monopolisitiques ou duopolistiques, selon que le régulateur décide de cette structure avant ou après que les entreprises n'annoncent leur caractéristique. Nous détaillons ici la version *ex ante* de leur modèle, lorsque le régulateur opte pour un marché duopolistique. C'est dans cette configuration que la concurrence par comparaison intervient.

Auriol et Laffont considèrent donc le cas d'un régulateur face à deux entreprises, supposées neutres au risque [30]. Le déroulement du jeu proposé est le suivant :

1. le régulateur choisit une structure duopolistique de marché ;
2. il investit doublement les coûts fixes (auprès de chaque entreprise) ;
3. les entreprises prennent connaissance du marché et de leur productivité β^i ;
4. le régulateur propose un menu de contrats $(t; q)$ qui conduit les entreprises à révéler leur identité ;
5. le contrat est exécuté : la production est réalisée par l'entreprise retenue, et les transferts sont versés.

Principe du modèle comparatif

Auriol et Laffont modélisent le paramètre source d'asymétrie d'information, de la même manière que les modèles en aléa moral. Ils considèrent en effet que l'identité de l'agent i, β^i, peut être décomposée en une composante commune et une autre, individuelle :

$$\beta^i = \alpha b + (1 - \alpha)\epsilon^i$$

où $b \in \{\underline{b}; \overline{b}\}$ avec ν la probabilité que b soit égal à $\underline{b}$; le terme αb s'interprète donc comme le coût marginal moyen du secteur,
les ϵ^i sont des variables aléatoires indépendantes et identiquement distribuées dans $[\underline{\epsilon}; \overline{\epsilon}]$,
$\alpha \in [0; 1]$ mesure la corrélation entre les β^i : pour $\alpha = 1$, $\forall i$, $\beta^i = b$ (les entreprises sont identiques), tandis que pour $\alpha = 0$, $\beta^i = \epsilon^i$ (les productivités sont indépendantes).

On a donc $\beta^i \in [\alpha\underline{b} + (1 - \alpha)\underline{\epsilon}; \alpha\overline{b} + (1 - \alpha)\overline{\epsilon}] = [\underline{\beta}; \overline{\beta}]$, intervalle que les auteurs divisent en deux intervalles disjoints correspondant chacun à une valeur de b. En posant $0 \neq \alpha = (\overline{\epsilon} - \underline{\epsilon})/(\overline{b} - \underline{b} + \overline{\epsilon} - \underline{\epsilon})$, on obtient la structure de caractéristiques décrite par la figure 3.2 :

[30] Ici, l'hypothèse de neutralité au risque a une portée bien moindre que chez Crémer et McLean, ce qui la rend acceptable.

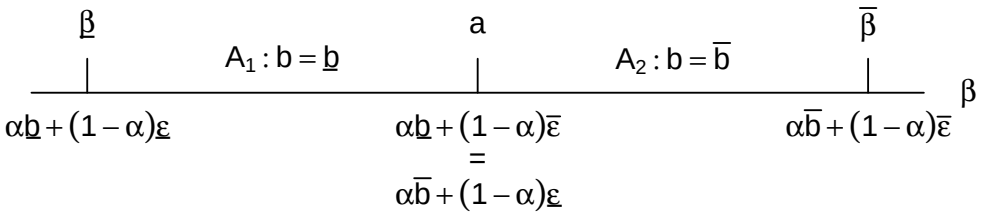

FIG. 3.2 – Support de distribution des caractéristiques des entreprises.
Source : Auriol & Laffont (1992) [56].

Cette disjonction des intervalles permet à l'entreprise qui découvre sa caractéristique, de déduire avec certitude la valeur de b et l'espérance de la productivité de l'autre agent. Cette structure des caractéristiques des entreprises est telle que les résultats de Crémer & Mc Lean (1988) [66] ne s'appliquent pas. Intuitivement, on pressent que les résultats de ce modèle devraient ressembler à ceux d'Holmström : l'impact de la composante commune du paramètre d'asymétrie d'information (b ou η) doit pouvoir être annulé, tandis que la composante particulière (ϵ), résiduelle, empêche d'atteindre une solution de premier rang. On retrouve d'ailleurs le fait qu'en l'absence de corrélation ($\alpha = 0$), la concurrence par comparaison n'est d'aucune utilité.

Résultats

Chaque entreprise choisit un contrat $(t; q)$, afin de maximiser son utilité espérée, conditionnellement à ses anticipations de la caractéristique de l'autre entreprise. L'équilibre est donc de type Nash-Bayesien, chaque entreprise ne connaissant que la composante commune de la caractéristique de l'autre.

Dans le jeu proposé par Auriol et Laffont, l'usage des comparaisons permet de réduire les incitations qu'ont les entreprises à se faire passer pour moins efficientes qu'elles ne sont. Notamment, les entreprises ne peuvent plus cacher la composante commune de leur caractéristique :

- Si $\beta^i \in A_2 =]a; \overline{\beta}]$, les deux entreprises sont de type $\overline{b}$; l'entreprise i n'a jamais intérêt à annoncer une caractéristique dans A_1 (cela reviendrait à annoncer un coût inférieur au sien).
- Si $\beta^i \in A_1 = [\underline{\beta}; a]$, les deux entreprises sont de type $\underline{b}$. En fixant des pénalités infinies lorsque les annonces de i et j diffèrent, le régulateur contraint les entreprises à révéler qu'elles sont efficientes, sans que cela lui coûte [31].

Cet effet de concurrence par comparaison, qui permet de tronquer l'espace des annonces possibles, présente un double intérêt :

1. Une réduction des rentes informationnelles des entreprises de type $\underline{b}$. En effet, celles-ci ne peuvent plus annoncer une caractéristique $\beta^i \in A_2$; cela réduit l'espace des rentes informationelles de $[\beta^i; \overline{\beta}]$ à $[\beta^i; a]$. L'asymétrie d'information,

[31] Cette hypothèse de pénalités infinies en cas d'annonces incompatibles est assez forte. Auriol (1993) [57] reconnaît le caractère extrême de cette hypothèse de punition non bornée.

et donc la rente informationnelle, ne portent alors plus que sur la composante particulière de la caractéristique, comme pour les entreprises de type $\overline{b}$ (pour lesquelles les rentes informationnelles ne sont pas modifiées).

2. Un relâchement des distorsions de production des entreprises de type $\overline{b}$ (nous reproduisons à la figure 3.3 l'illustration proposée par Auriol et Laffont). L'effort sous-optimal qui était demandé aux entreprises de type $\overline{b}$, afin de réduire les incitations des autres à dévier de leur contrat, n'est plus justifié. Les entreprises de type $\underline{b}$ ne pouvant plus se faire passer pour ce qu'elles ne sont pas, ces distorsions peuvent être relâchées. Ainsi, pour $\beta = a^+$, on retrouve la production de premier rang. Sur la figure 3.3, cela correspond au passage des pointillés aux tirets, sur l'intevalle $[a; \overline{\beta}]$.

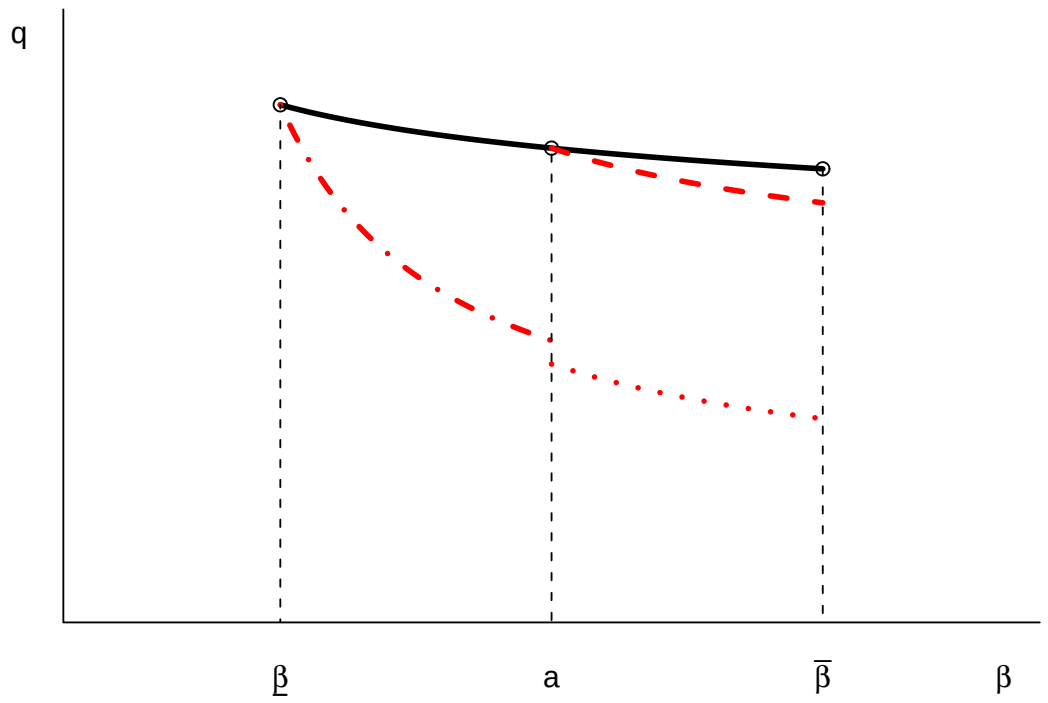

FIG. 3.3 – Relâchement des distorsions de productions grâce aux comparaisons ; —— : optimum de premier rang ; - · - · - : duopole et monopole ; - - - : duopole seul ; · · · · · · : monopole seul.
Source : Auriol & Laffont (1992) [56].

Dans le cas extrême où $\alpha = 1$, on retrouve le modèle de Shleifer : le régulateur parvient à capter tout le surplus, comme dans le modèle de Crémer et McLean dont les hypothèses sont alors vérifiées.

Auriol et Laffont montrent ainsi qu'en dépit de la duplication des coûts fixes, un marché duopolistique peut être préféré à une structure monopolistique. Outre le bénéfice qu'il retire de la comparaison des annonces, le régulateur profite également

d'un effet *d'échantillonnage* ; c'est-à-dire qu'il peut confier la production à l'entreprise dont les coûts sont les plus faibles (opportunité inexistante sur un marché monopolistique).

Ce modèle qui introduit de la concurrence par comparaison dans un cadre de sélection adverse a été progressivement étendu dans plusieurs directions. Dans les paragraphes qui suivent, nous présentons l'extension d'Auriol (1993).

La généralisation d'Auriol (1993)

Le modèle d'Auriol (1993) [57] relâche l'hypothèse de disjonction des intervalles A_1 et A_2, supports des distributions des caractéristiques des entreprises. En posant $\alpha \leqslant (\overline{\epsilon} - \underline{\epsilon})/(\overline{b} - \underline{b} + \overline{\epsilon} - \underline{\epsilon})$, Auriol obtient la structure de caractéristiques décrite par la figure 3.4 :

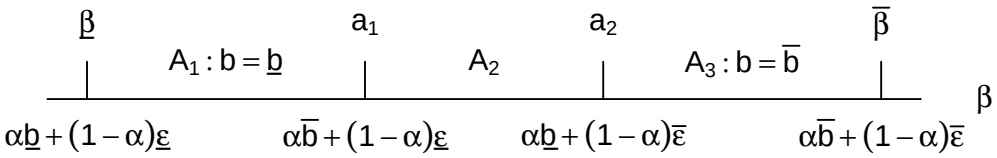

FIG. 3.4 – Support de distribution des caractéristiques des entreprises.
Source : Auriol (1993) [57].

Cette structure du modèle ne permet toujours pas d'appliquer les résultats de Crémer et Mc Lean. On retrouve, en revanche, une forte similitude avec ceux du modèle d'origine, à ceci près que l'observation de β^i ne permet plus de déduire systématiquement la valeur de b. Notamment, les entreprises dont la caractéristique est dans A_2 ignorent si la composante commune est $\overline{b}$ ou $\underline{b}$.

Cette indétermination partielle modifie légèrement les résultats du modèle d'origine, bien que les effets principaux soient préservés. Pour les mêmes raisons que précédemment, les entreprises sont contraintes aux annonces compatibles suivantes :

- Si $\beta^i \in A_3$, les deux entreprises sont de type $\overline{b}$; l'entreprise i n'a jamais intérêt à annoncer une caractéristique dans $A_1 \cup A_2$.
- Si $\beta^i \in A_2$, il y a indétermination, les deux entreprises peuvent aussi bien être de type $\underline{b}$ ou $\overline{b}$. Sous la menace de pénalités infinies en cas d'annonces incompatibles, l'entreprise i n'a pas intérêt à annoncer $\beta^i \in A_3$.
- Si $\beta^i \in A_1$, les deux entreprises sont de type $\underline{b}$. En annonçant $\beta^i \in A_3$, l'entreprise i est pénalisée par l'incompatibilité de son annonce avec celle de l'entreprise j. Elle est donc contrainte de restreindre ses annonces à $A_1 \cup A_2$.

De même que dans le modèle initial, il résulte de cette troncature de l'espace des annonces possibles :

1. Une réduction des rentes informationnelles. L'espace des rentes informationelles des entreprises qui observent une caractéristique $\beta^i \in A_1 \cup A_2$ est tronqué de $[\beta^i; \overline{\beta}]$ à $[\beta^i; a_2]$.

2. Un relâchement des distorsions de production des entreprises dont la caractéristique $\beta^i \in A_3$.

Ces résultats prolongent bien ceux du modèle d'Auriol et Laffont. Une différence intervient cependant au niveau de l'entreprise finalement chargée de la production. Auriol montre en effet que celle-ci n'est pas nécessairement la plus efficiente en terme de productivité, mais celle dont les coûts totaux (qui incluent les coûts informationnels) sont les plus faibles. Ainsi, Auriol montre que les principaux résultats du modèle d'origine sont maintenus et que l'effet de comparaison réduit toujours l'asymétrie d'information lorsque les caractéristiques des entreprises sont corrélées.

Cette première extension du modèle d'Auriol et Laffont illustre la souplesse de celui-ci, qui sera repris dans d'autres contextes. Nous verrons dans le chapitre suivant que des auteurs ont repris ce modèle afin d'étudier l'effet de la capture du régulateur ou encore les incitations à l'investissement.

Nous analysons à présent, dans la dernière section de ce chapitre, la concurrence par comparaison dans un cadre de sélection adverse doublé d'aléa moral. Nous verrons que le modèle de base d'Auriol et Laffont a été repris avec introduction d'une variable d'effort inobservable par le régulateur.

3.5 Concurrence par comparaison en aléa moral et sélection adverse

Dans cette section, nous analysons l'efficience des mécanismes comparatifs, en présence simultanée d'aléa moral (asymétrie d'information portant sur la variable endogène d'effort) et de sélection adverse (asymétrie d'information portant sur la variable exogène d'identité). Cette section est organisée comme les précédentes : nous rappelons d'abord les modèles classiques dans le cadre général, avant d'étudier dans le détail les pricipaux modèles de concurrence par comparaison.

3.5.1 Les modèles en aléa moral et sélection adverse

La problématique de l'aléa moral doublé de sélection adverse

Lorsque les deux sources d'asymétrie d'information sont couplées, les jeux que nous avons détaillés dans les sections précédentes interagissent. On considère généralement un jeu de la forme suivante :

1. les agents sont supposés neutres au risque[32] ;
2. le régulateur propose à l'entreprise une production fixe (normalisée à 1), en échange d'un transfert monétaire t ;
3. l'entreprise est caractérisée par sa productivité β que le régulateur n'observe pas ; elle peut fournir un effort e, également inobservable, moyennant une désu-

[32] Sauf (rare) mention contraire, les entreprises ne présentent pas d'aversion au risque.

tilité strictement convexe $\psi(e)$ [33] ; sa fonction de coût (observable *ex post*) s'écrit : $C = \beta - e$;

4. l'entreprise maximise son utilité $t - C(\beta, e) - \psi(e)$.

Du fait de la double asymétrie d'information, le régulateur ne peut ajuster le niveau du transfert au coût et à la désutilité à l'effort de l'entreprise. Nous avons rappelé qu'en situation d'aléa moral, un transfert à prix fixe incite les entreprises à l'effort. Mais en présence de sélection adverse, le régulateur ignore β et ne peut donc pas définir le montant efficace de ce transfert ; les entreprises jouissent alors de rentes informationnelles. Inversement, en situation de sélection adverse pure, un transfert basé sur les coûts de l'entreprise l'incite à révéler sa productivité. Mais en présence d'aléa moral, ce contrat ne fournit à l'entreprise aucune incitation à l'effort. Le régulateur est alors confronté à un nouveau dilemme entre fournir des incitations à la révélation d'information, et fournir des incitations à l'effort ; les unes étant donc incompatibles avec les autres. La solution consiste alors à proposer à l'entreprise un menu de contrats auto-sélectif que nous caractérisons dans le paragraphe suivant.

La référence des contrats individuels

Dans le contexte informationnel que nous considérons, et avec des hypothèses plus générales que celles que nous venons d'exposer, Laffont & Tirole (1986) [75] proposent un mécanisme incitatif d'allocation optimale. Ils montrent que le régulateur devrait offrir à l'entreprise un menu de contrats linéaires, fonctions du paramètre de productivité. Ces contrats comprennent une part de remboursement des coûts, et une part fixe. Ils varient entre les deux extrema que nous décrivons :

- Les entreprises efficientes ($\beta = \underline{\beta}$) sont incitées à choisir un pur contrat à prix fixe. Ce contrat, extrêmement incitatif, abandonne à l'entreprise une rente informationnelle importante et lui impose l'effort optimal.
- Les entreprises inefficientes ($\beta = \overline{\beta}$) sont incitées à choisir un contrat de type *cost plus*. Ce contrat, peu incitatif, ne laisse aucune rente informationnelle et n'impose à l'entreprise qu'un faible niveau d'effort.

Lorsque l'entreprise est faiblement averse au risque, Laffont et Tirole montrent que la part de remboursement des coûts augmente avec l'aversion au risque. Cela revient à offrir une part d'assurance à l'entreprise.

Dans ce contexte de double asymétrie informationnelle, il n'est, bien sûr, plus possible d'atteindre l'équilibre de premier rang de la solution en information parfaite et complète. Ceci résulte, d'une part, des rentes informationnelles cédées aux entreprises efficientes, et d'autre part, des distorsions imposées à l'effort des entreprises moins efficientes. Nous présentons brièvement dans les paragraphes suivants comment la présence de plusieurs agents peut améliorer cette solution mono-agent.

[33] Cette désutilité à l'effort correspond à l'investissement $R(c)$ du modèle de Shleifer.

L'apport du contexte multi-agent

Nous avons montré dans les sections précédentes comment la concurrence par comparaison pouvait améliorer les solutions au problème de la régulation, en aléa moral puis en sélection adverse. Dans le contexte que nous étudions à présent, on peut s'attendre à ce que la concurrence par comparaison améliore la solution des contrats individuels, en réduisant l'asymétrie d'information globale. En particulier, le jeu étant proche de celui en sélection adverse, il est raisonnable d'espérer que la concurrence par comparaison permette de réduire les rentes informationnelles et les distorsions à l'effort.

On doit les premiers résultats obtenus dans cette configuration à Demski & Sappington (1984) [68]. Leur modèle bi-agent propose d'atteindre une solution Pareto-optimale par un traitement asymétrique des agents qui sont corrélés et potentiellement averses au risque. Le jeu qu'ils établissent conduit un premier agent à révéler son identité (stratégie dominante), puis le second à révéler la sienne en meilleure réponse (équilibre de Nash). Laffont & Tirole (1993) [76] proposent aussi un modèle simple de concurrence par comparaison dans un tel contexte d'asymétrie d'information, avec deux agents neutres au risque. Ils montrent qu'un contrat qui offre $t^i = \psi(e^*) - (C^i - C^j)$ est une solution de premier rang. Le prix fixe incite l'entreprise à fournir l'effort optimal, tandis que la comparaison des coûts permet d'annihiler les effets de la composante commune d'identité. Enfin, mentionnons également la réflexion développée par Choné & Lesur (2001) [64] dans un tel contexte, selon que le régulateur observe, ou non, l'impact de l'hétérogénéité environnementale sur les coûts.

Dans les paragraphes qui suivent, nous détaillons le modèle d'Auriol (2000) [59] qui compare l'allocation obtenue en concurrence par comparaison, avec celle des contrats individuels dans un cadre multi-agent. Enfin, nous achevons cette section avec l'examen du modèle de Boyer & Laffont (2003) [63] qui compare les niveaux d'effort obtenus en concurrence par comparaison suivant la qualité de l'information perçue par le régulateur.

3.5.2 Le modèle d'Auriol (2000)

Dans un contexte de double asymétrie d'information, Auriol (2000) [59] reprend le modèle qu'elle avait établi avec Laffont en 1992 pour l'adapter au cadre de Laffont et Tirole (1993) [76] avec plusieurs agents [34]. Les hypothèses et notations que nous avons rappelées dans cette section et la précédente sont reprises. Auriol définit notamment toujours les caractéristiques de productivité par $\beta^i = \alpha b + (1-\alpha)\epsilon^i$, réparties dans deux intervalles disjoints. Elle précise également que la fonction de désutilité à l'effort, strictement convexe, vérifie $\psi'''(e) \geqslant 0$ [35]. Enfin, à défaut d'introduire le coût social

[34] Précisons qu'il n'est plus question dans ce nouveau modèle d'endogénéiser la structure de marché.

[35] Cette hypothèse relativement courante revient à considérer que la convexité de la désutilité à l'effort ne décroît pas, ce qui est tout à fait acceptable en soi. En outre, cette hypothèse favorise

des fonds publics, λ, elle considère que le régulateur maximise le surplus des seuls consommateurs.

Auriol établit les résultats suivants :

- Lorsque la corrélation des entreprises est parfaite ($\alpha = 1$), on retrouve la configuration et la solution du modèle de Shleifer. Il est alors possible d'atteindre l'optimum de premier rang dans ce cas très particulier.
- Pour $\alpha = 0$, les entreprises n'ont rien en commun ; les comparaisons n'ont alors pas d'utilité. Dans ce cas, l'optimum est défini par les contrats individuels de Laffont & Tirole (1986) [75].
- Lorsque la corrélation est imparfaite, on retrouve les résultats généraux obtenus dans Auriol & Laffont (1992) [56], c'est-à-dire que « tout se passe comme si αb [la composante commune de la caractéristique] était observable ». Moyennant l'hypothèse de détection-punition des annonces incompatibles (voir page 101), les rentes informationnelles sont tronquées et les distorsions sur l'effort des entreprises $b = \overline{b}$ atténuées.

Mais la cohérence du modèle d'Auriol naît surtout de ce qu'elle propose un mécanisme comparatif, doté des propriétés suivantes :

1. il est optimal ;
2. il prend la forme classique d'un menu de contrats linéaires ;
3. il atteint à chacune de ces extrémités selon que $\alpha = 1$ ou 0 la solution de Shleifer ou celle de Laffont et Tirole, respectivement.

Le menu de contrats linéaires offre des transferts de la forme :

$$t^i(C^i, C^{-i}) = a(\epsilon^i) + \left[b(\epsilon^i)C^i + (1 - b(\epsilon^i))\frac{\sum_{j \neq i} C^j}{N-1} \right]$$

où $a(\epsilon^i)$ est une prime fixe qui décroît avec ϵ,
et où $b(\epsilon^i) = 1 - \psi'(e(\epsilon^i))$ croît de 0 à 1 quand ϵ varie de $\underline{\epsilon}$ à $\overline{\epsilon}$.
Ainsi, le remboursement du coût s'effectue à travers une moyenne pondérée du coût final observé de l'entreprise et de celui de son entreprise jumelle (son étalon, défini par Shleifer). La pondération en faveur de l'un ou l'autre dépend de la caractéristique particulière de l'entreprise :

- **Les entreprises relativement efficientes ($\epsilon^i \longrightarrow \underline{\epsilon}$)** choisissent un contrat quasi-indépendant de leur coût ($b(\epsilon^i) \to 0$). Elles bénéficient d'une prime fixe élevée en échange de l'effort optimal qu'elles doivent fournir.
- **Les entreprises moins efficientes ($\epsilon^i \longrightarrow \overline{\epsilon}$)** choisissent un contrat de type *cost plus* ($b(\epsilon^i) \to 1$). L'effort qu'elles supportent est plus faible, et donc leur prime fixe de compensation aussi.

Enfin, une particularité de ce modèle doit être soulignée : il est robuste à la collusion, ce que nous discuterons davantage dans le chapitre suivant. Dans la continuité

la résolution des calculs en assurant la concavité de la fonction objectif du principal. Cela vient du fait qu'elle impose la convexité de la rente informationnelle de l'entreprise efficiente, par rapport au niveau d'effort fixé par le mécanisme incitatif : $\Phi(e) = \psi(e) - \psi(e - \Delta\beta)$.

des modèles en sélection adverse, Auriol (2000) [59] montre ainsi que la concurrence par comparaison améliore la solution des contrats individuels, lorsque l'effort est inobservable. Nous étudions enfin le récent modèle de Boyer & Laffont (2003) [63] qui présente des résultats intéressants dans un cadre différent.

3.5.3 Le modèle de Boyer & Laffont (2003)

Boyer & Laffont (2003) [63] étudient dans quelle mesure la concurrence par comparaison réduit les distorsions d'effort imposées aux entreprises moins efficientes [36]. Ce modèle présente l'intérêt majeur de se démarquer des modèles classiques de concurrence par comparaison afin de mieux correspondre à l'usage qu'il est fait en pratique des mécanismes comparatifs. Ainsi, les auteurs considèrent que l'apport informationnel des comparaisons repose sur l'observation éventuelle par le régulateur d'un signal corrélé avec le niveau d'efficience de l'entreprise. Cette information *ex ante* lui permet d'établir un contrat adapté, en particulier concernant les distorsions sur le niveau d'effort. En revanche, ce signal, non vérifiable, ne peut être « contractualisé » [37].

Le modèle de Boyer et Laffont repose sur les hypothèses classiques que nous avons déjà explicitées, à celles-ci près :

- ils ne considèrent que deux types d'entreprises : $\beta \in \{\underline{\beta}; \overline{\beta}\}$, avec ν la probabilité (connue du régulateur) que $\beta = \underline{\beta}$;
- la fonction de désutilité à l'effort, strictement convexe, vérifie $\psi'''(e) \geqslant 0$.

Lorsque les entreprises efficientes sont nettement majoritaires, les auteurs montrent qu'il vaut mieux renoncer au menu de contrats, au profit d'un unique contrat, établi pour des entreprises efficientes. En effet, compte tenu du faible nombre d'entreprises inefficientes, il devient trop coûteux de maintenir un menu avec un contrat $\underline{\beta}$ avec rentes informationnelles, et un autre $\overline{\beta}$ avec distorsions à l'effort. Il existe donc une valeur ν^* de la probabilité ν au delà de laquelle il devient optimal de renoncer à ce que les entreprises inefficientes produisent. Cela permet au régulateur de supprimer les rentes informationnelles des autres.

En pratiquant la concurrence par comparaison, le régulateur perçoit un ensemble de signaux (σ_i) qui lui permettent de réviser ses croyances relatives à la caractéristique de l'entreprise. Chaque signal σ_i le conduit à évaluer la probabilité conditionnelle que l'entreprise soit efficiente : $\hat{\nu}_i = \mathrm{P}(\underline{\beta}|\sigma_i)$. Selon l'information portée par le signal, on distingue :

- **Un signal favorable lorsque $\hat{\nu}_i > \nu_i$:** cela indique que l'entreprise serait plutôt de type efficient $(\underline{\beta})$. Dans ce cas, le régulateur doit renforcer davantage les distorsions à l'effort. Cela lui permet de réduire la rente informationnelle à céder à l'entreprise si elle est effectivement de type $\underline{\beta}$.
- **Un signal défavorable lorsque $\hat{\nu}_i < \nu_i$:** cela indique que l'entreprise serait plutôt de type inefficient $(\overline{\beta})$. Le régulateur peut alors relâcher en partie les

[36] Nous ne rapportons ici qu'une partie du travail présenté dans leur article. La concurrence par comparaison n'est qu'un moyen parmi les autres qu'ils analysent d'améliorer les incitations à l'effort.

[37] Cette modélisation lève donc les réserves que l'on peut avoir quant à la menace de détection-punition des annonces incompatibles, initialement proposée par Auriol & Laffont (1992) [56].

distorsions à l'effort, quitte à augmenter la rente informationnelle qu'il est moins probable qu'il ait à verser.

Les résultats théoriques du modèle s'intercalent entre deux *extrema* que nous caractérisons :

- Lorsque tous les signaux vérifient $\hat{\nu}_i < \nu^*$, l'information qu'ils véhiculent est relativement incertaine. Les incitations à l'effort sont, en espérance, inférieures à celles des contrats individuels [38].
- Lorsque tous les signaux favorables vérifient $\hat{\nu}_i > \nu^*$, ils sont très informatifs. Le régulateur peut se fier au signal qu'il observe, sans grand risque d'erreur. Dans ce cas, il n'a plus intérêt à proposer un menu de contrats, mais plutôt le contrat adapté au type pressenti de l'entreprise. Cela lui permet de ne pas céder de rente informationnelle, si le signal qu'il observe lui indique que l'entreprise est vraisemblablement de type $\underline{\beta}$. Inversement, s'il observe avec une quasi-certitude que $\beta = \overline{\beta}$, il proposera un contrat adapté, sans distorsion d'effort.

Entre ces deux régimes extrêmes, il est difficile d'établir l'impact de l'apport informationnel sur la contrainte d'incitation.

Les auteurs poursuivent ensuite avec l'examen de quatre cas particuliers auxquels nous conseillons vivement au lecteur de se reporter. Enfin, Boyer et Laffont montrent que les résultats peu encourageants obtenus lorsque les signaux ne sont pas suffisamment informatifs peuvent être améliorés, en considérant un effet d'équilibre général [39].

Conclusion : un apport informationnel salvateur

Ce chapitre nous a conduit à analyser plusieurs modèles théoriques de concurrence par comparaison, dans des configurations informationnelles élémentaires. Ce travail nous a permis de percevoir par quels moyens l'usage des comparaisons peut améliorer la régulation des entreprises. Il est notamment apparu que la concurrence par comparaison améliore la solution des contrats individuels en présence d'asymétrie d'information. Qu'il s'agisse d'aléa moral ou de sélection adverse, les mécanismes comparatifs améliorent l'information du régulateur. En effet, que l'asymétrie informationnelle porte sur une variable exogène ou endogène, les comparaisons permettent d'en distinguer la composante commune de la composante individuelle. Le régulateur peut alors proposer une meilleure assurance et des contrats plus incitatifs aux entreprises. Le résultat essentiel consiste à réduire les rentes informationnelles et les distorsions qui les accompagnent. La concurrence par comparaison semble donc être un outil efficace pour améliorer la régulation des entreprises.

[38] Ce résultat est obtenu moyennant une hypothèse classique, qui peut être alternativement une désutilité à l'effort quadratique, ou bien un coût social des fonds publics faible. Dans chaque cas, la concavité de la relation entre l'effort distordu $\overline{e}_i$ et les $\hat{\nu}_i$ évalués est assurée, et conduit à ce résultat en espérance.

[39] Considérant le coût social des fonds publics comme endogène, ils montrent que la concurrence par comparaison, en réduisant les rentes informationnelles, atténue la contrainte budgétaire. Cela permet parfois d'accroître les incitations à l'effort, au delà de la solution des contrats individuels.

Il est toutefois encore trop tôt pour faire confiance aux mécanismes de concurrence par comparaison. En effet, nous n'avons examiné que des modèles de dimension réduite, par rapport à la complexité de la relation d'agence régulateur-entreprises. D'une part, il s'agit d'une dimension temporelle réduite puisque nous n'avons analysé que des modèles statiques. Or, la régulation des industries de réseau fait apparaître une répétition des interactions entre le régulateur et les entreprises. Il convient donc de tenir compte des effets dynamiques susceptibles de se manifester dans un tel contexte. D'autre part, les modèles étudiés laissent entrevoir un certain nombre de limites (au premier rang desquelles, la menace de collusion), sans les traiter. Il apparaît donc bien nécessaire de consacrer un nouveau chapitre à la théorie microéconomique. Il s'agit, cette fois, d'étudier l'impact de phénomènes classiques (collusion des entreprises, capture, engagement du régulateur, *etc.*), susceptibles de réduire le bénéfice des comparaisons.

Cette analyse unidimensionnelle de la concurrence par comparaison selon différentes contraintes informationnelles justifie le recours à cette forme de régulation afin de combler les lacunes informationnelles du régulateur. Toutefois, il serait bien trop réducteur de projeter la relation régulateur-entreprises sur cette seule dimension. Aussi est-il trop tôt pour conclure que la concurrence par comparaison peut opportunément être mise en œuvre afin de réguler des industries de réseau. Nous nous attachons donc, dans le chapitre suivant, à examiner les résultats des mécanismes comparatifs, en considérant d'autres dimensions de la régulation.

Chapitre 4

Analyses bidimensionnelles

Nous analysons dans ce chapitre l'apport des mécanismes comparatifs dans d'autres dimensions de la régulation que celle des contraintes informationnelles. Il s'agit pourtant toujours de projeter l'introduction de ces mécanismes sur cette dimension informationnelle de la régulation. Mais nous considérons, à chaque fois, une dimension supplémentaire de la régulation.

Il s'agit notamment d'envisager des comportements d'entente explicite entre les agents économiques. Nous discutons notamment de la collusion explicite entre les entreprises, que nous avons déjà abordée au chapitre précédent. Nous nous intéressons également aux conséquences que pourrait avoir la capture du régulateur par les entreprises.

Nous considérons ensuite la dimension temporelle de la régulation, à travers l'analyse de modèles dynamiques, et non plus statiques. Dans ce contexte, une autre forme de collusion, tacite, cette fois, est susceptible de perturber les résultats des configurations élémentaires. Mais le régulateur, pas plus que les entreprises, n'est incité à respecter les règles du jeu statique, lorsque celui-ci est répété plusieurs fois. Nous en étudions les conséquences à travers l'effet de cliquet.

Enfin, nous prenons en compte les nécessaires investissements, afin d'évaluer dans quelle mesure la concurrence par comparaison incite les entreprises à investir. Nous distinguons alors les investissements productifs du bénéfice desquels l'entreprise peut parfois être dépossédée. Nous achevons ce chapitre par l'étude de la compatibilité entre un mécanisme comparatif et les incitations à investir dans la qualité du service offert.

4.1 Les comportements d'entente explicite

Sur les marchés concurrentiels, l'optimum parétien est naturellement atteint à l'équilibre qui s'impose entre une offre et une demande atomisées. Toutefois, lorsque la condition d'atomicité des offreurs n'est plus valide, ces derniers sont tentés de s'entendre au sein d'un cartel, afin de se partager une rente monopolistique. Une telle entente menace le bien-être des consommateurs. Aussi, des autorités, tels le Conseil

de la concurrence en France ou la Direction Générale de la concurrence de l'Union europénne, ont été instituées, afin de garantir la concurrence sur les marchés.

Les industries de réseau qui font intervenir un petit nombre d'offreurs sur des marchés régulés n'échappent pas à ces comportements d'entente. Dans la mesure où ces activités sont régulées spécifiquement, la formation de cartels peut être plus délicate. De ce fait, les entreprises sont tentées de s'assurer la bienveillance du régulateur à leur égard, en le capturant. Ce sont ces deux types d'ententes, collusion et capture, que nous étudions dans cette section, lorsque la concurrence par comparaison est mise en œuvre.

4.1.1 La menace de collusion

La collusion est le phénomène d'entente entre agents économiques, désireux de se soustraire, grâce à leur coopération, au régime concurrentiel qui leur est imposé. Ce comportement d'entente, qui nuit gravement au bien-être des consommateurs, est une pratique avérée de l'économie industrielle [1]. La collusion peut prendre des formes variées, difficilement détectables *ex post* : entente sur les prix, répartition des marchés, coordination des annonces, attribution de quotas de production...

Régulation et corrélation : la menace de la collusion

Ainsi que nous l'avons déjà souligné de nombreuses fois dans le chapitre précédent, les mécanismes de concurrence par comparaison semblent particulièrement sensibles à la collusion. Pouyet (2002) [85] rappelle que la concurrence par comparaison engendre des externalités informationnelles entre les concurrents. « En conséquence, ces concurrents vont souhaiter former un cartel collusif afin de coordonner leurs décisions et de contrer les effets d'extraction de rente des mécanismes de concurrence par comparaison. » Or, nous avons noté que plus les agents sont corrélés, plus la concurrence par comparaison est intéressante pour la collectivité. D'un autre côté, la réduction des rentes traditionnellement abandonnées aux entreprises motive les comportements d'entente. Et ceux-ci se développent d'autant plus facilement que les entreprises sont proches, et donc corrélées. On pressent donc une relation intrinsèque entre la puissance d'un mécanisme comparatif et les incitations à la collusion.

Au delà des incitations à la collusion, les effets mêmes de celle-ci sur l'efficience des mécanismes comparatifs sont à craindre. Dans la mesure où les incitations résultent de la comparaison des informations émises par les entreprises, la coordination de ces dernières annihile l'efficience du mécanisme. En effet, dans ce cas les informations sur lesquelles travaille le régulateur ne signalent plus la performance de l'entreprise. Au contraire, elles correspondent à des valeurs artificiellement choisies, afin de maximiser l'utilité collective d'un cartel.

Bien que la plupart des auteurs que nous avons déjà cités mentionnent la menace de collusion, seule Auriol (2000) [59] propose un modèle qui soit initialement robuste à

[1] Pour un exposé synthétique des problèmes de collusion, voir Linnemer & Souam (1997) [78].

la collusion. Nous proposons dans ce début de section d'analyser les effets généraux de la collusion sur les mécanismes comparatifs. Nous distinguons deux cas, selon que les entreprises forment le cartel collusif après ou avant la découverte de leur caractéristique de productivité. À partir du cadre proposé par Laffont & Martimort (2000) [73], nous envisageons donc l'impact de la collusion suivant qu'elle intervient *ex post* ou *ex ante*.

Collusion *ex post* : le modèle de Laffont & Martimort (2000)

Laffont & Martimort (2000) [73] proposent une relation d'agence, robuste à la collusion, bien que le principal exploite la corrélation entre deux agents. Ils considèrent le cas où les deux agents sont invités à révéler leur préférence pour la fourniture d'un bien public. Ils considèrent pour cela le cadre général de Crémer & McLean (1988) [66]. Dans le chapitre précédent, nous avions présenté le modèle de Crémer et McLean à l'aide de Pouyet (2002) [85] qui l'applique à la régulation. Nous avions alors précisé que l'article de Pouyet traitait plus particulièrement du rôle de la corrélation dans les incitations à la collusion. En effet, Pouyet a repris le modèle de Laffont et Martimort, afin de l'appliquer à une relation d'agence régulateur - entreprises. Nous nous référons donc plus explicitement au modèle de Pouyet pour expliquer le mécanisme envisagé par Laffont et Martimort.

On considère deux entreprises corrélées et neutres au risque ; leur caractéristique de productivité, discrète, peut prendre deux valeurs : $\beta \in \{\underline{\beta}; \overline{\beta}\}$. Le régulateur confie la production à l'entreprise la plus productive, ou, en cas d'annonces identiques, partage la production entre les deux agents. Laffont et Martimort imaginent que la collusion intervient par le truchement d'un « tiers ». Cet entremetteur propose aux entreprises de maximiser la somme de leurs utilités, en leur indiquant quelle caractéristique annoncer au régulateur. Il leur offre donc un contrat, dit « parallèle », suivant lequel elles doivent lui indiquer leur caractéristique[2], afin qu'il puisse optimiser les annonces à faire au régulateur. Ce contrat prévoit également les transferts par lesquels les gains de la collusion sont partagés entre les entreprises.

Rappelons les gains et pertes auxquels s'exposent les entreprises à l'issue du jeu, dans ce cadre de Crémer et McLean. Lorsque les caractéristiques sont positivement corrélées, le profit $\pi_{\beta_i;\beta_j}$ de l'entreprise i, de type β_i, face à l'entreprise j de type β_j, vérifie selon les cas :

$$\begin{aligned}
\pi_{\underline{\beta};\underline{\beta}} &\leqslant 0 \\
\pi_{\underline{\beta};\overline{\beta}} &> 0 \\
\pi_{\overline{\beta};\underline{\beta}} &< 0 \\
\pi_{\overline{\beta};\overline{\beta}} &\geqslant 0
\end{aligned}$$

[2] Cet agent, de même que le régulateur, ignore initialement les caractéristiques des entreprises ; nous verrons que cette asymétrie d'information a des répercussions sur la robustesse de la collusion.

Dans ce cadre, les incitations à former un cartel collusif dépendent donc des caractéristiques des entreprises de la façon suivante :

- **Lorsque les deux entreprises sont efficientes,** leur utilité *ex post* est négative ($\pi_{\underline{\beta};\underline{\beta}} \leqslant 0$ pour chacune). Supposons qu'une des entreprises annonce une caractéristique $\overline{\beta}$ qui n'est pas la sienne. Elle engendre alors une externalité informationnelle positive en faveur de l'autre, dont le profit $\pi_{\underline{\beta};\overline{\beta}}$ devient positif. Bien qu'elle-même continue d'obtenir un profit négatif ($\pi_{\overline{\beta};\underline{\beta}}$), l'utilité du cartel s'en trouve améliorée : $2\pi_{\underline{\beta};\underline{\beta}} < \pi_{\underline{\beta};\overline{\beta}} + \pi_{\overline{\beta};\underline{\beta}}$. Ce résultat est encore renforcé si les deux entreprises décident de mentir au régulateur, puisque l'utilité de chacune devient positive (on passe de $\pi_{\underline{\beta};\underline{\beta}} \leqslant 0$ à $0 \leqslant \pi_{\overline{\beta};\overline{\beta}}$).
- **Lorsque les entreprises ont des caractéristiques différentes,** elles ont intérêt à ce que l'entreprise efficiente se fasse passer pour inefficiente.
- **Lorsque les deux entreprises sont inefficientes,** il n'est pas dans l'intérêt du cartel que l'une d'elles mente au régulateur.

Il existe donc de multiples opportunités en faveur de la formation d'un cartel. Afin d'en annuler les conséquences, Laffont et Martimort définissent un *principe de robustesse faible à la collusion*. Le contrat « principal » offert par le régulateur est faiblement robuste à la collusion si, à son équilibre bayésien, le contrat parallèle optimal est nul (c'est-à-dire qu'il aboutit à reporter les caractéristiques réelles au régulateur). Ce principe est similaire au principe de révélation de Baron & Myerson (1982) [61]. Il s'agit pour le régulateur de se mettre à la place de l'entremetteur en intégrant un certain nombre de contraintes dans son programme d'optimisation. De cette façon, Pouyet note que « le contrat principal est tel qu'un cartel sera systématiquement formé, mais que l'entremetteur recommandera aux entreprises de révéler honnêtement leur caractéristique. »

L'introduction des contraintes de robustesse à la collusion dans le programme du régulateur conduit celui-ci à distordre la quantité à produire par deux entreprises inefficientes. On retrouve le résultat classique qui consiste à ne pas distordre par rapport à l'optimum de premier rang le travail des entreprises efficientes. Pouyet remarque néanmoins que ce résultat de second rang tend vers l'optimum avec la corrélation des entreprises[3].

Ainsi, dans une configuration de Crémer et McLean, où le régulateur exploite la corrélation entre les entreprises, la menace de collusion réduit l'efficience du mécanisme. Mais ce résultat doit être relativisé, dans la mesure où il s'atténue lorsque la corrélation augmente. Soulignons également que la tentation collusive dans ce cadre semble d'autant plus réduite que les entreprises sont corrélées. On se rappelle en effet que plus les entreprises sont semblables, moins le risque de quitter le jeu avec une désutilité élevée est grand. Inversement, deux entreprises faiblement corrélées risquent d'importantes pénalités dans le jeu principal. Elles seront davantage tentées

[3] Rappelons qu'il s'agit ici de corrélation positive. Mais Pouyet étudie également le cas de la corrélation négative entre les entreprises. Il montre que, dans ce cas, la collusion n'est pas très menaçante pour le régulateur, puisqu'incitations individuelles et collectives s'opposent alors systématiquement.

de recourir aux services d'un entremetteur, pour s'assurer contre le risque.

Collusion *ex ante* : le modèle de Tangerås (2002)

Tangerås (2002) [89] complête le travail de Laffont et Martimort en considérant un déroulement du jeu différent. Dans les modèles de Laffont-Martimort et Pouyet, nous avons vu que le cartel collusif se formait après que les entreprises aient découvert leur caractéristique. En ce sens, on parle de collusion *ex ante* lorsque les entreprises s'engagent à coopérer lorsqu'elles n'ont encore qu'une connaissance imprécise de leur productivité. Dans ce cas, il n'y a pas d'asymétrie informationnelle entre les entreprises, lors de la formation du cartel.

L'analyse de Tangerås s'inscrit dans le cadre du modèle d'Auriol (2000) [59]. Les entreprises ont la possibilité de souscrire un contrat parallèle à celui du régulateur. Avec des hypothèses légèrement différentes de celles que nous avons exposées dans la section 3.4, Tangerås retrouve des résultats similaires en l'absence de collusion. Bien que les entreprises ignorent leur caractéristique, elles anticipent la réduction de leurs rentes, due au mécanisme comparatif du régulateur. Elles sont donc tentées de s'entendre *ex ante*, afin d'annoncer une caractéristique commune $b = \overline{b}$ qui maximise leur rente informationnelle. Tangerås reprend le principe de robustesse faible à la collusion de Laffont et Martimort. L'intégration de contraintes supplémentaires permet de définir un contrat principal robuste à la collusion. En terme de bien-être, l'issue de ce contrat dépend alors de la possibilité qu'ont les entreprises de procéder à des paiements parallèles entre elles.

Lorsque les entreprises n'ont pas accès aux transferts parallèles, Tangerås montre que le régulateur peut s'opposer sans coût à la collusion. Il lui suffit de donner au jeu une configuration de « dilemme du prisonnier » en menaçant de récompenser l'entreprise qui annonce honnêtement $b = \underline{b}$ et de pénaliser l'autre, en cas d'annonces incompatibles[4]. En revanche, lorsque les entreprises peuvent procéder à des transferts entre elles, les contraintes de robustesse à la collusion deviennent coûteuses. Il n'est plus possible d'atteindre la solution classique de second rang : le régulateur est contraint de céder une rente aux entreprises. Tangerås montre que cette rente tend à compenser les gains de la concurrence par comparaison lorsque la corrélation entre les entreprises tend à être parfaite ($\alpha \to 1$).

Tangerås obtient ainsi des résultats opposés à ceux de Pouyet. Cette contradiction apparente s'explique par les hypothèses différentes de déroulement du jeu. Lorsque la collusion intervient *ex post*, le régulateur peut tirer parti de l'asymétrie d'information entre les entreprises. Notamment, tout refus du contrat parallèle par une entreprise envoie un signal informatif à l'autre. En revanche, lorsque la collusion intervient *ex ante*, le régulateur ne peut plus exploiter ce décalage informationnel. Les rentes demeurent coûteuses, et les incitations à la collusion sont d'autant plus fortes que

[4] Il s'agit donc de rompre la symétrie des pénalités infinies, proposées par Auriol & Laffont (1992) [56]. Cette disposition ne modifie en rien la solution à l'équilibre du jeu, puisqu'elle concerne un état instable du jeu.

les entreprises sont corrélées. Cela justifie l'écart observé entre les résultats des deux modèles.

Il est toutefois difficile de déterminer lequel de ces deux types de modèle est le plus pertinent. D'un côté, il semble plus concevable que le cartel collusif se forme après que les entreprises aient découvert leur caractéristique (*ex post*, plutôt qu'*ex ante*). D'un autre côté, la modélisation de la collusion *ex post* par Pouyet fait appel au cadre de Crémer et McLean qui semble peu opérationnel dans notre contexte de régulation. Enfin, le principe de robustesse à la collusion qui est mis en œuvre est « faible », c'est-à-dire qu'il ne garantit pas l'unicité de l'équilibre robuste à la collusion. Ainsi, le comportement d'entente explicite entre les entreprises réduit en général sensiblement les effets des mécanismes comparatifs.

En pratique, il semble souhaitable d'assurer, autant que possible, que les entreprises ne seront pas tentées par un cartel collusif. Par exemple, en augmentant le nombre d'entreprises impliquées dans le mécanisme de régulation, le régulateur réduit les opportunités de collusion explicite. En effet, la multiplication des agents accroît la probabilité de déviation de l'équilibre collusif, et diminue la stabilité du cartel. Une solution, plus sécurisante encore, consiste à comparer les annonces passées des entreprises, avant qu'elles n'aient pu établir un cartel.

Mais si les entreprises doivent être particulièrement suspectées d'entente, le régulateur n'est pas au-delà de tout soupçon, dans la mesure où il peut être capturé. Nous considérons dans les paragraphes qui suivent, comment l'entente entre le régulateur et les entreprises peut perturber le mécanisme comparatif.

4.1.2 La capture du régulateur

Indirectement, nous avons déjà abordé dans le chapitre précédent le cas où un régulateur pouvait être capturé. Bivand & Szymanski (1997) [62] [5] envisageaient dans un contexte multi-principal, la capture d'un régulateur par l'entreprise qui lui fait face. Nous avions noté l'externalité informationnelle négative qui résultait d'une entente entre un régulateur et une entreprise. Dans un cadre multi-principal, nous avions conclu qu'il était nécessaire que le mécanisme comparatif soit mis en œuvre par une agence commune des régulateurs. Mais la question de la capture de cette agence commune, assimilable à un unique régulateur, reste posée dans les mêmes termes que dans un contexte mono-principal. C'est le problème que nous examinons dans les paragraphes qui suivent.

Régulation et vénalité : la tentation de la capture

La nouvelle économie de la réglementation a intégré la question de l'intérêt personnel du régulateur en proposant un cadre intellectuel qui réinterprète la problématique de la capture. Le régulateur est ainsi modélisé comme un collecteur d'information pour

[5] Nous avons discuté leur modèle à propos de l'hétérogénéité endogène. Ils envisagent le cas où le régulateur subventionne l'entreprise de façon excessive (voir page 95).

le compte du gouvernement ; il bénéficie donc lui-même vis-à-vis de ce dernier d'une asymétrie informationnelle dont il peut tirer parti. Un régulateur vénal cherchera à favoriser certaines entreprises en dissimulant de l'information, en échange de discrètes compensations.

Le modèle de Contreras & Rickman (2004)

Contreras & Rickman (2004) [65] reprennent le modèle d'Auriol & Laffont (1992) [56] en intégrant une contrainte politique. Le gouvernement qui définit le jeu et distribue les transferts n'a pas les moyens, en termes d'expertise, de discerner les caractéristiques des entreprises[6]. Il confie alors cette tâche à un régulateur, chargé de lui transmettre ses observations. Mais celui-ci n'observe qu'imparfaitement la corrélation entre les entreprises, α ; il arrive qu'il ne puisse pas déterminer la valeur de ce paramètre. Le problème naît de ce que les entreprises peuvent capturer le régulateur. Elles ont intérêt à ce que le régulateur rapporte au gouvernement qu'il n'a pu observer leur corrélation ; cela préserve l'asymétrie d'information et les rentes associées. Les auteurs comparent leurs résultats à ceux d'Auriol et Laffont, suivant que le régulateur est désintéressé, ou bien vénal et donc « capturable » par les entreprises.

Contreras et Rickman caractérisent l'imparfaite supervision des entreprises par les performances suivantes :

- le régulateur observe un signal $\sigma = \alpha$ avec une probabilité ξ, et $\alpha > 0$ avec une probabilité z ;
- le régulateur observe un signal non informatif $\sigma = \emptyset$ avec une probabilité $(1-\xi)$, et $\alpha = 0$ avec une probabilité $(1-z)$; de plus, $\alpha = 0 \Rightarrow \sigma = \emptyset$.

Dans le cas d'un régulateur désintéressé, le gouvernement est donc utilement informé avec une probabilité ξz. Cette structure d'observation imparfaite réduit déjà les résultats d'Auriol et Laffont, que l'on ne retrouve que pour $\xi z = 1$.

Lorsque le régulateur est vénal, il peut prétendre auprès des entreprises à une part k de leur rente $\hat{s}$. Cette part dépend des coûts de transaction associés à la capture λ_c :

- lorsque ces coûts de transaction sont importants (ce qui correspond à une détection du transfert), $\lambda_c \to \infty$, le régulateur ne peut être capturé : $k \to 0$;
- lorsque les transferts des entreprises au régulateur s'effectuent sans coût, $\lambda_c \to 0$, le régulateur est facilement capturé : $k \to 1$.

Afin d'éviter la capture d'un régulateur vénal, le gouvernement doit donc lui céder un transfert incitatif. Le montant de ce transfert, $\lambda \xi z k \hat{s}$, peut s'avérer très élevé lorsque k est proche de 1. Pour réduire cette rente, le gouvernement est contraint de distordre la production - comme en information asymétrique avec les rentes informationnelles. En diminuant la quantité à produire, le gouvernement réduit les rentes informationnelles potentielles des entreprises ($\hat{s}$), et donc, *in fine*, la rente du régulateur.

[6] Les auteurs justifient cette structure en faisant remarquer qu'« en l'absence de régulateur, bien que le gouvernement puisse suspecter une certaine corrélation, les entreprises peuvent toujours argumenter (devant un tribunal) que la corrélation est nulle, et le gouvernement n'a aucune information pour contre-argumenter ».

Finalement, Contreras et Rickman montrent qu'avec un régulateur vénal, le gain résultant des comparaisons est réduit. Cependant, l'effet d'échantillonnage défini par Auriol et Laffont [7] préserve l'intérêt qu'il peut y avoir en faveur d'un marché duopolistique.

Ainsi, les comportements d'entente explicite entre les agents économiques de la relation de régulation sont susceptibles de porter préjudice aux mécanismes comparatifs. Il ne s'agit cependant pas d'un résultat exceptionnel, dans la mesure où les phénomènes de collusion et de capture sont susceptibles de perturber n'importe quelle forme de régulation économique.

En pratique, la régulation peut aussi être perturbée par des comportements d'entente tacite. En effet, lorsque le jeu statique est répété successivement, les entreprises peuvent être incitées à dévier de l'équilibre concurrentiel. L'introduction de la dimension temporelle dans la relation régulateur-entreprises est donc nécessaire afin de poursuivre notre évaluation de la concurrence par comparaison.

4.2 Les effets dynamiques

Après une section consacrée aux comportements d'entente explicite, nous abordons donc la dimension temporelle de la régulation. En pratique, la relation régulateur-entreprises se réduit rarement à une seule période de jeu. Ceci est d'autant plus vrai lorsqu'il s'agit de la régulation d'une industrie de réseau, dont l'activité a vocation à perdurer. Mais la théorie microéconomique nous laisse entrevoir quelques effets dynamiques, susceptibles de réduire l'efficacité de la concurrence par comparaison.

À la section précédente, nous avons mentionné la collusion tacite, que nous détaillons dans les paragraphes suivants. Mais outre ce phénomène, la capacité d'engagement limitée du régulateur contribue aussi à réduire le bénéfice attendu de la régulation statique. Nous verrons à travers l'effet de cliquet la réaction des entreprises lorsqu'elles anticipent que le régulateur adoptera un comportement opportuniste.

4.2.1 La collusion tacite

Ainsi que nous l'avons précisé, la collusion entre les entreprises n'intervient pas forcément aussi explicitement que dans la section précédente. Elle peut également apparaître tacitement, lorsque les entreprises interagissent régulièrement. Contrairement à la collusion explicite, la collusion tacite n'est pas condamnable en soi, mais ses effets - tout aussi néfastes - doivent être anticipés et évités.

[7] Rappelons ce qu'est l'effet d'échantillonnage : lorsque le régulateur opte pour un marché duopolistique, il choisit de confier la production à l'entreprise qui s'annonce comme étant la plus productive. Avec un marché duopolistique, il a donc plus de chance de faire face à une entreprise dotée d'une caractéristique particulière efficiente (ϵ faible) qu'avec un marché monopolistique.

Régulation et répétition : l'apparition de nouveaux équilibres

À court terme, les mécanismes de régulation sont conçus afin que l'unique équilibre du jeu corresponde à l'optimum souhaité par le régulateur. L'équilibre est déterminé en conjecturant que chaque entreprise adopte un comportement qui maximise son utilité espérée. Dans un modèle statique, aucun agent n'a intérêt à dévier de ce couple stratégie-équilibre. Étant *a priori* le seul à dévier, il anticipe une perte d'utilité, et donc, se conforme à la stratégie souhaitée par le régulateur. Les contraintes d'incitation, introduites dans le programme d'optimisation du régulateur, permettent d'atteindre ce résultat.

Mais lorsque le jeu est répété sur plusieurs périodes, les entreprises cherchent désormais à maximiser leur utilité de long terme. Or, rien ne garantit que la solution qui maximise l'utilité de long terme soit la répétition de la solution du jeu statique, qui maximise l'utilité de court terme. La dynamique du jeu fait donc apparaître de nouveaux équilibres sous-optimaux. Dans ce contexte, les entreprises peuvent maintenir des prix élevés, en s'accordant implicitement sur les conséquences néfastes qu'aurait une déviation de cet équilibre collusif [8].

Classiquement, lorsque le jeu est répété de nombreuses fois, les agents convergent rapidement vers l'équilibre collusif, qu'ils maintiennent pendant l'essentiel du jeu. En revanche, lorsque le jeu dynamique approche de son terme, les agents retrouvent intérêt à dévier de l'équilibre collusif. En effet, la déviation vers l'équilibre concurrentiel (qui ne peut intervenir qu'une fois) assure à celui qui dévie en premier une très forte utilité. Quelques périodes avant la fin du jeu dynamique, l'équilibre redevient donc concurrentiel. Dans les paragraphes qui suivent, nous analysons un exemple, testé expérimentalement, de collusion tacite entre agents soumis à la concurrence par comparaison.

Le modèle de Potters, Rockenbach, Sadrieh *et al.* (2004)

Potters, Rockenbach, Sadrieh *et al.* (2004) [84] testent expérimentalement l'intuition suivant laquelle plus le régulateur extrait de surplus en appliquant la concurrence par comparaison, plus les entreprises sont tentées par la collusion tacite. Reprenant le cadre du modèle de Shleifer dans une configuration bi-agent, ils considèrent deux schémas de concurrence par comparaison :

- un mécanisme « discriminant » qui correspond à la règle de tarification suggérée par Shleifer, à partir des coûts annoncés : $p_i \leqslant c_{j \neq i}$;
- un mécanisme « uniforme », intermédiaire entre le précédent et un contrat *cost plus* : $p_i \leqslant (c_i + c_{j \neq i})/2$. Ce deuxième schéma est évidemment moins incitatif que le premier.

Ces deux mécanismes sont caractérisés par un équilibre de Nash symétrique non-coopératif, avec un prix, des coûts et une utilité de l'entreprise faibles. Mais il existe un autre équilibre, collusif, avec des prix et coûts plus élevés, assurant aux entreprises une utilité maximale.

[8] À propos de cette forme de collusion, voir Ivaldi, Jullien, Rey *et al.* (2003) [72].

Potters, Rockenbach, Sadrieh *et al.* comparent expérimentalement le comportement de binômes d'agents économiques, selon qu'ils sont soumis à l'un ou l'autre des mécanismes.

Quelques précisions s'imposent, concernant l'unique mise en œuvre des techniques de l'économie expérimentale dont nous faisons part dans ce travail. À défaut de pouvoir porter un quelconque jugement sur la méthodologie employée, nous décrivons brièvement l'expérience réalisée. Les agents qui participent à l'expérience sont des étudiants de licence, majoritairement en économie-gestion. Ils ne participent qu'à une série de jeu répété, et n'ont jamais pris part à de telles expériences auparavant. Ils sont d'abord associés en binôme aléatoirement. Puis, ils doivent répéter 50 fois de suite le même jeu, correspondant à un mécanisme discriminant (16 binômes) ou uniforme (13 binômes). Les agents peuvent effectuer des simulations numériques, afin de connaître rapidement l'impact de telle décision sur leur utilité. À l'issue des 50 épreuves, ils reçoivent un équivalent monétaire des gains qu'ils ont accumulés.

Les résultats de cette expérience montrent que les joueurs soumis à la pure concurrence par comparaison annoncent des coûts significativement plus élevés (proches des annonces collusives) que ceux soumis au mécanisme uniforme. L'analyse des auteurs démontre que le comportement des joueurs du mécanisme discriminant est significativement plus collusif que celui des autres. Du fait de ces déviations collusives, le bien-être moyen est supérieur avec le mécanisme uniforme, pourtant moins incitatif. Ainsi, la répétition d'un jeu de concurrence par comparaison conduit les agents à adopter un comportement collusif, et ce d'autant plus que le mécanisme est puissant. Les auteurs notent cependant que l'augmentation du nombre d'agents impliqués dans le modèle, devrait réduire cet écart. Cela est dû à la dilution de la part *cost plus* dans le mécanisme uniforme qui tend vers le mécanisme discriminant.

Mais les entreprises ne sont pas les seuls agents économiques à vouloir tirer parti de la répétition dans le temps du jeu statique. Le régulateur qui apprend la caractéristique des entreprises lors de la première période de jeu est également tenté par l'opportunisme. En l'occurence, il est incité à exploiter ultérieurement cette information, à moindre coût. Nous analysons dans les paragraphes qui suivent l'effet de cliquet : comment les entreprises, et donc les mécanismes comparatifs, s'adaptent à la capacité d'engagement limitée du régulateur ?

4.2.2 L'effet de cliquet

Avec l'effet de cliquet, nous abordons la question de la capacité d'engagement limitée du régulateur, que nous retrouverons dans la section suivante, consacrée aux investissements. Là encore, la prise en compte de la crédibilité du régulateur constitue un des apports fondamentaux de la nouvelle économie de la réglementation.

Régulation et révélation : la menace de *ratchet effect*

Dans un cadre statique de sélection adverse, nous avons rappelé que les entreprises étaient amenées à révéler leur productivité, en échange d'une rente informationnelle. Toutefois, lorsque l'on considère la régulation dans un cadre dynamique, le modèle statique ne peut pas être transposé simplement à chaque période. En effet, à l'issue de la première période, le régulateur étant informé de la productivité de l'agent, il est incité à exploiter celle-ci sans coût. Ainsi, plutôt que de maintenir un contrat révélateur coûteux (alors qu'il n'a plus rien à apprendre), il préférera offrir un contrat sans rente informationnelle. Mais l'agent anticipe ce comportement opportuniste ; il préfère alors réduire sa rente en première période, afin de conserver une part d'information stratégique pour la suite. Le problème naît donc de ce que le régulateur ne peut pas s'engager crédiblement à ne pas agir de façon opportuniste. De ce fait, l'agent, craignant une réévaluation du contrat, tend à ne révéler que tardivement sa caractéristique. C'est ce qu'on appelle l'effet de cliquet (*ratchet effect*).

Les contrats de long terme de Faure-Grimaud & Reiche (2003)

Dans un modèle de sélection adverse particulier, Faure-Grimaud & Reiche (2003) [69] montrent que la conclusion de contrats de long terme (qu'ils spécifient) résout le problème d'engagement. Dans le cadre de Crémer & McLean (1988) [66], ils montrent qu'une série de contrats de court terme permet d'atteindre la solution de premier rang à chaque période. Ils considèrent un modèle avec deux agents neutres au risque, dont les caractéristiques $\beta^i \in \{\underline{\beta}; \overline{\beta}\}$ sont corrélées. Le mécanisme qu'ils proposent consiste à offrir en première période un contrat comparatif révélateur de long terme. Afin de prévenir l'effet de cliquet, la contrainte d'incitation des entreprises efficientes fait intervenir la somme actualisée des utilités qu'elles auraient si elles s'annonçaient inefficientes. Ce contrat fixe la production à son niveau optimal, et ne laisse, en moyenne, aucune rente aux entreprises. Lors des périodes suivantes, toute l'information ayant été révélée, le régulateur se contente de maintenir le niveau de production optimal, et de rembourser les coûts correspondants [9].

Ce mécanisme, qui extrait la totalité des rentes dès la première période, n'est plus conditionné par la capacité d'engagement du régulateur. Toutefois, ces résultats séduisants reposent sur la neutralité au risque des entreprises qui sont menacées d'une pénalité élevée à la première période. En effet, nous avons vu avec le modèle de Pouyet (2002) [85] qu'une entreprise qui reporte une caractéristique inefficiente $\overline{\beta}$, tandis que l'autre s'annonce efficiente avec $\underline{\beta}$, est fortement pénalisée. Cette entreprise est pourtant susceptible d'être honnête [10].

[9] Ce traitement dissymétrique justifie, selon les auteurs, les rares mises en œuvre observées de la concurrence par comparaison. Celle-ci n'intervient, en fait, qu'à la première période.

[10] Rappelons que nous ne sommes pas, ici, dans la même configuration qu'Auriol & Laffont (1992) [56]. La caractéristique β du modèle de Faure-Grimaud et Reiche n'est pas commune aux deux entreprises, contrairement au b d'Auriol et Laffont. Il est donc probable que deux agents annoncent honnêtement des caractéristiques différentes.

Conscients du coût élevé supporté par une entreprise pénalisée en première période, les auteurs envisagent que le montant des transferts négatifs soit limité. Ils montrent alors que seul un contrat de long terme, qui étale la pénalité tout au long de la période, permet de retrouver la solution de premier rang. Signalons également qu'ils proposent une variante à leur mécanisme, lorsqu'une seule entreprise produit sur le marché. À condition qu'une deuxième entreprise menace d'entrer plus tard sur le marché, ils montrent qu'un contrat de long terme permet, là encore, d'atteindre l'optimum de premier rang[11].

De même que pour les modèles statiques élémentaires, nous préférons nous concentrer sur un modèle qui n'impose pas que les agents neutres au risque participent à une loterie. Nous détaillons donc, dans les paragraphes qui suivent, le modèle de Meyer & Vickers (1997) [80].

Le modèle de Meyer & Vickers (1997)

Meyer & Vickers (1997) [80] considèrent un mécanisme de comparaison bi-agent, dans une configuration d'aléa moral et de sélection adverse *ex ante*. Rappelons que dans un contexte de sélection adverse *ex ante*, la conclusion du contrat intervient avant que l'entreprise ait connaissance de sa caractéristique exogène. Meyer et Vickers distinguent l'usage, implicite ou explicite, des comparaisons par le régulateur. Le recours « implicite » aux comparaisons consiste à fixer les gains de productivité d'un mécanisme de *price cap* à l'aide de celles-ci[12]. Le recours « explicite » aux comparaisons correspond à une mise en œuvre plus formelle de la concurrence par comparaison[13].

Cas des comparaisons implicites : Meyer et Vickers considèrent que chaque entreprise i, neutre au risque, est chargée d'une production unitaire pour deux périodes. Sa fonction d'utilité s'écrit :

$$U_i = \sum_{t=1}^{2} p_{i,t} - C_{i,t} - \psi(e_{i,t}) = \sum_{t=1}^{2} p_{i,t} - k + (\beta^i - e_{i,t} + \theta_{i,t}) - \psi(e_{i,t})$$

où p, C et $\psi(.)$ désignent respectivement le prix, le coût de production et la désutilité à l'effort,
k correspond aux coûts fixes,
β est la caractéristique de la firme,
e est l'effort fourni par l'entreprise,
et θ est l'aléa exogène d'espérance nulle.

[11] L'idée consiste à réguler avec des contrats *cost plus* l'opérateur historique tant qu'il est seul sur le marché. Lorsqu'un nouvel entrant apparaît, le contrat devient comparatif, comme dans le cadre duopolistique, avec la possibilité de pénalités rétroactives. Anticipant cela, l'opérateur historique révèle honnêtement sa productivité, dès la première période.

[12] Cette pratique correspond à celle du régulateur britannique chargé de la distribution et de l'assainissement de l'eau.

[13] Signalons également que Meyer et Vickers déclinent leur modèle selon qu'il est appliqué à la régulation d'entreprises ou à la relation dirigeant-actionnaires.

Ici, le régulateur utilise les comparaisons de façon implicite afin de fixer le prix pour la prochaine période. Il fixe ce prix au niveau du coût espéré, conditionnellement au résultat antérieur de l'entreprise, et à celui de sa concurrente. Dans le cas des contrats individuels, les auteurs montrent que l'effort de première période est sous-optimal, en raison de l'effet de cliquet. En effet, cet effort contribue à réduire le prix de deuxième période, à cause de l'information qu'il véhicule sur la caractéristique β. Ainsi, le recours aux comparaisons peut être bénéfique s'il apporte au régulateur une information plus précise sur la caractéristique β. Le régulateur fixe alors le prix de deuxième période en accordant moins de poids au résultat de l'entreprise en première période. Dans ce cas, l'effet de cliquet est atténué.

Meyer et Vickers démontrent que l'usage implicite des comparaisons améliore la solution des contrats individuels pour $\mathrm{corr}(\beta^i;\beta^j) > \mathrm{corr}(\theta_i;\theta_j)$. L'effet de cliquet est atténué par les comparaisons lorsque les caractéristiques des entreprises sont davantage corrélées que les aléas [14].

Cas de la concurrence par comparaison : Nous considérons à présent le cas où le régulateur assoit explicitement le contrat d'une entreprise sur les performances d'un concurrent, en plus des siennes. Un tel contrat engendre, d'une part, un effet statique d'assurance contre le risque grâce à la concurrence par comparaison. D'autre part, il introduit un effet de cliquet dynamique. Meyer et Vickers montrent que le surplus peut être additivement décomposé en deux termes, correspondant chacun à l'un de ces effets. Les agents pouvant désormais présenter de l'aversion au risque, le terme d'assurance contre le risque est toujours positif. En revanche, l'effet de cliquet peut compenser ce gain.

Meyer et Vickers ne parviennent pas à caractériser analytiquement les conditions favorables à la concurrence par comparaison. Toutefois, ils remarquent qu'en plus de quelques cas particuliers, le surplus est positif dans les conditions suivantes :

- $\mathrm{corr}(\beta^i;\beta^j) > \mathrm{corr}(\theta_i;\theta_j) > 0$; on retrouve la condition précédente : la corrélation des caractéristiques doit être supérieure à celle des aléas,
  et
- les entreprises sont faiblement neutres au risque.
  ou
- $\mathrm{var}(\theta) \geqslant \left[1/(\sqrt{3}-1)-1\right] \mathrm{var}(\beta)$; la variance des aléas ne doit pas être trop inférieure à celle des caractéristiques.

Les auteurs proposent ensuite quelques simulations numériques de résultats en fonction de ces paramètres.

Nous retenons donc, une fois de plus, qu'une forte corrélation entre les entreprises est nécessaire pour assurer l'efficience du mécanisme comparatif. Notons cependant

[14] Ce résultat est valable pour des corrélations positives, mais les auteurs l'expriment de façon générale, quel que soit le signe des corrélations. Signalons également que ce résultat est l'opposé de celui obtenu pour la relation dirigeant-actionnaires. En effet, dans ce cas, une réputation de forte productivité permet au dirigeant d'augmenter sa rémunération sur les périodes suivantes. Il est ainsi incité à maximiser son résultat de première période de façon à laisser croire qu'il est très productif.

qu'il est possible de supprimer totalement l'effet de cliquet, en rompant la dépendance entre les coûts de l'entreprise et sa rémunération. Cela correspond au modèle préconisé par Shleifer. En pratique, un nombre élevé d'entreprises impliquées dans le mécanisme permet de diluer l'effet individuel dans la moyenne du secteur.

L'introduction de la dimension temporelle dans notre analyse de la concurrence par comparaison fait donc apparaître de nouveaux effets dont le régulateur doit se prémunir. Il convient en particulier d'être particulièrement vigilant face à l'apparition de la collusion tacite. L'effet de cliquet, en revanche, peut facilement être atténué lorsque la concurrence par comparaison est mise en œuvre judicieusement. Dans ce cas, les mécanismes comparatifs réduisent les effets de la capacité limitée d'engagement du régulateur. Mais cette capacité limitée d'engagement peut avoir d'autres conséquences, notamment en terme d'incitations à investir. C'est la raison pour laquelle nous analysons dans la dernière section la question des investissements, lorsque les entreprises sont régulées par la concurrence par comparaison.

4.3 Les incitations à investir

La régulation, telle que nous l'avons abordée jusqu'ici, vise à encadrer les marchés de façon à maximiser le surplus collectif produit au cours du jeu. Toutefois, les progrès technologiques permettent aux entreprises d'améliorer le surplus créé par leur activité marchande. Pour cela elles doivent investir, notamment en matière de recherche et développement, afin d'améliorer leur productivité ou la qualité de leur production. Ainsi, il est nécessaire que la forme de régulation à laquelle les entreprises sont soumises ne réduise pas leurs incitations à investir, en même temps qu'elle réduit leurs rentes. Nous analysons donc dans cette section la compatibilité des mécanismes comparatifs avec les incitations à l'investissement. Nous distinguons d'abord les investissements de productivité, que le régulateur peut être tenté de s'approprier, puis les investissements en faveur de la qualité.

4.3.1 Les investissements de productivité

Dans les modèles que nous avons discutés jusqu'à présent, la technologie utilisée par les entreprises était modélisée par un paramètre d'efficience exogène, β. Cette caractéristique, fixe dans les modèles statiques, peut néanmoins être améliorée par des investissements de productivité. Dans les paragraphes qui suivent, nous détaillons dans quelle mesure la concurrence par comparaison favorise ces investissements. Trois contributions ont traité cette problématique entre 1997 et 1999. Nous verrons que deux concepts, « *hold-up* » et *spillovers*, qui se rejoignent en concurrence par comparaison, menacent de désinciter les entreprises à investir

Régulation et investissements : la menace de « *hold-up* »

En information complète, le régulateur observe la caractéristique productive des entreprises. Cela lui permet d'extraire totalement la rente informationnelle dont elles

seraient susceptibles de jouir. Les entreprises n'ont donc aucune incitation à investir pour améliorer leur productivité. Elles anticipent en effet qu'elles seront expropriées de leur retour sur investissement par un « *hold-up* » du régulateur. Ce problème de sous-investissement naît de ce que le régulateur ne dispose pas d'une capacité d'engagement suffisante, pour assurer l'entreprise d'investir sans risque d'expropriation. Cette question peut être appréhendée comme la transposition sur le long terme de l'effet de cliquet étudié précédemment. Nous avons vu que l'effet de cliquet incite l'agent à réduire son effort (de court terme) sous la menace d'une expropriation de sa rente informationnelle. De même, l'agent peut être incité à réduire son effort à long terme (ses investissements), par crainte de ne pas en récolter les bénéfices. Le régulateur fait ainsi face à un dilemme, entre l'extraction de la rente informationnelle (efficacité à court terme) et les investissements (efficacité à long terme).

Inversement, Laffont & Tirole (1993) [76] soulignent qu'en information incomplète, « lorsque le régulateur n'observe pas [la caractéristique] β avant d'offrir un contrat, l'entreprise est en partie protégée d'une expropriation de son investissement. »

Cette question des incitations à l'investissement se doit d'être particulièrement examinée sous un régime de concurrence par comparaison. Dans la mesure où la concurrence par comparaison permet de réduire l'asymétrie d'information et les rentes associées, elle est susceptible de conduire à un sous-investissement similaire à celui observé en information complète. Sobel (1999) [88] confirme cet impact négatif de la concurrence par comparaison sur l'investissement. Lorsque l'apport informationnel de la concurrence par comparaison est faible [15], il montre que le régulateur préférera inciter l'entreprise à maintenir un niveau d'investissement élevé, quitte à obtenir moins d'information. En revanche, si le coût social des fonds publics limite le recours aux contrats révélateurs classiques, la concurrence par comparaison devient pertinente et domine les effets négatifs sur l'investissement.

Nous montrons dans les paragraphes suivants comment les travaux de Powell & Szymanski (1997) [86] et surtout ceux de Dalen (1998) [67] éclairent ce résultat, en distinguant les caractéristiques des investissements.

Comparaison et investissements : le problème des *spillovers*

Powell & Szymanski (1997) [86] évaluent dans quelle mesure l'investissement bénéficie non seulement à l'entreprise qui le met en œuvre, mais aussi aux autres entreprises du secteur. Il s'agit de la problématique des *spillovers* (débordements), selon laquelle certains investissements dégagent une externalité positive auprès des autres entreprises [16]. Il est assez intuitif qu'un mécanisme d'évaluation de la performance relative incite les entreprises à investir, afin d'améliorer leur seule productivité. C'est la raison

[15] Cela peut être le cas s'il existe des mécanismes alternatifs et que le coût social des fonds publics est faible.

[16] Dans le cas d'une entreprise intégrée verticalement, il peut s'agir d'investissements sur une infrastructure utilisée par d'autres entreprises ; il peut aussi s'agir de certaines dépenses de recherche et développement, du recours à des consultants en management, *etc.*

pour laquelle les résultats de la concurrence par comparaison peuvent être perturbés par la présence de *spillovers*.

Powell et Szymanski reprennent le cadre du modèle de Shleifer en explicitant la façon dont le coût marginal peut être réduit par un investissement. Dans un contexte bi-agent, ils définissent la fonction de coût de l'entreprise i par :

$$c_i = 1 - \alpha e_i - (1 - \alpha) e_j$$

où e_k résulte de l'investissement $R(e_k) = e_k^2/2$ de l'entreprise k.
Le paramètre α correspond à la distribution des bénéfices de l'investissement entre les entreprises. Les auteurs montrent d'abord que si les investissements dégagent des externalités (apparition de *spillovers* pour $\alpha < 1$), la concurrence par comparaison ne permet plus d'atteindre l'optimum de premier rang du modèle de Shleifer. Le niveau d'investissement est alors d'autant moins optimal que les *spillovers* sont élevés (c'est-à-dire que α est faible).

Powell et Szymanski analysent ensuite le cas où le régulateur propose un mécanisme intermédiaire entre la concurrence par comparaison et le contrat *cost plus*. Ils étudient les distorsions qu'il convient d'appliquer à la tarification pour retrouver l'optimum. Toujours dans le cadre d'un modèle de Shleifer bi-agent, ils proposent donc de maintenir un transfert $T_i = R_j$, mais de fixer les prix $p_i = \beta c_i + (1 - \beta) c_j$ [17]. Les auteurs établissent alors qu'en fixant $\beta^* = 1 + 1/(1 - 2\alpha)$, l'optimum peut être atteint quel que soit $\alpha \neq 0,5$. Dans le cas le plus réaliste où $0,5 < \alpha \leqslant 1$ (l'entreprise bénéficie de l'essentiel de ses investissements), $\beta^* \leqslant 0$. La tarification qui semblait atténuée par l'introduction de la composante *cost plus*, renforce au contraire la composante *yardstick competition*, β étant négatif. Nous détaillons à présent le modèle de Dalen qui synthétise les deux problématiques de « *hold-up* » et de *spillovers*.

Synthèse : le modèle de Dalen (1998)

Dalen (1998) [67] reprend comme cadre général le modèle d'Auriol & Laffont (1992) [56]. Il considère toutefois que le coût de production peut être réduit par l'effort exercé par l'entreprise, que le régulateur n'observe pas. Il s'agit donc d'un cadre de sélection adverse doublé d'aléa moral, semblable à celui d'Auriol (2000) mais dans une configuration bi-agent. Dalen modélise les *spillovers* en distinguant :

- **Les investissements spécifiques à la firme :** en l'absence d'externalités (cas $\alpha = 1$ de Powell et Szymanski), l'investisement contribue à améliorer la composante particulière de la caractéristique de l'entreprise, ϵ_i ;
- **Les investissements spécifiques à l'industrie :** en présence d'externalités « totales » (cas $\alpha = 0,5$ de Powell et Szymanski), l'investissement bénéficie à la composante commune de la caractéristique, b.

[17] Cette forme hybride rappelle le menu de contrats proposé par Auriol (2000) [59] (voir page 107).

Intuitivement, il faut s'attendre à ce que la concurrence par comparaison favorise les investissements spécifiques à la firme, qui lui permettent d'accroître sa rente informationnelle. Inversement, les investissements spécifiques à l'industrie sont improductifs pour l'entreprise, puisque le régulateur extrait la rente informationnelle portant sur la composante commune de la caractéristique b.

Dalen considère d'abord le cas où le régulateur s'engage à traiter les entreprises avec des contrats individuels [18]. Il retrouve les résultats classiques de cette configuration, avec un niveau d'investissement inférieur à l'optimum [19].

Lorsque le régulateur a recours à la concurrence par comparaison (ou lorsqu'il ne peut s'engager à s'en dispenser), Dalen retrouve les résultats statiques de la concurrence par comparaison. Il confirme par ailleurs notre intuition :

- le niveau d'investissements spécifiques à l'industrie est nul ;
- le niveau d'investissements spécifiques à la firme est supérieur à celui obtenu avec les contrats individuels.

Le premier résultat s'explique par le « *hold-up* » du régulateur qui tirerait parti des *spillovers*. Le second résultat s'explique par la forte incitation qu'ont les entreprises, à dégager une rente informationnelle à partir de la composante de leur caractéristique qui demeure cachée.

Ainsi, chaque entreprise est tentée de réorienter ses investissements, afin qu'ils lui profitent davantage et moins aux autres. Cela lui permet de se démarquer des autres et donc de réduire l'efficience de la concurrence par comparaison. Aussi, lorsque les investissements du secteur sont majoritairement spécifiques à la firme, la concurrence par comparaison est toujours plus efficiente que les contrats individuels. En revanche, lorsque les investissements sont spécifiques à l'industrie, les gains du modèle statique d'Auriol et Laffont peuvent être compensés par le sous-investissement [20]. Ce sous-investissement a comme conséquence d'accroître la probabilité que les entreprises soient de type $\overline{b}$.

L'étude des incitations à investir en productivité montre donc qu'une trop forte proximité entre les entreprises crée des externalités sur les investissements, qui réduisent l'efficience de la concurrence par comparaison. Nous analysons à présent quelles sont les conséquences de la concurrence par comparaison sur la qualité du service offert.

4.3.2 Les investissements de qualité

Malgré les doutes qui se font jour autour de la qualité des services libéralisés, Auriol (1998) [58] note que peu de travaux théoriques portent sur la qualité dans un

[18] Il montre que cet engagement est nécessaire, dans la mesure où la concurrence par comparaison bénéficie toujours au régulateur, *ex post*.

[19] Dans ce cas, l'entreprise sait que le régulateur anticipe son investissement et réduira sa rente *ex post* ; voir Laffont & Tirole (1993) [76], p. 101.

[20] Le fait que les entreprises soient peu incitées à investir lorsque les externalités sont importantes rappelle la problématique des biens collectifs.

environnement régulé. Bien que la qualité soit un élément clef du fonctionnement économique des industries de réseau, elle ne se laisse pas appréhender facilement. Ses multiples dimensions et l'impossible vérification de certaines d'entre elles ont généralement conduit les théoriciens et les praticiens à travailler davantage sur la régulation des prix et quantités.

Régulation et qualité

Auriol montre que l'introduction de la concurrence sur le réseau d'infrastructure de l'opérateur historique induit des comportements de passager clandestin de la part des nouveaux entrants. Cet opportunisme entraîne alors un sous-investissement de l'opérateur historique sur le réseau. On retrouve le même mécanisme que dans le cas des *spillovers* précédemment évoqués. Dans la même logique que ses travaux précédents, Auriol analyse la structure endogène d'un marché de services substituables, en tenant compte de la qualité, selon qu'elle est vérifiable ou non. Toutefois, elle n'introduit pas la possibilité pour le régulateur de mettre en œuvre un mécanisme comparatif ; nous ne détaillerons donc pas davantage ce modèle.

Le modèle de Tangerås (2002)

La question de l'impact de la concurrence par comparaison sur la qualité de la production se pose dans les mêmes termes que pour d'autres formes de régulation. En accroissant la pression sur les coûts des entreprises, le régulateur ne risque-t-il pas d'inciter ces dernières à réduire la qualité ? Cette interrogation est d'autant plus vive concernant la concurrence par comparaison que celle-ci est appliquée à des activités sensibles à la qualité des prestations, tels les soins hospitaliers ou la distribution d'eau. C'est dans cette logique que s'inscrit la réflexion de Tangerås (2002) [90].

Tangerås reprend le modèle d'Auriol & Laffont (1992) [56] dans une configuration de sélection adverse, doublée d'aléa moral. Il introduit un paramètre de qualité discret $s_i \in \{0; \overline{s}\}$ dans la fonction de coût de l'entreprise i qui devient : $C_i = (\beta^i + s_i - e_i)q_i$. Le régulateur et les consommateurs perçoivent un signal imparfaitement informatif sur le niveau de qualité mis en œuvre.

L'utilité des consommateurs augmente avec la qualité perçue. De ce fait, la demande en faveur d'un produit dont la qualité perçue est $\overline{s}$ est supérieure à celle d'un produit classique ($s = 0$). Lorsque les entreprises investissent $\overline{s}$, leur profit est affecté par la hausse de coût correspondante et par l'accroissement de la demande. Il est néanmoins nécessaire qu'elles soient incitées par le régulateur à investir ; celui-ci doit alors augmenter le transfert versé lorsqu'il perçoit $\overline{s}$.

Dans le cas de la régulation des contrats individuels, Tangerås obtient des résultats cohérents avec les modèles classiques. Notamment, le contrat optimal est distordu en effort, mais aussi en qualité, par rapport à l'optimum en information symétrique. Il existe donc un seuil de productivité, $\beta^I \in [\underline{\beta}; \overline{\beta}]$, en-deçà duquel les entreprises n'investissent pas en qualité. En revanche, les entreprises efficientes investissent afin d'améliorer leur qualité, ce qui leur permet d'accroître leur rente in-

formationnelle. Par ailleurs, ces entreprises fournissent davantage d'effort que si elles n'augmentaient pas leur qualité. Le surplus des consommateurs et la rente des entreprises augmentent donc avec le niveau de qualité. Mais la hausse du coût social de production qui en résulte rend incertain l'effet sur le surplus global.

Lorsque le régulateur applique la concurrence par comparaison, Tangerås retrouve les résultats classiques d'Auriol et Laffont. Avec ce mécanisme, la variation du surplus global, consécutive à la hausse de la qualité, est supérieure à celle des contrats individuels. Cela signifie qu'il est socialement moins coûteux d'investir en qualité sous le mécanisme comparatif proposé, plutôt que sous un régime de contractualisation individuelle. La concurrence par comparaison a donc un effet positif sur les incitations à l'investissement, lorsqu'il s'agit d'améliorer la qualité observable de la production [21].

La concurrence par comparaison induit donc une variation de la répartition des investissements par rapport au cas de la régulation individuelle. D'une manière générale, les investissements favorisés améliorent particulièrement la production de l'entreprise (sa productivité ou sa qualité). Inversement, les investissements qui dégagent d'importantes externalités ne sont plus mis en œuvre par les entreprises. Avant d'introduire un régime de concurrence par comparaison, le régulateur doit donc évaluer la compatibilité de cette réorientation théorique des investissements avec l'activité régulée.

Conclusion : des limites théoriques à anticiper

Ce chapitre nous a permis d'évaluer quelques limites des mécanismes comparatifs, en considérant d'autres dimensions que celle des contraintes informationnelles. Nous avons relevé que certains phénomènes classiques de la régulation sont susceptibles d'advenir lorsque la concurrence par comparaison est appliquée au marché. Ils en limitent alors les résultats positifs. Ainsi, la concurrence par comparaison, qui s'était révélée très efficiente pour réduire l'asymétrie d'information, paraît vulnérable face aux comportements d'entente explicite, à la dynamique ou aux incitations à investir.

Notons que pour ces différentes limites théoriques, les effets n'agissent pas tous dans le même sens. Certains sont renforcés par la corrélation des entreprises, d'autres, au contraire, s'atténuent lorsque les agents comparés sont proches. Nous pensons qu'il n'y aurait pas de sens à conclure généralement et définitivement sur la pertinence des mécanismes comparatifs. Nous avons montré dans ce chapitre que ce système de régulation peut être pertinent en fonction des caractéristiques des marchés, et des conditions de sa mise en œuvre.

[21] Précisons que Tangerås analyse un autre cas de concurrence par comparaison, lorsque les agents sont, par ailleurs, en situation de concurrence. Cela correspond au cas des patients qui peuvent choisir entre plusieurs hôpitaux en milieu urbain, ou aux étudiants qui ont le choix entre plusieurs universités. Dans ce cas, l'optimum économique conduit à privilégier la qualité lorsque les agents sont productifs. La meilleure qualité offerte par un établissement efficient permet d'attirer davantage de consommateurs, qui bénéficient de ses services à un coût socialement inférieur.

Aussi, la meilleure façon de tirer parti de la concurrence par comparaison consiste à en maîtriser les principes, les effets et les limites. Si l'on ajoute à cela une connaissance solide des marchés susceptibles d'adopter cette forme de régulation, la probabilité d'une rencontre heureuse débouchant sur une proposition d'application sérieuse n'est pas si faible. C'est ce que nous présentons dans le chapitre suivant.

Conclusion : Des progrès théoriques favorables à la mise en œuvre

Cette deuxième partie de notre travail, consacrée à l'analyse des modèles théoriques de concurrence par comparaison, nous amène à conclure les résultats suivants :

- Les mécanismes comparatifs permettent de réduire l'asymétrie d'information, et donc, *a priori*, de rendre la régulation plus efficace.
- Suivant les conditions de leur mise en œuvre sur les marchés [1], ces mécanismes sont susceptibles d'être limités par des contraintes non informationnelles.

Nous récapitulons dans le tableau C2.1 (page suivante) la trentaine de modèles analysés [2], dans leur ordre d'apparition dans ce travail. Pour chaque modèle, identifié par ses auteurs, nous précisons :

- l'année de publication ;
- le cadre informationnel ;
- le nombre d'entreprises impliquées ;
- un éventuel point paticulièrement étudié ;
- le modèle de référence.

Ces modèles ont été développés sur une vingtaine d'années, sous l'impulsion de la nouvelle économie de la réglementation. Dans un souci répété de fournir des bases normatives aux mécanismes comparatifs existants, l'analyse théorique a proposé des modèles de plus en plus opérationnels [3]. Les principaux modèles développés sont les suivants :

- **Le modèle d'Holmström** est relativement opérationnel, mais dans un cadre d'aléa moral pur avec aversion au risque des agents. Or ce contexte informationnel, adapté à la théorie des incitations, n'est sans doute pas le plus pertinent pour modéliser la régulation.
- **Le modèle de Shleifer** a été inspiré par l'expérience Medicare ; il pèche néanmoins par ses hypothèses trop fortes.
- **Le modèle de Crémer et McLean - Pouyet** manque également de pertinence en raison du recours exagéré à l'hypothèse de neutralité au risque.

[1] Nous les avons précisées au chapitre 4.

[2] Signalons que d'autres auteurs ont analysé la concurrence par comparaison, notamment en l'appliquant à la relation électorale entre citoyens et hommes politiques.

[3] Rappelons qu'un modèle est opérationnel lorsqu'il remplit les trois conditions de cohérence, pertinence et mesurabilité.

auteurs des modèles	année	cadre	nombre d'entreprises	point particulier	modèle de référence
Baiman & Demski	1980	A	2	-	-
Holmström	1982	A	n	-	Holmström
Mookherjee	1984	A	2	-	Holmström
Lazear & Rosen	1981	A	2→n	tournois de rang	-
Nalebuff & Stiglitz	1983	A, S*	2→n	tournois de rang	-
Green & Stokey	1983	A, S*	n	tournois de rang	-
Shleifer	1985	A	n	-	Shleifer
Bivand & Szymanski	1997	A	2	dépendance spatiale	-
Crémer & McLean	1988	S	n	enchères	-
McAffe & Reny	1992	S	n	enchères	Crémer & McLean
Pouyet	2002	S	2	-	Crémer & McLean
Auriol & Laffont	1992	S	2	-	Auriol & Laffont
Auriol	1993	S	2	-	Auriol & Laffont
Demski & Sappington	1984	A, S	2	dissymétrie	-
Laffont & Tirole	1993	A, S	2	-	-
Chone & Lesur	(2001)	A, S	n	hétérogénéité	Auriol & Laffont
Auriol	2000	A, S	n	-	Auriol & Laffont
Boyer & Laffont	2003	A, S	n	-	-
Laffont & Martimort	2000	S	2	collusion expl.	Crémer & McLean
Pouyet	2002	S	2	collusion expl.	Crémer & McLean
Tangerås	2002	A, S	2	collusion expl.	Auriol & Laffont
Conteras & Rickman	(2004)	S	2	capture	Auriol & Laffont
Potters, Rockenbach, Sadrieh *et al.*	2004	-	2	collusion tacite	Shleifer
Meyer & Vickers	1997	A, S*	2→n	effet de cliquet	-
Faure-Grimaud & Reiche	(2003)	S	2	effet de cliquet	Crémer & McLean
Dalen	1998	A, S	2	investissement prod.	Auriol & Laffont
Powell & Szimanski	1997	S	2	investissement prod.	Shleifer
Sobel	1999	S	n	investissement prod.	-
Tangerås	(2002)	A, S	2	investissement qual.	Auriol & Laffont

TAB. C2.1 – Récapitulatif des modèles analysés ;
A : aléa moral ; S : sélection adverse ; S* : sélection adverse *ex ante*.

- **Le modèle d'Auriol et Laffont** est devenu *le* modèle de référence en sélection adverse. Il a été initialement inspiré par l'apparition de duopoles dans les industries de réseau libéralisées. Afin d'évaluer son opérationalité, il est plus pertinent de retenir le modèle d'Auriol (2000) qui intègre, en plus, un effort inobservable.
- **Le modèle de Boyer et Laffont** est sans doute le modèle le plus opérationnel, compte tenu de sa simplicité. Les auteurs reconnaissent d'ailleurs s'être inspirés des mécanismes comparatifs instaurés dans les secteurs des télécommunications et de l'éléctricité.

Cette interaction entre la théorie et la pratique a donc été particulièrement féconde. Si elle a permis à la théorie de s'enrichir de modèles toujours plus pertinents, elle a également contribué à la mise en œuvre des mécanismes comparatifs. Dans la troisième partie, nous abordons donc la question pratique de l'application de ces mécanismes, plus particulièrement sur les marchés ferroviaires.

Troisième partie

La concurrence par comparaison appliquée aux chemins de fer : justification et résultats

Introduction : Les mises en œuvre de mécanismes comparatifs

> « Les citoyens témoins des dysfonctionnements de l'appareil d'État, ne fonctionneront plus dans la société du xxie siècle comme leurs ancêtres élevés dans l'esprit du service public. Ne pas le reconnaître et ne pas adapter nos institutions à cette réalité priverait la France d'une partie importante des bénéfices de la globalisation des marchés et des progrès technologiques. »
>
> Jean-Jacques Laffont (1999),
> « Étapes vers un État moderne : une analyse économique »,
> *rapport au Conseil d'Analyse Économique.*

Depuis la mise en place de Medicare et la diffusion des modèles économiques, les applications de mécanismes comparatifs se sont multipliées. Elles ont la double caractéristique de recouvrir des secteurs d'activité variés, et de faire intervenir les comparaisons de diverses façons. On peut en effet distinguer une large variété d'usages des comparaisons, qui offrent des incitations croissantes :

- **Le « *benchmarking* assisté » :** Il s'agit d'une forme très peu contraignante du recours aux comparaisons. En général, les entreprises concernées participent avec le régulateur à la démarche. À l'issue du travail de comparaison, les résultats indiquent les faiblesses des entreprises, et celles-ci définissent un plan d'amélioration, avec l'assistance du régulateur.
- **La *sunshine regulation* :** Littéralement, il s'agit de la « régulation sous les projecteurs ». Dans ce cas, les comparaisons n'exercent qu'une pression concurrentielle indirecte. La diffusion des résultats entraine un effet de réputation, en ce sens qu'aucune entreprise ne peut se permettre d'apparaître trop longtemps inefficiente.
- **L'aide à la régulation :** Les comparaisons peuvent enfin intervenir directement dans la régulation, sans pour autant en constituer le socle. Le régulateur recourt alors ponctuellement aux comparaisons, afin de réduire l'asymétrie d'information. Il peut s'agir de résoudre un point technique particulier, ou bien de déterminer les gains de productivité exigibles, dans le cadre d'un mécanisme *price cap*. Taylor & Weisman (1996) [119] proposent notamment une extension des mécanismes de *price cap* classiques, afin de prendre en compte, notamment, la performance relative estimée par comparaisons. Bouf & Péguy (2001) [93]

expliquent comment les mécanismes comparatifs peuvent compléter utilement la concurrence pour le marché.

– **La régulation des coûts :** Les comparaisons servent alors directement à fixer les coûts, selon un mécanisme prédéterminé. Cette forme très incitative d'application se rapproche des modèles normatifs que nous avons étudiés dans la partie précédente. C'est sans doute le seul type de mécanisme comparatif que l'on puisse réellement dénommer « concurrence par comparaison ».

Le tableau ci-dessous détaille les principales mises en œuvre des mécanismes comparatifs :

secteur d'activité	pays	usage	références
hôpitaux	multiples	divers	voir note 2
bus	Norvège	régulation	Dalen & Gómez-Lobo (2003) [100]
eau	UK	aide	Cowan (1997) [99]
électricité	Amérique du Sud	aide	Estache, Rossi & Ruzzier (2004) [102]
armement	US	régulation	Riordan & Sappington (1989) [117]
infrastructures	Australie	*benchmarking*	Lawrence, Houghton & George (1997) [107]
tours Petronas	Malaisie	*sunshine*	-
infrastructure ferroviaire	UK	*benchmarking*	Kennedy & Smith (2004) [104]
chemins de fer	Japon	régulation	voir section 5.1

Tab. I3.1 – Principaux mécanismes comparatifs.

Les mécanismes comparatifs sont davantage appliqués à la régulation des industries de réseau. Mais c'est afin de réguler la tarification hospitalière qu'ils ont été le plus développés. Ces mécanismes, inspirés de l'exemple américain Medicare, prévoient de rembourser les hôpitaux en fonction des pathologies diagnostiquées et traitées. Outre aux États-Unis, ces systèmes de « tarification à la pathologie », en cours d'introduction en France, fonctionnent déjà au Québec, en Australie et dans de nombreux pays européens [1]. Selon les États, ces mécanismes sont plus ou moins incitatifs [2].

Dans les industries de réseau, les mécanismes comparatifs sont mis en œuvre de façon plus ou moins incitative. Dalen & Gómez-Lobo (2003) [100] décrivent comment la concurrence par comparaison est appliquée à plusieurs services de bus en Norvège. Ils montrent que les entreprises ainsi régulées réduisent leur inefficience plus rapidement que les autres. Dans le secteur de la distribution et de l'assainissement d'eau,

[1] Royaume-Uni, Allemagne, Suisse, Scandinavie.

[2] La bibliographie française (donc non exhaustive) relative à ce sujet est déjà importante. Elle présente l'intérêt de mêler approche théorique et mise en pratique. On peut notamment citer les travaux de Sylvain Pichetti, David Bardey, Brigitte Dormont et Romain Lesur, ainsi que la thèse d'Élisabeth Sage (1999) à Paris 9. Pour un exposé des mises en œuvre, voir le Hors série de juillet 2002 des *Dossiers solidarité et santé* : « La tarification à la pathologie : les leçons de l'expérience étrangère », La documentation française.

en Grande-Bretagne, les comparaisons permettent au régulateur de fixer les gains de productivité du mécanisme de *price cap* [3]. La même idée est développée dans un contexte multi-principal pour la régulation de l'industrie électrique en Amérique Latine par Estache, Rossi & Ruzzier (2004) [102].

Les mécanismes comparatifs sont également appliqués à la régulation des infrastructures. En Australie, le *Bureau of Industry Economics* a mené une analyse de *benchmarking* international, concernant les gestionnaires d'infrastructure des industries de réseau. Lawrence, Houghton & George (1997) [107] décrivent le travail conjoint du régulateur et des entreprises, ainsi que ses résultats. Depuis, la *Productivity Commission* australienne a ponctuellement recours aux comparaisons, lorsqu'elle est confrontée à une trop forte asymétrie d'information. De façon plus incitative, l'État malaisien a confié la construction de chacune des tours jumelles Petronas à une entreprise différente. Cette *sunshine competition* a incité l'entreprise initialement en retard, à fournir suffisamment d'effort pour achever sa tour en même temps que l'autre.
Enfin, de nombreuses opportunités ont été signalées pour introduire des mécanismes comparatifs. Estache *et al.* (2002) [101] montrent notamment qu'il serait pertinent d'appliquer la concurrence par comparaison à la régulation des ports mexicains. Kennedy & Smith (2004) [104] présentent les réflexions menées en Grande-Bretagne, en vue de comparer entre elles les performances locales du gestionnaire d'infrastructure ferroviaire. Si les britanniques n'ont pas encore procédé à l'introduction d'un mécanisme comparatif, nous verrons dans cette partie que les chemins de fer japonais sont, eux, régulés par la concurrence par comparaison.

Ce bref exposé illustre la variété d'usage des mécanismes comparatifs concernant tant les activités régulées, que le mode d'intégration des comparaisons dans la régulation. Cela démontre la souplesse de ces mécanismes et l'intérêt que tout régulateur est susceptible de leur porter. Enfin et surtout, cela prouve la capacité des économistes à apporter des solutions adaptées à la régulation des industries de réseau.

Forts des résultats obtenus dans la partie précédente et de ces observations, nous étudions dans cette dernière partie l'application de la concurrence par comparaison aux chemins de fer. Notre objectif consiste à proposer un mécanisme opérationnel, susceptible d'être appliqué sur certains marchés ferroviaires français. Dans le chapitre 5, nous analysons d'abord la concurrence par comparaison telle qu'elle est appliquée aux chemins de fer japonais et telle qu'elle est susceptible d'être mise en œuvre en France. Dans le chapitre 6, nous précisons notre proposition d'application, que nous validons sur des données des années 90.

[3] Voir Cowan (1997) [99] et la thèse d'Élisabeth Sage.

Chapitre 5

La concurrence par comparaison appliquée aux chemins de fer

Dans ce premier chapitre consacré aux applications ferroviaires de la concurrence par comparaison, nous analysons comment les modèles théoriques peuvent être mis en pratique. Dans une première section, nous considérons le cas des chemins de fer japonais, soumis depuis près de vingt ans à la concurrence par comparaison. En dépit de l'hétérogénéité des entreprises, un mécanisme incitatif a pu être mis en œuvre, de façon très pragmatique.

Dans la section suivante, nous proposons les bases d'un mécanisme comparatif, afin de réguler les transports ferroviaires régionaux en France. Nous affinons progressivement cette proposition, en tenant compte des résultats successivement établis dans la partie précédente. Cela nous conduit à définir le cadre de régulation envisagé et les effets attendus de son application.

5.1 L'expérience japonaise

L'étude des marchés ferroviaires japonais présente un grand intérêt, compte tenu de leur diversité. En effet, certaines liaisons (urbaines ou interurbaines) attirent un trafic exceptionnel en raison de la très forte densité de population rencontrée. En revanche, il existe également des lignes ferroviaires rurales, dont l'exploitation est adaptée à leur particularité. Le secteur ferroviaire japonais est caractérisé par :

- une multiplicité des entreprises de transport ferroviaire (plus d'une centaine) ;
- l'intégration verticale de ces entreprises, qui gèrent elles-même leur infrastructure ;
- l'absence de concurrence sur les voies [1] ;
- l'absence de mise en concurrence pour le marché.

[1] Il existe cependant quelques lignes parallèles, exploitées par des opérateurs distincts, permettant ainsi une certaine concurrence sur le marché.

Signalons également la prédominance des trafics voyageurs sur les trafics marchandises[2], ainsi que les importantes pertes de parts de marché du mode ferroviaire, à l'image de ce qu'ont connu les chemins de fer européens[3] (voir Terada (2001) [120] pour une présentation du secteur ferroviaire japonais).

Afin de caractériser les opérateurs ferroviaires japonais, nous reprenons (voir Shoji (2001) [118]) la classification établie par le Ministère des transports japonais qui distingue :

- **Les 6 JRs (Japan Railways) :** issues de la dé-intégration horizontale de l'ancien monopole public JNR en 1987, les JRs ont chacune un statut d'entreprise privée. Toutefois, seules les trois plus importantes ont ouvert leur capital au secteur privé.
- **12 opérateurs urbains :** il s'agit des compagnies municipales chargées d'exploiter les réseaux métropolitains.
- **TRTA :** la RATP tokyoïte.
- **Les 15 majors :** ce sont les plus grosses des compagnies privées ; elles exploitent essentiellement des services de banlieue dans les grandes agglomérations.
- **Les 6 quasi-majors :** elles ont les mêmes caractéristiques que les précédentes, mais opèrent sur des réseaux de plus petite taille.
- **Les minors :** ce sont les autres compagnies de transport de voyageurs.
- **JR Freight :** il s'agit de l'opérateur fret de référence puisqu'il achemine 99 % du trafic ferroviaire de marchandises sur les réseaux des 6 JRs.
- **15 opérateurs de fret privés :** marginaux, ils ne transportent que 1 % du trafic marchandises.

Si les opérateurs de fret sont soumis à une forte concurrence intermodale, les compagnies ferroviaires de services aux voyageurs disposent, elles, d'un fort pouvoir de marché qui rend nécessaire leur régulation. Depuis la fin des années 1980, le Ministère des transports régule les tarifs des opérateurs ferroviaires par la concurrence par comparaison. Initialement conçu afin de réguler les tarifs des 15 majors et de TRTA, le mécanisme a été étendu aux six JRs, ainsi qu'à neuf opérateurs urbains. Ainsi, il concerne à présent 31 opérateurs, représentant 95 % des services ferroviaires en terme de trafic. Trois objectifs ont été assignés au mécanisme comparatif :

1. Améliorer l'efficience des opérateurs : il s'agit évidemment de tirer parti des incitations fournies par tout mécanisme comparatif. Nous montrons dans les paragraphes qui suivent, comment les incitations interviennent dans le système japonais.

[2] 389 milliards de voyageurs.kilomètres pour 23 milliards de tonnes.kilomètres au Japon en 1998, contre 65 milliards de voyageurs.kilomètres pour 55 milliards de tonnes.kilomètres en France, la même année.

[3] La part de marché voyageurs du mode ferroviaire, exprimée en voyageurs.kilomètres, est passée de 67 % en 1965 à 27 % en 1998 ; celle des marchandises, exprimée en tonnes.kilomètres, est passée de 31 % en 1965 à 4 % en 1998. En France, on passe, respectivement pour les mêmes années, de 76 % à 9 % pour les voyageurs (en incluant le trafic métro afin d'homogénéiser ces comparaisons) et de 53 % à 27 % pour les marchandises.

2. Réduire les coûts de la régulation : l'administration du Ministère des transports reconnaît le faible niveau des coûts de transaction d'un mécanisme comparatif. Par ailleurs, la mise en œuvre de cette forme de régulation par l'administration japonaise s'efforce au maximum de réduire les coûts de transaction.
3. Assurer la transparence des révisions tarifaires : afin que les opérateurs ferroviaires acceptent le mécanisme de régulation, le Ministère des transports a pris soin d'insister sur la transparence du système.

5.1.1 Technique d'application

La mise en œuvre de la concurrence par comparaison par le Ministère des transports japonais afin de réviser les tarifs des opérateurs est assez originale ; nous reprenons dans ce qui suit l'analyse détaillée d'Okabe (2004) [114]. Cette application de la concurrence par comparaison vise à fixer un niveau de tarif plafond en deçà duquel les opérateurs peuvent plus ou moins librement fixer leurs tarifs.

Les comparaisons interviennent dans le calcul d'une simulation des charges du compte de résultat (hors éléments exceptionnels), face à laquelle les tarifs sont ajustés, de façon à équilibrer raisonnablement le compte de résultat (qui prévoit la réalisation d'un profit) :

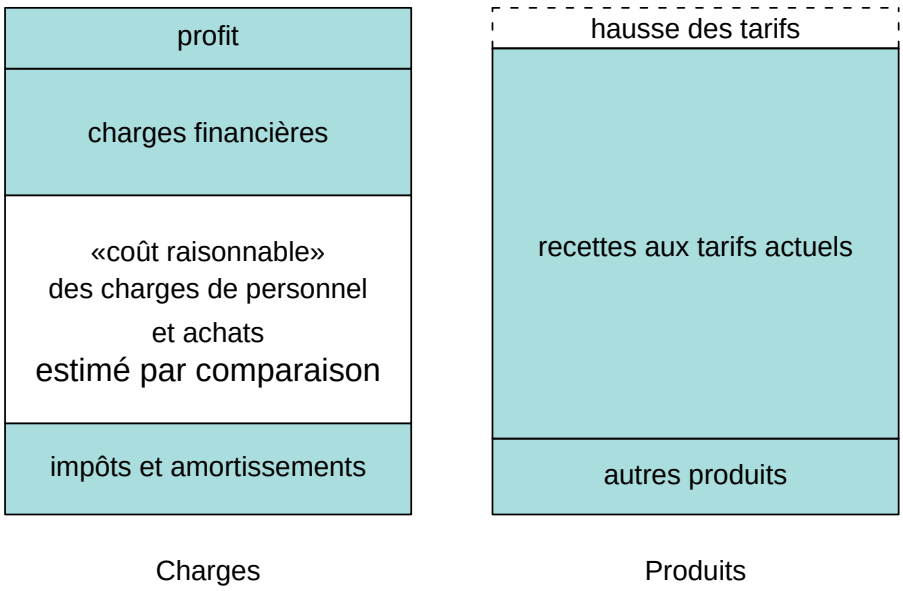

Fig. 5.1 – Compte de résultat reconstitué. Source : Okabe (2004) [114].

Afin de construire ce compte de résultat, le compte des charges comprend :
- une estimation des impôts et dotations aux amortissements ;
- une estimation des charges de personnel et achats ;
- une estimation des charges financières ;

– une estimation du profit.

C'est le montant estimé des charges de personnel et achats qui fait l'objet d'un traitement comparatif que nous explicitons.

Le compte des produits comprend :
– une estimation des produits du trafic aux tarifs actuels ;
– une estimation des autres produits ;
– une éventuelle hausse des recettes tarifaires, afin d'équilibrer le compte.

Le montant estimé de ce que devraient être les charges de personnel et achats d'un opérateur dépend des performances des autres opérateurs de sa catégorie (JRs, Majors ou urbains). Il est appelé « coût raisonnable », et est calculé de la façon suivante :

1. Les charges observées de personnel et achats sont affectées entre 5 postes du tableau 5.1 : voies, caténaire, matériel roulant, circulation des trains et gares.

Postes de charges	Équipements	Principales variables environnementales
voies	longueur de voie	voitures.km par km de voie
caténaire	longueur de caténaire	km d'EMU [4] par km de caténaire
matériel roulant	parc de matériel roulant	parcours moyen des voitures
circulation des trains	longueur de ligne	trains.kilomètres par km de ligne
gares	nombre de gares	nombre de voyageurs par gare

Tab. 5.1 – Variables mobilisées pour calculer le coût de référence.
Source : Okabe (2004) [114].

2. Les coûts unitaires sont calculés pour chaque opérateur, en divisant chacun des coûts précédents par l'équipement correspondant (voir tableau 5.1).
3. Pour chaque catégorie d'opérateurs, ces coûts unitaires sont régressés sur les variables environnementales, susceptibles de les affecter (voir tableau 5.1 pour les principales d'entre elles [5]). Le terme constant de la régression correspond au « coût unitaire de référence ».
4. Un « coût de référence » est calculé pour chaque opérateur. Il est obtenu en multipliant (vectoriellement) les « coûts unitaires de référence », estimés à l'étape précédente, par les équipements associés.

Le « coût de référence » calculé correspond donc au coût non biaisé auquel devrait produire l'opérateur. Un opérateur dont le coût observé est inférieur à ce coût de référence, est performant, et inversement. Le « coût raisonnable », sur lequel repose la hausse de tarif, dépend de l'écart entre le coût observé de l'opérateur, et le coût de référence estimé.

[4] Unités Multiples Électriques.

[5] Suivant la catégorie de l'opérateur, d'autres variables explicatives peuvent intervenir. Par exemple, les coûts unitaires de maintenance de la voie sont régressés par le pourcentage d'ouvrage d'art (pour les Majors) ou par la quantité de chute de neige (pour les JRs), en plus de la densité de trafic.

5.1.2 Principe du mécanisme incitatif

Le mécanisme incitatif mis en œuvre comporte, d'une part, des incitations statiques à la performance (suivant l'écart entre le coût observé et le coût de référence), et d'autre part des incitations dynamiques à la performance (suivant la variation de l'écart entre les deux coûts).

Incitations statiques

Les incitations statiques agissent de la façon suivante (voir figure 5.2 ci-dessous) :

- Si le coût observé est supérieur au coût de référence (opérateur peu performant), le coût raisonnable est égal au coût de référence, à charge pour l'opérateur de combler l'écart d'une manière ou d'une autre.
- Si le coût observé est inférieur au coût de référence (opérateur performant), le coût raisonnable est égal à la demi-somme du coût observé et du coût de référence. L'opérateur bénéficie donc, en plus de la compensation de ses coûts, d'un gain égal à la moitié de l'écart entre le coût de référence et le coût observé.

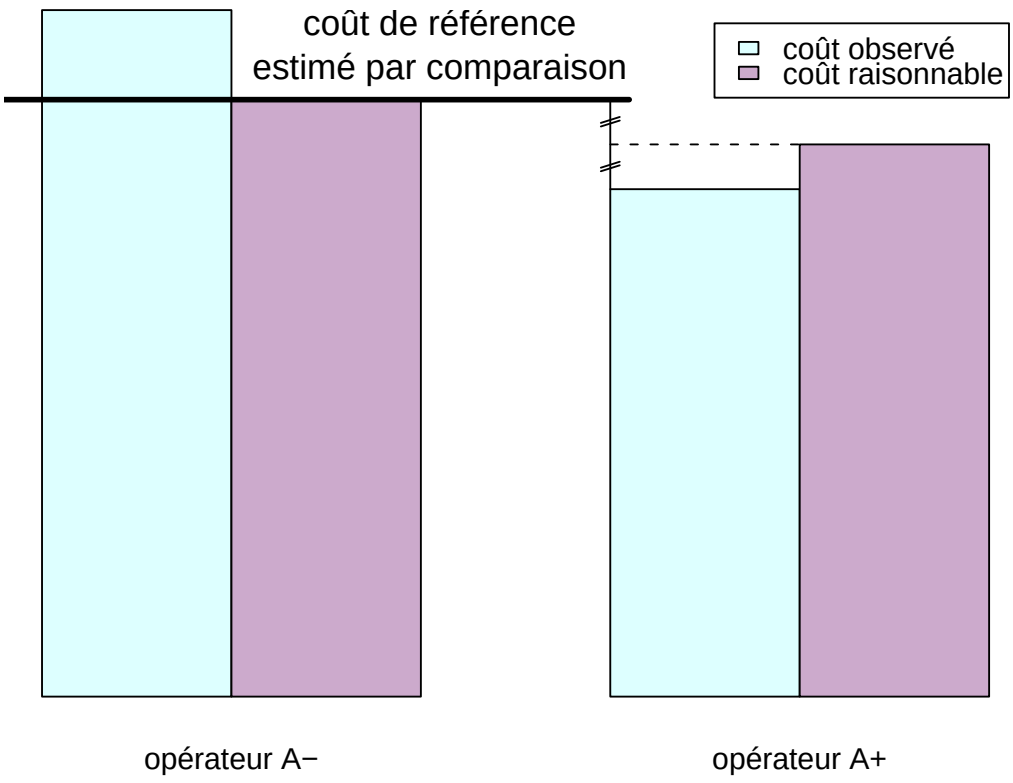

FIG. 5.2 – Mécanisme d'incitation statique pour un opérateur peu performant (A-) ou performant (A+).
Source : Okabe (2004) [114].

Ce mécanisme s'inspire donc du modèle de Shleifer corrigé de l'hétérogénéité dans la mesure où les prix sont fixés par comparaison à la moyenne des charges de personnel et achats des entreprises de la même catégorie. Il paraît néanmoins plus contraignant puisqu'un opérateur performant ne bénéficie que d'une moitié de l'écart entre le coût de référence sectoriel et son coût observé (le jeu n'est pas à somme nulle). Ce système incite donc les entreprises à l'effort afin de réduire leurs coûts de personnel et achats. Toutefois, ce mécanisme ne compare que les coûts marginaux, alors que Shleifer compare également les investissements. Compte tenu de la technique d'application retenue (l'équilibre du compte de résultat reconstitué), le système devrait inciter les entreprises à surinvestir pour améliorer leur efficience. En effet, d'une part l'amortissement des investissements est garanti, d'autre part, une entreprise efficiente (dont les coûts marginaux sont inférieurs à ceux de ses concurrents) bénéficie de la moitié de l'écart entre le coût de référence sectoriel et son coût observé. Il est cependant très difficile de repérer ces surinvestissements potentiels.

Incitations dynamiques

Les incitations dynamiques visent à stimuler un opérateur peu performant, mais dont les coûts se réduisent, et à mettre en garde un opérateur performant, mais dont les coûts augmentent. Le coût raisonnable est augmenté (ou diminué), en pourcentage, de l'opposé de la demi-différence entre les années N et N-1 du rapport coût observé / coût de référence.
Ainsi, considérons un opérateur efficient, dont le rapport coût observé / coût de référence passe de 0,8 à l'année N-1, à 0,9 à l'année N. Son coût raisonnable sera modifié (à la baisse) de -5 % ($= -1/2 \times (0,9 - 0,8)$).
Inversement, considérons un opérateur inefficient, dont le rapport coût observé / coût de référence passe de 1,2 à l'année N-1, à 1,1 à l'année N. Son coût raisonnable sera modifié (à la hausse) de 5 % ($= -1/2 \times (1,1 - 1,2)$).

Enfin, le mécanisme comporte également des incitations non-financières, engendrées par la *sunshine regulation*. En effet, le Ministère diffuse les résultats des comparaisons qui sont largement médiatisés. De leur côté, les entreprises efficientes ne manquent pas de mettre en avant leur bonne performance.

Mizutani (1997) [111] a analysé empiriquement les effets de la concurrence par comparaison à ses débuts. Bien que ses résultats ne soient pas toujours très significatifs, Mizutani met en évidence deux principaux résultats lorsque la concurrence par comparaison est appliquée : la réduction des coûts d'exploitation, et celle de leur variance. Ce double constat va de pair avec les effets attendus du mécanisme mis en œuvre. La réduction des coûts est le principal effet attendu de la mise en concurrence par comparaison. La réduction de la variance des coûts d'exploitation est une conséquence logique de la concurrence qui tend à faire converger les prix vers l'équilibre concurrentiel (ou un équilibre collusif !). Dans le contexte japonais, cette réduction de la variance peut aussi être expliquée par la dimension dynamique du mécanisme incitatif. Mizutani montre aussi que l'amélioration de la performance

augmente progressivement, les meilleurs opérateurs ayant commencé à réduire leurs coûts plus tard.

5.1.3 Une mise en œuvre pragmatique

La mise en œuvre de ce mécanisme régulateur est marquée par le pragmatisme d'une administration, soucieuse de réduire les coûts de transaction et de faire accepter la concurrence par comparaison aux opérateurs. Ainsi, le mécanisme a été singulièrement enrichi en 1995-96 afin d'être appliqué aux JRs et aux opérateurs urbains. La décomposition des charges de personnel et achats en 5 postes a alors permis une meilleure prise en compte des biais environnementaux. Par ailleurs, afin de réduire l'instabilité liée aux révisions annuelles de tarifs, l'administration fixe désormais les tarifs pour une durée de trois ans. Cette disposition, demandée par les entreprises, permet de réduire encore les coûts de transaction. Suivant la même logique, la transparence du mécanisme (qui n'utilise que des données déjà connues et publiées dans un autre cadre) réduit l'incertitude des opérateurs et le coût de la régulation. Enfin, l'application de la concurrence par comparaison a permis de réduire le coût de la régulation économique du secteur ferroviaire. Auparavant partagée entre plusieurs autorités locales, la régulation est maintenant exercée par une agence commune des régulateurs locaux.

Vue de France, la régulation économique des chemins de fer japonais impressionne. D'abord, parce qu'elle a le mérite d'exister. Ensuite, parce que les préconisations théoriques y ont été intelligemment et pragmatiquement déclinées. Cela a permis d'inciter correctement les opérateurs à réduire leurs coûts, tout en prenant soin de minimiser les coûts de la régulation. L'expérience japonaise ne peut que renforcer l'idée d'introduire la concurrence par comparaison sur les marchés ferroviaires français.

5.2 Une proposition d'application aux TER français

Malgré les différences entre les marchés, la régulation économique du transport ferroviaire de voyageurs au Japon peut être riche d'enseignements pour les chemins de fer français. Le trait commun principal entre les deux pays est le monopole légal accordé aux opérateurs ferroviaires. Les étendues verticale et horizontale de ces monopoles légaux ne sont toutefois pas identiques[6]. Néanmoins, il convient de retenir de l'expérience japonaise qu'un mécanisme de concurrence par comparaison peut être mis en œuvre à partir du moment où les performances de plusieurs opérateurs ferroviaires semblables (appartenant à la même catégorie) peuvent valablement être comparées. Dès lors, c'est sur la base de ce critère de comparabilité que nous proposons de passer

[6] Monopole sur les activités voyageurs (transport longue distance *et* régional) et certaines missions de gestion de l'infrastructure en France. Monopoles intégrés (transport de voyageurs longue distance *ou* régional et gestion de l'infrastructure associée) au Japon. Monopole national en France, *versus* monopoles locaux au Japon.

en revue les différents marchés ferroviaires français, susceptibles d'être régulés par un mécanisme comparatif.

5.2.1 Quelques options pour une application en France

Comparaisons internationales

La première idée qui vient à l'esprit consiste à comparer dans la lignée de Bouf & Péguy (2001) [93] les performances des opérateurs ferroviaires historiques européens, dont les différents trafics présentent quelques similitudes. Toutefois, les réformes institutionnelles et réglementaires mises en œuvre par les différents États européens ne permettent plus d'assurer l'homogénéité des comparaisons. Il y a d'abord une grosse difficulté à collecter des données fiables. Ensuite, les activités des opérateurs ferroviaires sont devenues hétérogènes, ce qui rend très délicat les comparaisons de performances. Ainsi, certains opérateurs ferroviaires assument encore de nombreuses missions de service public (cas de la France), tandis que d'autres n'assurent plus que les trafics rentables (cas de la Suède). Voir à ce sujet Nash & Shires (2000) [113]. D'une manière générale, la menace des effets de dépendance spatiale [7] dissuade toute tentative de comparaison fine entre les opérateurs ferroviaires européens actuels.

Mais quitte à renoncer à un mécanisme comparatif précis, il peut être opportun de mettre en œuvre des pratiques comparatives moins incitatives, mais plus ciblées, suivant une logique de *benchmarking* assisté. Les principaux acteurs du système ferroviaire français auraient sans doute intérêt à considérer d'autres modèles d'organisation que ceux dont ils ont hérité. De telles démarches ont déjà été mises en œuvre par RFF, concernant le cadencement des trains ou les politiques de maintenance et régénération de l'infrastructure, avec la Suisse pour étalon. Ces initiatives devraient être généralisées en considérant l'autre étalon ferroviaire mondial qu'est le Japon.

Comparaisons intranationales

Il est toutefois possible de s'inspirer du modèle japonais de régulation intranationale que nous avons décrit dans la section précédente. Sur le modèle de la déintégration des JNR en six compagnies régionales JRs, on peut imaginer comparer les performances des trois réseaux Grandes Lignes (Atlantique, Nord et Sud-Est). Toutefois, les comparaisons de performances, entre seulement trois agents, ne seront pas suffisamment robustes pour être intégrées à un mécanisme de régulation.

En revanche, il semble se dessiner une perspective plus favorable du côté des missions de gestion de l'infrastructure confiées à l'opérateur historique. Dans Bouf & Lévêque (2005) [92], nous avons analysé dans quelle mesure la concurrence par comparaison pouvait utilement être mise à profit pour réguler les infrastructures de transport. Nous avons abouti aux trois résultats suivants :

[7] Rappelons que ceux-ci sont induits par l'asymétrie d'information accrue entre un régulateur national et un opérateur étranger. Cette asymétrie informationnelle porte essentiellement sur la réglementation, les financements et les choix comptables des différents opérateurs.

- Le recours à la concurrence par comparaison est limité concernant la construction des infrastructures. Cela est dû à l'importante hétérogénéité géologique et géographique qui impacte directement les coûts de construction (l'ampleur des travaux de génie civil à engager).
- L'usage des comparaisons, afin d'évaluer la performance dans l'allocation des capacités d'infrastructure, est envisageable. Toutefois, cela requiert une forme souple, dans une logique de « *benchmarking* assisté », qui ne peut être appliquée qu'à des réseaux très similaires.
- La concurrence par comparaison peut parfaitement être mise en œuvre pour inciter les opérateurs chargés de la maintenance d'infrastructures à réduire leurs coûts.

Ainsi, partant de ce dernier résultat, il est tout à fait concevable qu'un maître d'ouvrage comme RFF ait recours à la comparaison des performances des 150 établissements d'équipement (EVEN) de son gestionnaire délégué, afin d'inciter ce dernier à progresser. L'expérience japonaise montre bien, dans le cas d'opérateurs intégrés, que les charges de maintenance de l'infrastructure (voie et caténaire) peuvent être régulées grâce à la concurrence par comparaison.

C'est finalement sur le transport régional que nous choisissons de travailler. En effet, la configuration actuelle des marchés associés s'y prête tout à fait, et la régulation économique y est peu développée. Nous montrons que les différents modèles théoriques analysés dans la partie précédente peuvent y être appliqués sans difficulté. Nous étudions enfin le sens que l'on souhaite donner à la performance d'un opérateur de transport ferroviaire régional.

5.2.2 Perspective d'application régionale

À première vue, les modèles théoriques de concurrence par comparaison, qui font intervenir plusieurs agents, semblent fortement éloignés de la réalité des marchés régionaux [8] du transport ferroviaire, lesquels sont toujours monopolisés par la SNCF. De plus, les opérateurs ferroviaires régionaux ne bénéficient d'aucun statut particulier, n'étant pas même des filiales de l'établissement public ; ce ne sont que des « activités ». Toutefois, ce paradoxe peut être levé si on considère le contexte de libéralisation des marchés et, en particulier, la récente « régionalisation » de ces marchés, intervenue en 2002.

Historique de l'émergence des régulateurs régionaux en France :

Depuis 1984, les régions peuvent ajouter (en les subventionnant) des services ferroviaires, en contractant avec la SNCF. Ces contrats transféraient à la région non seulement le risque industriel sur le(s) train(s) créé(s) sur une liaison donnée, mais aussi le risque commercial sur la liaison. Ces contrats, peu incitatifs, furent

[8] Nous ne considérons pas ici les cas particuliers de l'Île-de-France et de la Corse dont la réglementation est différente.

suspendus en 1994, en raison de la facture financière présentée aux régions par la SNCF après la récession de 1993.

En 1997, une phase d'« expérimentation » commence : 6, puis 7 régions deviennent les régulateurs de leurs services de transport (les Transports Express Régionaux - TER). Elles en définissent les horaires, la tarification, et les subventionnent.

La loi Solidarité et Renouvellement Urbain de décembre 2000 régionalise les TER, en généralisant cette décentralisation. Les régions se voient alors attribuer les compétences et les finances nécessaires à la régulation de leurs TER.

Les vingt régions ont signé en 2002 des conventions de délégation de leurs services TER avec la SNCF, pour une durée comprise entre 5 et 10 ans.

Une régulation actuellement peu incitative

Depuis une vingtaine d'années qu'elles ont la possibilité d'intervenir sur leur marché ferroviaire, les régions administratives françaises ont progressivement acquis les compétences d'un régulateur classique. Il leur manque toutefois la possibilité de choisir un autre opérateur que la SNCF. En attendant que la concurrence pour le marché soit introduite, la concurrence par comparaison semble être un mécanisme régulateur justifié et adapté au cadre réglementaire actuel. En effet, les 20 régions ne disposent pas d'une expertise suffisante pour négocier avec l'opérateur historique qui leur est en outre imposé.

Les conventions signées sont généralement assez incitatives concernant le trafic : l'opérateur assume majoritairement le risque commercial. Elles intègrent également correctement la qualité de service avec des systèmes de type *bonus-malus*. En revanche, elles manquent de façon flagrante d'incitations à la maîtrise des coûts. Certes, la rémunération de l'opérateur est forfaitaire, mais on est bien loin du rigoureux modèle de *price-cap* indexé sur $RPI - X$ [9]. D'une part, l'index des prix est une agrégation d'index [10] qui augmentent plus vite que le taux d'inflation. D'autre part, aucun gain de productivité n'est exigé de l'opérateur, c'est-à-dire que $X = 0$ [11].

Quels opérateurs et quel régulateur ?

Le mécanisme que nous proposons consisterait donc à mettre en concurrence par comparaison les 20 activités régionales TER de la SNCF, qui seraient « les opérateurs ». Cette régulation pourrait être développée sous l'impulsion d'une agence commune des régulateurs régionaux. Ce « régulateur » pourrait regrouper les compétences régionales au sein du GART (le Groupement des Autorités Responsables de Transport)

[9] Où *RPI*, *Retail Price Index*, désigne le taux d'inflation, et X les gains de productivité.

[10] Elle est constituée, suivant les conventions, à partir d'indices du coût du travail dans différents secteurs, d'indices salariaux et de divers indices de prix, dont ceux de l'énergie.

[11] La convention signée par la région Centre prévoit toutefois une baisse annuelle de 1 % des charges de structure, lesquelles ne représentent que 18 % des charges forfaitisées...

ou de l'ARF (Association des Régions de France). Une telle agence commune permet de mutualiser les informations et les compétences de ses membres, tout en maintenant l'essentiel de la régulation au niveau régional. Le processus de comparaison devrait permettre d'inciter les opérateurs à la performance, et d'accroître la capacité d'expertise des régulateurs régionaux.

La question se pose de savoir sous quelle forme il faut inciter les opérateurs. Si les modèles de la littérature introduisent un mécanisme financier lié au résultat des comparaisons, il n'est pas évident qu'il s'agisse de l'option la plus pertinente concernant notre cas d'étude.

Quelles incitations ?

En théorie, l'espérance d'un gain et la menace d'une pénalité incitent financièrement chaque entreprise à dépasser ses concurrents virtuels. Mais dans notre cas, les opérateurs ne sont que des « activités » (ne disposant pas de l'autonomie budgétaire) au sein de la même entreprise. Dans le cas d'un jeu à somme nulle, au niveau global de la SNCF, les gains et pénalités du jeu vont se compenser mutuellement au sein du budget de l'entreprise. Aussi, quelles que soient les performances de chaque opérateur, le résultat financier du mécanisme est le même : un remboursement *ex post* de tous les coûts. D'une manière générale, la configuration du marché revient à considérer que les opérateurs forment un « cartel » légal. Il convient donc surtout de mettre en œuvre un mécanisme incitatif (financièrement ou non), qui soit robuste à la collusion.

L'expérience de la régionalisation prouve que les incitations financières peuvent être pertinentes dans ce contexte de contractualisation région-SNCF. Des mécanismes similaires (dans le cadre de la qualité de service) motivent, au sein de l'entreprise, les équipes de travail et leurs managers, qui bénéficient d'un effet de réputation. Ainsi, les résultats des comparaisons pourraient être utilisés en complément de l'actuel mécanisme de régulation afin d'offrir aux régulateurs une capacité d'expertise renforcée. Un tel usage des comparaisons devrait néanmoins contribuer au développement de comportements concurrentiels pour les raisons suivantes :

- D'abord, un effet de réputation intervient. Aucun opérateur ne peut se permettre d'avoir une image bien moins bonne que les autres. Cet effet apparaît dès que le régulateur diffuse largement les résultats des comparaisons.
- Ensuite, un effet de croyance peut survenir. Avec l'ouverture progressive des marchés, les régulateurs devraient s'intéresser de plus en plus aux résultats du mécanisme, en vue de la préparation de futurs appels d'offre. En montrant son attachement au mécanisme, et la crédibilité qu'il y accorde, le régulateur peut induire un effet concurrentiel.
- Enfin, dans le cadre actuel, les résultats des comparaisons permettront aux régulateurs de négocier de véritables gains de productivité (le X du *price-cap*) dans la logique d'aide à la régulation exposée en introduction de cette partie. Nous proposons comment procéder dans ce cas, de manière très concrète, dans la conclusion générale de ce travail.

Ainsi, les régulateurs régionaux auraient beaucoup à gagner à mettre en œuvre la concurrence par comparaison entre les 20 activités TER de la SNCF, chacun étant libre de fixer les incitations qu'il souhaite fournir à son opérateur en cohérence avec le résultat des comparaisons. Il s'agit donc d'une proposition peu coercitive en soi, dans la logique traditionnelle de la régulation économique des chemins de fer en France.

Un avenir prometteur

Lorsque les régions disposeront de capacités d'expertise accrues et d'éléments de régulation désagrégés, le mécanisme comparatif pourra être mis en œuvre à une échelle plus fine. On pourrait notamment imaginer comparer les sous-réseaux de lignes voisines que nous avons mis en évidence dans la première partie. Ce travail, plus exigeant en terme de données à mobiliser, permettra une régulation plus fine des marchés, en prenant mieux en compte l'impact des variables environnementales.

La concurrence par comparaison apparaît aussi comme un complément utile à la future concurrence pour le marché, comme le signalent Bouf & Péguy (2001) [93]. La concurrence par comparaison permettrait de maintenir une pression concurrentielle, en l'absence d'enchérisseurs face à l'opérateur historique. Réciproquement, les enchères de la concurrence pour le marché permettent de changer un opérateur dont les comparaisons révèlent l'insuffisante performance, sans interrompre le service. Ces deux mécanismes de régulation sont donc complémentaires et se renforcent en partie l'un l'autre. Voyons maintenant plus précisément dans quelle mesure les modèles microéconomiques de concurrence par comparaison sont susceptibles de s'appliquer (lorsqu'il s'agit de bénéfices), ou non (lorsqu'il s'agit de limites), à notre cas d'étude.

5.2.3 Applicabilité des modèles théoriques

Nous avons vu qu'il semble, *a priori*, pertinent d'appliquer la concurrence par comparaison aux marchés régionaux du transport ferroviaire de voyageurs. Il convient d'abord de vérifier plus précisément si, dans ce contexte particulier, l'apport informationnel des modèles théoriques permet d'améliorer la régulation. Nous examinons ensuite dans quelle mesure les limites théoriques, liées aux autres dimensions de la régulation, menacent notre proposition d'application.

Un apport informationnel salvateur

Les modèles théoriques que nous avons analysés dans la dimension des contraintes informationnelles mettent, à peu près tous, l'accent sur la nécessaire corrélation entre les agents comparés. Sur ce point, notre proposition de comparer les vingt activités TER de la SNCF relève du cas d'école, tant la corrélation y est élevée. Indépendamment de toute considération environnementale, même l'hypothèse d'identité des agents serait acceptable, si nécessaire.

Concrètement, l'enjeu auquel font face les régulateurs régionaux relève autant d'une problématique de sélection adverse que d'aléa moral. Le principe de *price cap* adopté se heurte à l'ignorance des régulateurs quant à la productivité de leur

opérateur. L'asymétrie d'information naît de la méconnaissance de la caractéristique cachée, β, des entreprises. En 2002, un audit des comptes régionaux a permis de fixer la rémunération initiale des opérateurs. Mais depuis, l'accroissement de la demande et les gains de productivité ont permis aux activités régionales de se constituer des rentes informationnelles. Le problème vient de ce que les régions n'ont actuellement pas les moyens de réduire ces rentes. Par ailleurs, nous montrons à la section 6.3 que le niveau d'effort mis en œuvre par les opérateurs varie selon les régions. Mais ce paramètre étant inobservable, les régulateurs ne parviennent pas à inciter les opérateurs à fournir un effort plus optimal.

Théoriquement, compte tenu de la réglementation interne de l'entreprise, la corrélation entre les activités régionales est parfaite ; on aurait $\alpha = 1$, en reprenant le paramètre d'Auriol & Laffont (1992) [56]. En pratique, les performances variables de l'encadrement conduisent à l'apparition de caractéristiques particulières ϵ. Dans le contexte de sélection adverse que nous avons développé à la section 3.4, où la fonction de coût s'écrit $C = [\alpha b + (1-\alpha)\epsilon]q$, les paramètres introduits correspondent aux réalités suivantes :

- **La composante commune de la productivité, b** reflète notamment les modes d'exploitation réglementaires des trains, et la base salariale identique entre les régions. Elle prédétermine largement la productivité des opérateurs régionaux.
- **La composante particulière de la productivité, ϵ** illustre l'efficience de la gestion locale, notamment l'optimisation des ressources.

Si l'on considère les divers aléas susceptibles d'influencer les performances réalisées par nos vingt activités TER, là encore les conditions d'application de la concurrence par comparaison semblent optimales. En effet, en référence au modèle d'Holmström (1982) [71], l'aléa qui perturbe le niveau de production x résultant d'un effort a s'écrit : $x = a + \eta + \epsilon$. Le rapport de variances entre la composante commune de l'aléa et la composante particulière est alors, dans notre cas, particulièrement élevé :

- **La composante commune de l'aléa, η** modélise les modifications du règlement du travail de la SNCF, les variations des prix ou encore les mouvements sociaux nationaux.
- **La composante particulière de l'aléa, ϵ** reflète les perturbations régionales de la production. Celles-ci peuvent être dues à la météo, aux grèves et incidents locaux. Rappelons que certaines techniques économétriques, les frontières stochastiques (que nous détaillons dans le dernier chapitre), permettent d'atténuer cet impact.

Cette réflexion nous amène à discuter l'acceptabilité de l'hypothèse, fréquente, de neutralité au risque. En soi, cette hypothèse est généralement admise lorsqu'elle concerne des industries de réseau réglementées. Dans la première partie, nous avons précisé (page 52) que le statut d'EPIC de la SNCF la préservait de tout risque, du temps où son monopole n'était pas remis en question. Si la libéralisation rend l'entreprise un peu plus averse au risque, la gestion hasardeuse de nombreux dossiers montre que la SNCF demeure toujours quasi-neutre au risque. Précisons cependant

que le concept de neutralité au risque ne peut être que relatif. Une entreprise peut être considérée comme neutre au risque, sans pour autant qu'elle accepte de participer à une loterie à la Crémer & McLean (1988) [66]. Aussi, dans le contexte que nous proposons, l'hypothèse de neutralité au risque de la SNCF ne paraît pas exagérée.

Voyons à présent dans quelle mesure notre proposition d'application de la concurrence par comparaison est sensible aux limites théoriques, présentées au chapitre précédent.

Des limites théoriques peu menaçantes

Nous précisons brièvement, dans les paragraphes qui suivent, pourquoi les limites théoriques repérées ne semblent pas menacer notre application de la concurrence par comparaison.

- **Collusion explicite :** elle constitue évidemment la menace principale à la mise en œuvre de la concurrence par comparaison. Dans la mesure où les agents comparés sont les activités d'une même entreprise, rien n'est plus naturel pour eux que de coordonner leurs actions afin de maximiser la rente de l'entreprise. Soulignons toutefois que :
  - le nombre important d'agents comparés (vingt) complexifie les possibilités d'entente ;
  - la comparaison de données comptables réduit les opportunités de modification de l'information ;
  - sur des marchés relativement stables dans le temps, toute dérive collusive peut être facilement détectée par le régulateur ;
  - la détection d'une entente aurait un effet très négatif sur la réputation de l'entreprise, à la veille de l'ouverture des marchés ;
  - la collusion nécessite de maîtriser l'ensemble des paramètres concourant à l'estimation de la performance, dont les variables environnementales exogènes.

  Compte tenu de ces éléments, la collusion paraît déjà moins menaçante. Toutefois, afin d'assurer les régulateurs régionaux contre tout risque de collusion, il semble souhaitable d'utiliser des données passées, que l'entreprise ne peut plus manipuler.
- **Capture du régulateur :** c'est peut-être la principale faiblesse de notre proposition. Celle-ci prévoit une agence commune des régulateurs régionaux, ce qui suppose donc la coopération de chacun d'eux. Or il ne faut pas exclure qu'une majorité politique régionale puisse être capturée par l'opérateur ferroviaire, ce qui réduirait l'efficience de notre proposition.
- **Collusion tacite :** compte tenu de la facilité avec laquelle une entente explicite peut voir le jour, il est peu vraisemblable que la collusion tacite limite notre proposition.
- **Effet de cliquet :** un effet de cliquet est tout-à-fait susceptible d'advenir. Les différentes activités régionales de la SNCF peuvent convenir de réduire leurs efforts, afin de ne pas en être expropriées lors des comparaisons ultérieures.

Mais, la prochaine ouverture des marchés et le mécanisme de *price cap* des régions inciteront l'entreprise à ne pas trop retenir ses efforts.

- **Investissements productifs :** deux éléments contribuent à atténuer d'éventuelles craintes liées à un désinvestissement productif. D'abord, la SNCF s'est déchargée de la plupart des investissements, soit sur les régions (qui financent le matériel roulant), soit sur RFF et les collectivités locales (qui financent l'infrastructure). Par ailleurs, les investissements dans le transport régional de voyageurs ne présentent pas d'externalité qui ne soit pas internalisée.
- **Qualité :** la qualité est un paramètre multidimensionnel et donc difficilement appréhendable. S'il n'est pas dans l'intérêt des opérateurs de réduire la sécurité des trains, d'autres volets de la qualité - insuffisamment régulés - pourraient être pénalisé par une régulation plus contraignante sur les seuls coûts. Ainsi, plus encore qu'actuellement, les opérateurs pourraient réduire l'offre en termes de places disponibles par train en deçà du minimum de confort admissible.

Ainsi, le mécanisme dont nous proposons la mise en œuvre semble peu menacé par les différentes limites théoriques de la concurrence par comparaison. Le contexte réglementaire et historique dans lequel il s'inscrit le préserve de nombreux effets négatifs. Par ailleurs, une utilisation judicieuse de l'information disponible permet au régulateur de se prémunir contre la menace de la collusion. Aussi, dès lors que l'ensemble des régulateurs régionaux concourent à mettre en œuvre ce mécanisme, les incitations théoriques devraient être effectives. Cette analyse conforte donc notre proposition, que nous poursuivons plus avant en nous interrogeant sur la mesure de la performance des opérateurs qu'il convient de retenir.

5.2.4 Mesure de la performance des opérateurs ferroviaires

Avant de pouvoir procéder à l'estimation de la performance des opérateurs, il convient de définir précisément celle-ci. Que souhaite-t-on mesurer ? Comment peut-on le mesurer ? Nous élaborons une réponse à la première de ces questions à travers l'analyse théorique de l'efficience, tandis que la seconde question nous conduit à affiner notre choix de mesure de la performance, en prévenant les effets de dépendance spatiale.

Productivité ou efficience ?

La comparaison des performances des TER se doit de refléter les performances des différentes activités régionales de la SNCF, corrigées de l'hétérogénéité externe susceptible de les biaiser. Cela est d'autant plus nécessaire que le modèle vise à inciter localement la SNCF à améliorer son efficience. Le modèle doit donc être le plus robuste possible, afin de crédibiliser le mécanisme de concurrence par comparaison.

Nous choisissons de ne pas évaluer la performance des différentes activités TER par une estimation de leur productivité respective. Telle que la définissent Lovell (1993) [110] et Waters (2000) [121], il s'agit de tout ratio du type *output(s)/input(s)*. La mesure de la productivité peut être partielle (lorsqu'un output est mis en regard d'un *input*) ou totale, lorsque le ratio fait intervenir des agrégations d'*outputs*

et d'*inputs*. Oum, Waters & Yu (1999) [115] ainsi que Nash & Shires (2000) [113] donnent de nombreux exemples de telles mesures de la productivité dans les chemins de fer. Toutefois, les mesures de productivité ont l'inconvénient d'être affectées par divers facteurs : l'efficience, bien sûr, mais aussi les nombreuses variables environnementales qui influent sur la productivité. Mesurer la performance avec des ratios de productivité est source de biais importants, que la rigidité de l'approche ne permet pas de corriger de façon satisfaisante. C'est pourquoi nous nous tournons vers une estimation de la performance, approchée en tant qu'efficience.

Selon Lovell (1993) [110] et Waters (2000) [121], l'efficience est le résultat de la comparaison, entre les valeurs observées et optimales, des *outputs*, coûts, revenus ou profits. Usuellement, l'efficience est estimée par des techniques de frontière qui estiment d'abord l'efficience optimale (la frontière) qu'il est possible d'atteindre, puis les écarts entre les résultats des agents, et cet optimum. Avant de détailler au chapitre suivant la méthodologie suivie pour élaborer cette frontière, interrogeons-nous d'abord sur le type d'efficience que nous souhaitons évaluer.

Le choix de l'efficience coût

L'efficience peut être approchée suivant différents objectifs, exprimés en terme de production, d'allocation, d'équité, d'équilibre budgétaire ou de politique macroéconomique. Perelman & Pestiau (1994) [116] précisent qu'en dépit de l'incompatibilité et de la difficile mesurabilité de ces objectifs, « le seul objectif dont l'atteinte n'entrave pas celle des autres est l'efficience productive. Produire trop peu ou employer trop de facteurs de production par rapport à ce qui est techniquement faisable ne peut être justifié par un autre objectif quelconque ».

Nous avons précisé dans la section précédente que l'actuel mécanisme régulateur pâtit de l'absence d'incitation à la réduction des coûts. Afin de renforcer l'actuelle régulation, il apparait donc souhaitable d'évaluer l'inefficience sous la forme d'une surconsommation de facteurs de production, au delà du nécessaire à la production d'une certaine quantité de service. Nous faisons donc le choix restrictif de considérer la performance sous l'angle productif. Depuis Farell (1957) [103], le cadre analytique d'étude de l'efficience a été largement développé. Kumbhakar & Lovell (2000) [106], proposent quatre types d'efficience que nous rappelons brièvement :

- **L'efficience technique :** elle renseigne sur l'aptitude de l'entreprise à optimiser sa production en terme de quantité de facteurs de production consommés. On la désigne également par le terme d'efficience productive. Sur la figure 5.3 qui représente l'isoquante à deux facteurs de production (x_1 et x_2) d'une entreprise i, l'efficience technique correspond au rapport OT/Oi.
- **L'efficience coût :** elle rend compte de l'efficacité avec laquelle l'entreprise minimise ses coûts pour produire une certaine quantité de service. Elle peut être décomposée multiplicativement par le terme d'efficience technique défini ci-dessus et par un terme d'efficience allocative. Celui-ci correspond à l'utilisation dans des proportions rationnelles des différents facteurs de production, compte tenu de leurs prix unitaires. Sur la figure 5.3 de la page suivante, l'efficience

allocative correspond au rapport OA/OT et l'efficience coût au rapport OA/Oi. Pour l'entreprise i, compte tenu de sa production y et des prix relatifs de x_1 et x_2, l'efficience coût est maximale au point C (OC/OC= 1). Cet équilibre correspond à l'égalisation du taux marginal de substitution technique avec le rapport des prix des facteurs de production (pente de la tangente AC).

- **L'efficience revenu :** elle permet d'évaluer la capacité de l'entreprise à maximiser ses revenus compte tenu de ses prix de vente et des facteurs de production utilisés. L'efficience revenu est destinée à être utilisée par les entreprises, dans une logique de *benchmarking*.
- **L'efficience profit :** elle sert à estimer l'aptitude de l'entreprise à maximiser ses profits à partir des prix des facteurs de production et des produits. L'efficience profit peut être retenue par des actionnaires désireux de mettre en concurrence par comparaison l'entreprise dans laquelle ils ont placé leur capital.

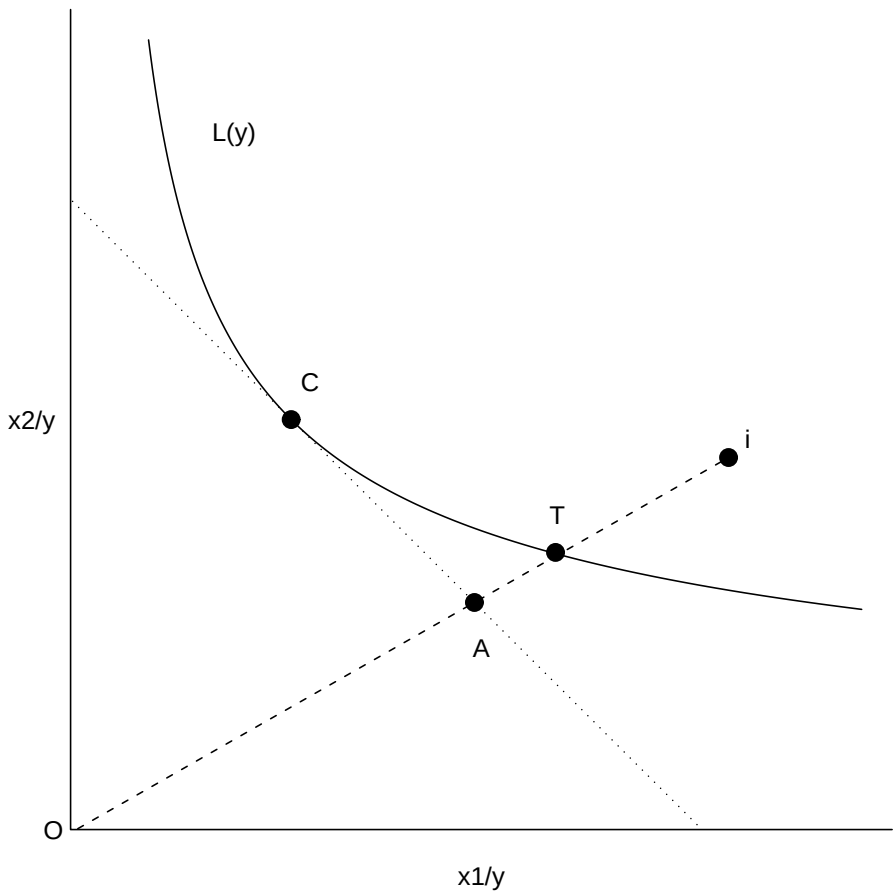

FIG. 5.3 – Décomposition de l'efficience coût pour la production de y par l'entreprise i, avec deux facteurs de production x_1 et x_2.
Source : Farell (1957) [103].

De toute évidence, les concepts d'efficience revenu et d'efficience profit ne correspondent pas à nos besoins. Il reste donc à choisir entre la simple efficience technique ou bien l'efficience coût qui tient compte, en plus, de l'efficience allocative. C'est cette dernière mesure de l'efficience que nous retenons pour trois raisons.

1. du point de vue analytique d'abord, prendre en compte l'efficience allocative en plus de l'efficience technique paraît souhaitable. Cela permet d'évaluer financièrement l'impact d'une surconsommation de certains facteurs de production, comme, par exemple, le recours à un, voire deux, contrôleur(s) pour accompagner un train, là où aucun, voire un seul, contrôleur serait suffisant.
2. du point de vue économétrique ensuite, il est plus rigoureux de fonder notre estimation de l'efficience sur la base des coûts (endogènes) auxquels correspondent une certaine production (exogène). En effet, dans un cas comme le nôtre la quantité de service produit est exogène à l'entreprise (définie par la région), bien qu'elle repose encore en partie sur un héritage non-marchand endogène à l'entreprise.
3. du point de vue de la disponibilité des données enfin, la mesure de l'efficience technique exige de connaître les quantités de facteurs de production consommées (nombre d'agents de conduite, de contrôleurs, volume d'énergie...). Or les données relatives aux ressources humaines ne sont pas connues, dans la mesure où le personnel non administratif n'est pas affecté à l'activité TER. Ainsi un conducteur tracte aussi bien des trains régionaux que nationaux, de voyageurs ou de fret, tandis qu'un guichetier vend tous les types de billets commercialisés par la SNCF. Il paraît donc impossible de développer une estimation fine de l'efficience technique.

Nous retenons donc pour estimer la performance des opérateurs ferroviaires, l'efficience coût, qui évalue à la fois l'efficience technique et l'efficience allocative des agents. Il faut à présent préciser quel *output* introduire face aux coûts pris en compte, pour estimer l'efficience coût de nos opérateurs. Afin de détailler ces choix, nous devons anticiper les biais qui pourraient survenir en présence des effets de dépendance spatiale, définis dans la partie précédente.

Prévenir les effets de dépendance spatiale

Nous avons vu au chapitre 3 que les effets de dépendance spatiale étaient une source d'hétérogénéité externe dont il fallait tenir compte dans un contexte comme le notre. Rappelons que ceux-ci proviennent des politiques particulières que les régulateurs régionaux peuvent poursuivre. Il s'agit de tenir compte du fait que chaque régulateur régional peut pratiquer une politique particulière, susceptible d'influencer les performances de son opérateur, et donc celles du référentiel de comparaison. Ces effets de dépendance spatiale sont d'autant plus difficilement contrôlables qu'ils ne sont pas toujours observables.

Dans notre cas, ces effets proviennent de ce que le taux d'accessibilité au service de transport ferroviaire régional est laissé à la libre appréciation du régulateur.

Contrairement à d'autres services en réseau (distribution d'eau, énergie, télécommunications...), il est impossible de desservir la totalité de la population régionale. Aussi, si l'on compare les trafics ferroviaires dans leur globalité, les différentes évaluations politiques de la pertinence de ce mode de transport peuvent biaiser les comparaisons. Si une région souhaite développer le transport ferroviaire sur des lignes où le trafic est structurellement faible, son opérateur pourra sembler peu performant. Inversement, si elle mise essentiellement sur les axes à fort trafic, son opérateur semblera performant si on ne prend pas garde au choix de l'*output*. Il faut donc s'efforcer d'évaluer la performance de l'opérateur, et non celle des politiques publiques qui dictent ses prestations. Nash (2000) [112] souligne l'impact du contexte d'exploitation et de l'intervention des pouvoirs publics sur l'efficience des entreprises ferroviaires. Notons que ce risque de distorsion non corrigible est cependant amoindri lorsque les comparaisons sont menées par une agence commune des régulateurs, ainsi que nous le proposons.

Considérant cet aspect de l'hérogénéité externe, posons-nous la question de savoir si l'on doit considérer comme demandeur du service TER l'autorité régionale ou bien les voyageurs. La pertinence de cette question n'est pas à démontrer, compte tenu de la contribution financière majoritaire de la région et de son implication dans la définition du service (c'est elle qui fixe le niveau de demande finale). C'est ce que précisent Nash & Shires (2000) [113] [12]. Cette question du client à considérer nous conduit à nous interroger sur l'étendue que nous souhaitons conférer à l'efficience de nos opérateurs.

Lorsque la Région est considérée comme demandeuse du service TER, la demande est exprimée en trains.kilomètres (TrK), et l'efficience coût ne prend pas en compte la performance commerciale de l'opérateur. Ce modèle présente deux avantages majeurs. D'une part, il nous permet d'exclure en partie la question de l'adéquation de l'offre à la demande [13] dans l'évaluation de la performance de l'opérateur. Ceci paraît souhaitable, dans la mesure où le régulateur régional définit l'offre, et où l'on souhaite n'évaluer que la performance de l'opérateur. D'autre part, ce modèle se justifie par le fait que c'est l'autorité régionale qui bénéficie des gains de productivité de l'opérateur ; c'est à ce niveau qu'il est possible d'observer une élasticité entre la

[12] « Les pouvoirs publics exercent également, par leur intervention, une influence décisive sur les performances. En plus de leur intervention dans les domaines de l'emploi et de l'investissement, ils interviennent fréquemment dans les décisions de tarification et de production des sociétés de chemins de fer. Pour ces entreprises ferroviaires, les mesures de performances fournissent typiquement des informations portant à la fois sur la performance des gestionnaires et sur le cadre institutionnel dans lequel ils agissent. Pour ce qui concerne le service voyageurs, il n'est pas rare que les pouvoirs publics assurent un contrôle effectif sur les horaires en terme de fréquence du service sur chaque itinéraire, soit dans le cadre d'un accord formel de franchisage, soit par le biais d'une obligation de service public. Il n'est pas interdit de considérer, dans une telle situation, que les pouvoirs publics deviennent le client, et que la production de l'entreprise ferroviaire est constituée par un certain niveau de service, plutôt que par le transport d'un certain nombre de personnes. [...] On peut dès lors penser qu'il pourrait s'avérer souhaitable de changer de manière plus radicale l'unité de production, en remplaçant les voyageurs.kilomètres ou les tonnes.kilomètres par les trains.kilomètres. »

[13] En effet, par convention contractuelle entre le régulateur régional et l'opérateur, l'offre de ce dernier doit correspondre à la demande définie par le premier.

demande et le prix payé par l'exécutif régional (la subvention).

En revanche, lorsque les voyageurs sont considérés comme demandeurs, la demande est exprimée en voyageurs.kilomètres (VK), et l'efficience coût prend en compte la performance commerciale de l'opérateur. Le taux de remplissage des trains induit alors un effet de dépendance spatiale important car c'est en définitive le régulateur régional qui juge s'il souhaite maintenir en circulation des trains peu fréquentés. Cela soulève la question de la pertinence des comparaisons de l'efficacité commerciale de la SNCF entre des régions aussi disparates que Rhône-Alpes, la Picardie et le Limousin, où les besoins de déplacements et les alternatives modales sont très différents. De plus, le trafic régional n'est pas directement lié au coût régional du service [14], et les voyageurs ne bénéficient pas directement des gains de productivité de l'opérateur. Aussi, il nous semble plus pertinent de développer un modèle fondé sur la logique précédente.

Conclusion : choix d'une mesure de la performance

Ainsi, les effets de dépendance spatiale nous conduisent à élaborer un modèle dont la demande émanerait de la région. Nous retenons l'efficience coût comme estimateur de la performance. Cela nous conduit à mettre en regard les coûts des opérateurs face à leur production de trains.kilomètres. Cette efficience est estimée à l'aide d'une frontière, méthodologie couramment retenue dans le domaine des transports (voir notamment les travaux de Estache & *al.* (2002) [101] et de Dalen & Gomez-Lobo (2003) [100] déjà cités). Cela nous permet d'évaluer, à travers des scores d'efficience, les efficiences technique et allocative. Nous nous rapprochons donc par ce choix de la régulation japonaise, bien que celle-ci porte sur un ensemble d'activités plus vaste, incluant la gestion de l'infrastructure.

Mais le coût des opérateurs dépend encore partiellement des cahiers des charges imposés par les régulateurs régionaux, dépendance qu'il faudra corriger. Il nous reste enfin à préciser la variable coût que nous utiliserons dans le modèle, compte tenu de nos objectifs et des données disponibles.

Le coût des services TER est classiquement décomposé en deux éléments :
- les charges d'exploitation forfaitisées, notées C_1, sur lesquelles la SNCF assume le risque industriel ;
- les charges notées C_2, dont la SNCF n'a pas la maîtrise (péage d'infrastructure, charges de capital liées au matériel roulant...). Celles-ci sont refacturées *ad valorem* à la région.

Aussi, le coût que nous retenons pour estimer l'efficience de nos opérateurs doit correspondre aux charges C_1.

Ces charges devront néanmoins être minorées des charges routières, généralement incluses dans le forfait. En effet, selon les régions, entre 0 % et 35 % des véhicules.kilo-

[14] Les voyageurs ne contribuent directement qu'aux recettes commerciales (qui ne couvrent qu'un tiers des coûts), fixées par les régions sur la base des tarifs nationaux et des contributions tarifaires supplémentaires qu'elles accordent.

mètres TER étaient réalisés en car en 2003, à un coût bien moindre qu'en train. Nous ne tenons évidemment pas compte ici de ces services routiers, qui n'ont pas vocation à être durablement gérés par la SNCF. Les charges C_1 devront également être minorées des pénalités pour non-circulation des trains (qui figurent dans un autre poste de charges). Cela permet de rendre compte, de la façon la plus exhaustive qui soit, de ce que coûte le service TER produit au cours d'une année. Ce coût couvrira ainsi les charges suivantes :

- conduite ;
- accompagnement (contrôle) ;
- énergie ;
- entretien (gares et matériel roulant) ;
- vente des titres de transport ;
- manœuvre ;
- gestion ;
- structure SNCF ;
- rémunération de l'exploitant.

À l'issue de ce chapitre, nous proposons donc d'appliquer la concurrence par comparaison aux marchés régionaux de transport ferroviaire. Mise en œuvre par une agence commune des régulateurs régionaux, cette forme de régulation doit donc inciter les opérateurs à réduire leurs coûts, à l'image des résultats obtenus dans les chemins de fer japonais. Pour y parvenir, nous suggérons d'évaluer la performance des opérateurs à travers l'efficience coût. Nous proposons d'estimer une frontière de coût. Celle-ci doit mettre en regard les charges d'exploitation endogènes (C_1), et, au titre d'*output* correspondant, le nombre de trains.kilomètres produits à la demande des régions. Ce cadre général, construit afin de compléter la régulation actuelle, contourne les principales limites théoriques des mécanismes de concurrence par comparaison, notamment les effets de dépendance spatiale. Il nous reste à affiner le modèle économique de nos comparaisons, puis à le décliner en termes économétriques afin de produire des résultats transparents.

Chapitre 6

Estimation de la performance par frontière

L'analyse développée au chapitre précédent nous a permis de définir :

- le cadre d'application de notre proposition de concurrence par comparaison (la régulation des 20 activités TER de la SNCF) ;
- la mesure de la performance que nous souhaitons évaluer (l'efficience coût, estimée par une frontière de coût).

Dans ce chapitre, nous précisons d'abord le modèle économique, la spécification de la fonction de coût, et les différentes variables environnementales susceptibles de l'affecter. Nous détaillons ensuite le modèle économétrique de frontière auquel nous avons recours. Nous achevons enfin ce travail par une interprétation économique des résultats estimés économétriquement. Cette ultime étape nous permet alors d'attribuer à un manque d'effort l'inefficience observée dans certaines régions.

6.1 Le modèle économique

Nous commençons par présenter un modèle économique complet des marchés régionaux du transport ferroviaire de voyageurs. Il s'agit ainsi de modéliser, d'une part, la relation économique entre les régions et la SNCF et, d'autre part, la performance du service fourni, évaluée par l'efficience coût des opérateurs. Dans un second temps, nous verrons que nous sommes contraints de nous restreindre à la deuxième partie du modèle (modélisation de la performance), afin de pouvoir estimer correctement notre frontière d'efficience.

6.1.1 Structure du modèle complet

La modélisation de l'ensemble de la relation économique entre chaque opérateur régional et son régulateur a d'abord semblé pertinente. Elle vise à tenir compte du contexte économique au sein duquel notre mesure de la performance s'inscrit. Il s'agit notamment d'introduire une élasticité prix-demande dans la fonction de demande de la région. Nous reprenons dans les paragraphes qui suivent et en annexe le modèle que

nous avions proposé dans Lévêque (2003) [108]. Nous définissons d'abord les variables du modèle et les relations qui les lient, avant de détailler les limites rencontrées et de nous concentrer sur la seule frontière d'efficience.

Les variables du modèle

Nous décrivons brièvement les variables qui interviennent dans la modélisation économique d'un marché régional de transport ferroviaire. À ce stade, nous ne faisons intervenir que les variables économiques en précisant leurs moyenne et coefficient de variation pour 1998 ; les variables environnementales mobilisées interviendront ultérieurement.

- ***TrK*** **:** offre de services TER mise en œuvre par la SNCF ; elle est exprimée en trains.kilomètres, et supposée égale à la demande contractuelle de la région ($\overline{X} = 5,58.10^6$; $\sigma/\overline{X} = 59$ %) ;
- ***VK*** **:** trafic exprimé en voyageurs.kilomètres ($\overline{X} = 370.10^6$; $\sigma/\overline{X} = 74$ %) ;
- ***R*** **:** recettes perçues par l'opérateur : il s'agit des recettes commerciales (tirées de la vente des titres), des compensations tarifaires pour tarifs sociaux, et des recettes annexes : publicité, amendes, *etc.* ($\overline{X} = 32,0.10^6$€ 2004 ; $\sigma/\overline{X} = 74$ %) ;
- ***CT*** **:** compensations tarifaires destinées à combler le manque à gagner sur les tarifs sociaux et régionaux (indisponible en 1998) ;
- $\boldsymbol{C_1}$ **:** charges de production du service TER pour l'opérateur ($\overline{X} = 54,9.10^6$€ 2004 ; $\sigma/\overline{X} = 61$ %) ;
- $\boldsymbol{C_2}$ **:** charges exogènes, remboursées par la région *ad valorem* (péages d'infrastructure, charges de capital liées au matériel roulant, *etc.*) ($\overline{X} = 24,6.10^6$€ 2004 ; $\sigma/\overline{X} = 62$ %) ;
- ***S*** **:** subvention versée par la région à l'opérateur pour le service de transport ($\overline{X} = 47,5.10^6$€ 2004 ; $\sigma/\overline{X} = 56$ %) ;
- ***s*** **:** subvention unitaire au train.kilomètre ($\overline{X} = 8,71$€ 2004 ; $\sigma/\overline{X} = 14$ %) ;
- ***C*** **:** coût de revient à la collectivité et aux voyageurs du service TER de l'opérateur. *Ex post*, on a : $C = R + S - C_2$. Ce coût correspond à la somme des coûts C_1 et de la rémunération de l'opérateur.

Ces variables économiques sont dépendantes les unes des autres. Nous précisons à présent les relations fonctionnelles que nous considérons.

Les relations entre les variables

Nous continuons de nous concentrer sur les variables économiques du modèle, tout en mentionnant les variables environnementales susceptibles d'intervenir dans les relations fonctionnelles. La notation $\boldsymbol{= f(.)}$ indique une relation de dépendance, sans qu'il s'agisse à chaque fois de la même fonction $\boldsymbol{f}$.

- $\boldsymbol{TrK = f(s)}$: la demande contractuelle de la région dépend du montant de subventions fixé par l'opérateur. Derrière cette trivialité économique se cache une hypothèse très forte, concernant l'importance du caractère marchand dans la relation région-opérateur. Le montant de subvention publique est, bien sûr,

un élément capital dans la décision des autorités régionales de mettre en place de nouveaux services TER. Mais l'essentiel des actuels services TER provient directement de l'héritage transféré de l'État aux régions. Or ces services ont été maintenus au fil des ans pour diverses raisons non économiques (niveau de trafic acceptable, demande politique, *etc.*), en l'absence de toute relation marchande.
Même à présent, l'utilité qu'ont les autorités régionales du train.kilomètre par euro public investi varie très largement avec la composition politique du Conseil Régional. Ainsi, le nombre de trains.kilomètres est plutôt déterminé exogènement par des éléments géographiques, morphologiques et politiques (la population régionale, sa distribution dans l'espace, *etc.*). Cela fait craindre l'apparition de distorsions difficilement corrigibles [1].

- $\boldsymbol{VK = f(TrK)}$: dans la logique de notre modèle, le trafic voyageurs dépend de l'offre ferroviaire (notamment en terme de fréquence de dessertes). Cette fonction de demande ne fait pas intervenir de prix, dans la mesure où les conditions tarifaires (prix au voyageur.kilomètre) sont identiques pour l'ensemble des services TER. Toutefois, les régions proposent des réductions de tarif, susceptibles de favoriser un surcroît de trafic. Ces réductions sont remboursées à l'opérateur à travers les compensations tarifaires (CT). Il conviendra donc de tester si une forme $f(TrK, CT)$ n'expliquerait pas mieux le trafic.
Enfin, d'autres variables environnementales expliquant le niveau de trafic devront être introduites. Elles devront expliquer le niveau de trafic relativement faible en Limousin ou, au contraire, relativement fort en Picardie, Nord-Pas-de-Calais ou PACA. Là encore, des distorsions difficiles à corriger s'annoncent. Par exemple, le fort trafic picard est dû à l'attrait de la région capitale toute proche, mais néanmoins hors de la région Picardie.
- $\boldsymbol{R = f(VK)}$: les recettes dépendent du trafic de façon quasi linéaire. On peut vérifier par régression qu'une forme linéaire $R = r\,VK$ peut être tout à fait valablement retenue ($R^2 = 100\%$) [2].
- $\boldsymbol{C_1 = f(TrK)}$: nous avons présenté, dans la section 1.1, la théorie relative à la définition de fonctions de coût. Nous y avons justifié pourquoi, dans un travail comparatif basé sur les données comptables de la SNCF, l'absence des prix unitaires du travail et des autres facteurs de production (quasi-identiques d'une région à l'autre, mais indisponibles), était acceptable. Dans le cas présent, il ne s'agit plus d'estimer une fonction de coût total, mais une fonction de coût réduite aux seules charges de production de l'opérateur. Afin d'estimer

[1] En effet, certaines régions comme le Limousin ou l'Auvergne ont hérité d'un service très étoffé par rapport à la demande de transport des habitants. Par ailleurs, nous avons pu vérifier que les indicateurs de richesse régionale (PIB, revenu fiscal) ne contribuent pas à expliquer la demande régionale en trains.kilomètres.

[2] Ce résultat est logique, dans la mesure où le prix d'un billet plein tarif est une fonction affine par morceaux du parcours kilométrique. Ce résultat indique que la distribution des distances de déplacement est similaire dans chaque région. De plus, les compensations tarifaires (incluses dans les recettes) sont calculées de façon suffisamment précise pour combler le manque à gagner sur les tarifs sociaux. Enfin, les recettes annexes s'expliquent en partie par le trafic et en partie par l'efficacité commerciale de l'opérateur ; elles ne représentent qu'une faible part des recettes totales.

la performance productive des opérateurs, nous considérons comme produit le nombre de trains.kilomètres. Toutefois, afin de conserver une certaine similitude avec les fonctions de coût établies pour définir les rendements d'échelle, nous introduisons la variable longueur de lignes (LL) au titre de quasi-produit [3]. Cette variable peut rendre compte de l'influence de l'étendue du réseau sur les coûts de production de l'opérateur.
Nous retenons une fonction de coût correspondant à une fonction de production de type Cobb-Douglas. En effet, les variations observées des charges C_1, en fonction du nombre de trains.kilomètres, suivent une forme log-linéaire :

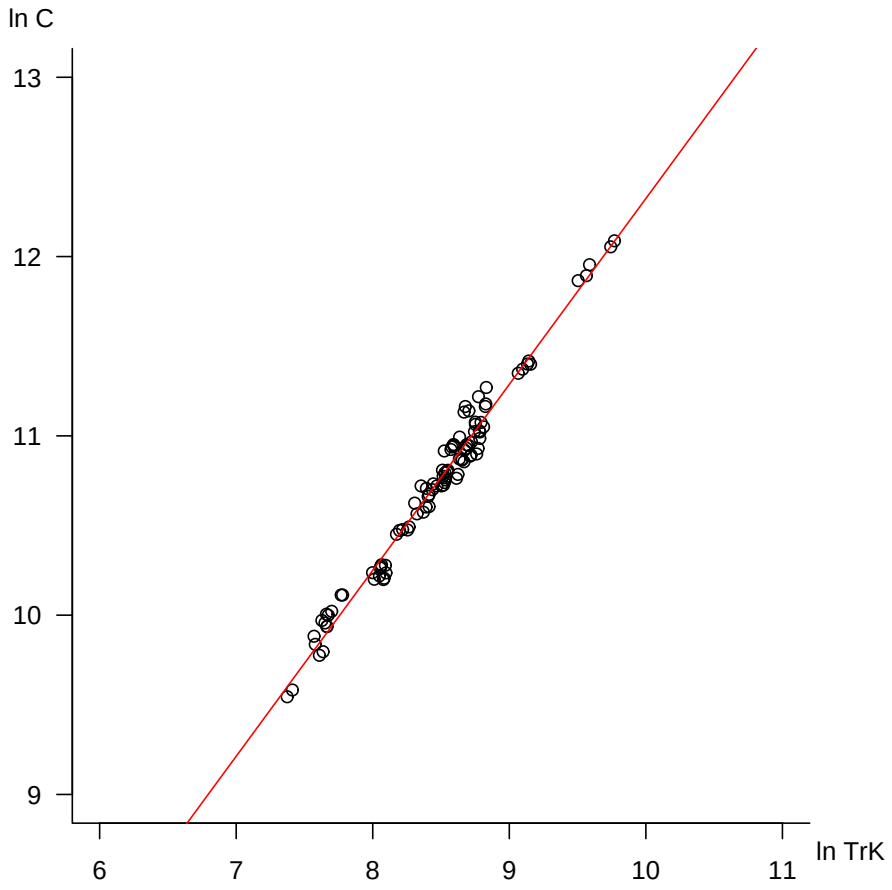

FIG. 6.1 – Variation du coût en fonction de la production : observations des 20 régions (1993-98 hors 1996) et droite de régression ($R^2 = 98\%$).

Il n'est donc pas justifié de recourir à une forme fonctionnelle plus souple, de type translog. Nous définissons donc notre fonction de coût C_1 de la façon suivante :

$$\ln C_1 = \beta_0 + \beta_{TrK} \ln TrK + \beta_{LL} \ln LL$$

[3] Il ne s'agit plus ici à proprement parler d'un produit, dans la mesure où notre fonction de coût (réduite aux coûts d'opération) n'inclut plus les charges d'infrastructure.

- $\boldsymbol{C_2 = f(TrK, VK)}$: les charges exogènes dépendent du nombre de circulations, auxquelles correspondent une redevance d'infrastructure et une valeur du parc de matériel roulant. Elles dépendent aussi du trafic, qui rend compte indirectement de la catégorie tarifaire des voies utilisées et de la lourdeur du matériel roulant. Toutefois, ces charges - notamment les péages d'infrastructure - sont difficilement modélisables, compte tenu du grand nombre de facteurs exogènes qui contribuent à en fixer le niveau.
- $\boldsymbol{s}$: la subvention unitaire au train.kilomètre est la variable prix autour de laquelle s'ajuste l'équilibre économique.
- $\boldsymbol{S = s.TrK}$
- $\boldsymbol{C = R + S - C_2}$: le coût des services TER de l'opérateur pour la collectivité et les voyageurs est la somme de la subvention versée par la région et des recettes R, diminuée des charges exogènes C_2 (inclues dans S).

Cette analyse des relations fonctionnelles entre les variables économiques du modèle, nous indique la forme de ce dernier. Toutefois, nous avons mis en évidence un certain nombre de difficultés, en terme de mesurabilité, susceptibles de s'opposer à l'opérationalité du modèle.

Synthèse et limites

Nous récapitulons dans ce paragraphe les relations entre les variables du modèle, illustrées par la figure 6.2 de la page suivante. Le schéma propose un mécanisme aboutissant à offrir TrK trains.kilomètres pour un coût C :

1. Un équilibre prix s - demande TrK s'établit entre une région et la direction locale de la SNCF. Il correspond à une subvention totale de la région, d'un montant $S = s.TrK$.
2. Afin de réaliser son engagement contractuel, la SNCF offre aux voyageurs TrK trains.kilomètres, ce qui lui coûte directement C_1.
3. Cette offre ferroviaire, ainsi que les compensations tarifaires, contribuent à générer un volume de déplacements, le trafic, VK.
4. Ce trafic engendre un volume de recettes commerciales $R = r\,VK$.
5. Ces recettes, la subvention totale versée par la région et les charges exogènes C_2 définissent le coût total C payé par la collectivité et les voyageurs à l'opérateur, pour la seule exploitation des TrK trains.kilomètres. C'est à partir de cette variable que nous estimons la frontière de coût et l'efficience des opérateurs.

Nous détaillons en annexe (page 229) la poursuite de ce modèle, notamment les propositions de spécification des fonctions objectif et demande, ainsi que la résolution du problème de maximisation. En effet, nous ne pouvons poursuivre plus loin ce modèle pour les raisons suivantes :

- **La difficile spécification de certaines fonctions :** Ainsi que nous l'avons déjà précisé, les fonctions de demande de la région (TrK) et des voyageurs (VK) font appel à de très nombreuses variables exogènes. Or nous ne disposons pas de modèle adapté à ce cadre agrégé qui inclut à la fois des relations

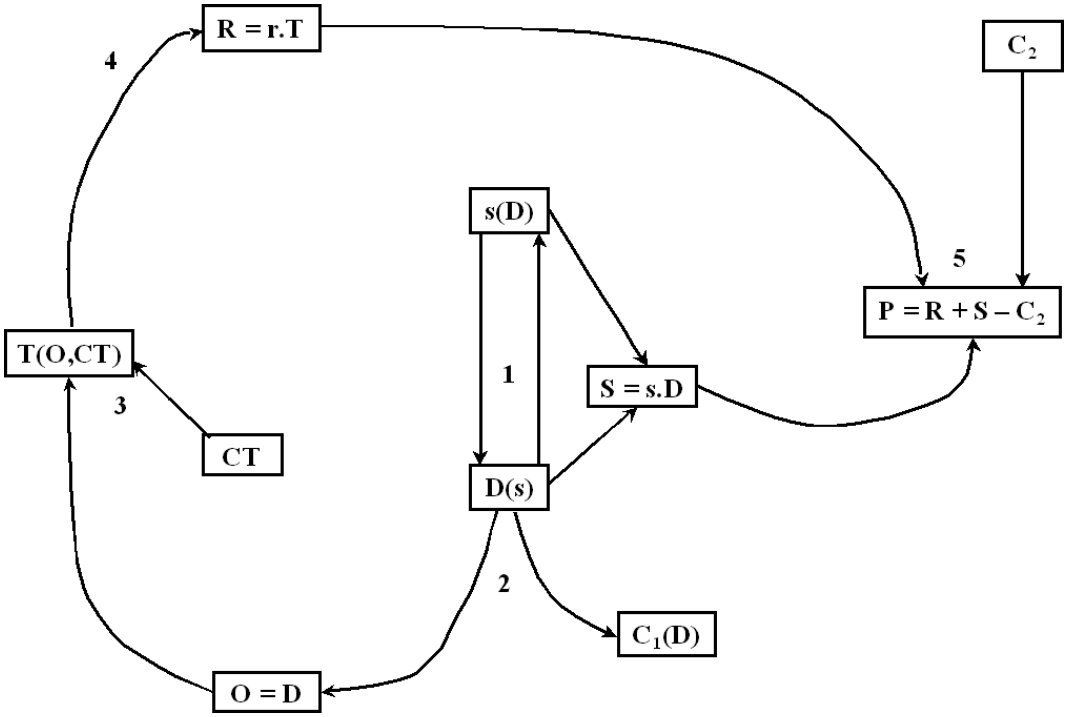

FIG. 6.2 – Structure du modèle.

périurbaines et interurbaines. Il faudrait donc recourir à l'agrégation des fonctions de demande de chaque liaison, ce qui complexifie le modèle et réduit sa mesurabilité.
Pour des raisons similaires, l'estimation d'une fonction de coût C_2 se révèle très délicate. Compte tenu du mode de détermination des péages d'infrastructure, il n'apparaît pas possible de modéliser ces charges dans un cadre agrégé, simple. De plus les fortes variations temporelles des redevances d'infrastructure (doublement entre 2002 et 2005) ne facilitent pas leur intégration au modèle.

- **La non-mesurabilité du modèle :** Le modèle économétrique issu de cette modélisation fait intervenir de très nombreux paramètres à estimer (plus d'une dizaine, dans le cas le plus simple). Il faut donc disposer d'un nombre conséquent d'observations pour estimer ce modèle complet. Or le nombre limité d'opérateurs (20) et le jeune âge de cette configuration des marchés n'offrent qu'une soixantaine d'observations en 2005.
- **L'insuffisance du caractère marchand :** Ainsi que nous l'avons déjà précisé, ce modèle est justifié par la prise en compte d'une élasticité de la demande régionale au coût du service. Or, il ne semble pas qu'une telle élasticité soit apparente, tant les services régionaux reposent encore sur un héritage non marchand.

Si la cohérence du modèle établi est acceptable, quoique discutable, la description précédente prouve que nous butons sur un double problème de pertinence et de mesurabilité. D'un côté il semble difficile de proposer une approximation du réel qui soit de qualité. D'un autre côté, l'évaluation des quantités mises en jeu et des paramètres des équations semble délicate avec si peu de recul. Un tel conflit rend donc notre modèle inopérationnel. Pour dépasser cet obstacle, nous sommes contraints de réduire l'objectif du modèle. Nous abandonnons donc la modélisation du contexte économique, afin de privilégier l'estimation de l'efficience. Ainsi, notre travail se trouve restreint à la construction de la frontière de coût, spécifiée par la forme fonctionnelle $C = R + S - C_2 = \beta_0 TrK^{\beta_{TrK}} LL^{\beta_{LL}}$. Nous discutons à présent les variables environnementales, susceptibles d'influer sur l'efficience des opérateurs ; nous étudions dans quelle mesure l'hétérogénéité introduite doit être corrigée.

6.1.2 Correction de l'hétérogénéité

Nous avons vu à la section 3.3 qu'il était nécessaire de tenir compte de l'environnement externe aux opérateurs, pour ne pas distordre le mécanisme comparatif. Afin de « comparer ce qui est comparable », nos scores d'efficience doivent donc tenir compte des variables environnementales qui en expliquent la part exogène, et biaisent les comparaisons. Il est possible de distinguer quatre ensembles de variables environnementales de natures différentes. Nous les discutons en précisant l'impact attendu de chaque variable sur la fonction de coût, évalué à l'aide d'une estimation de l'élasticité[4] :

- éléments liés à l'infrastructure disponible :
  - **La qualité de la voie** (relief, sinuosité, capacité) détermine la vitesse des trains et influe sur l'efficience technique. En effet, à effectif identique, le nombre de rotations du matériel et des agents augmente avec la vitesse. L'élasticité estimée coût-vitesse est de l'ordre de 20 %. Toutefois, les vitesses moyennes sont suffisamment homogènes pour ne pas engendrer de biais dans les comparaisons.
  - **L'électrification** d'une ligne permet de recourir à du matériel roulant à traction électrique. Les coûts d'exploitation sont alors inférieurs (de 0,3 € par train.kilomètre en moyenne) à ceux d'un matériel similaire à traction thermique (voir tableau 6.1 page suivante). Ce facteur influe toutefois très peu sur l'efficience allocative (élasticité de l'ordre de 1 %). Bien que la part de trains.kilomètres à traction thermique varie fortement entre les régions, l'impact estimé au niveau des charges totales de l'opérateur est négligeable.
  - **La localisation des établissements** (du matériel et de la traction, ou d'entretien et de maintenance) par rapport à l'ensemble des activités de transport, peut induire des distorsions entre les régions. Suivant les régions, la part des parcours à vide est plus ou moins importante (1 % dans le Limousin, contre 14 % en Haute-Normandie en 1998). En revanche, l'impact estimé sur les coûts est faible (élasticité de l'ordre de 3 %). Cependant, nous

[4] Le détail des calculs d'élasticité figurent en annexe, page 233.

Rame tractée (une locomotive et cinq voitures)			
diesel		électrique	
heures pleines	heures creuses	heures pleines	heures creuses
7,5	5,8	7,2	5,4

Automoteur (premier élément)			
diesel		électrique	
heures pleines	heures creuses	heures pleines	heures creuses
5,6	3,9	5,4	3,7

Automoteur (élément supplémentaire)	
diesel	électrique
1,9	1,6

TAB. 6.1 – Prix des circulations au train.kilomètre en € 2001 en fonction du type de matériel, de l'énergie de traction et de la période (hors péage d'infrastructure). Source : SNCF Convention Pays de la Loire.

pensons que le taux de parcours effectué à vide est une source d'inefficience relativement endogène. Il appartient aux opérateurs de la réduire par une remise en cause de leur organisation spatiale.

- éléments liés au matériel roulant :
  - **L'âge et la lourdeur du matériel roulant utilisé** augmentent les coûts d'exploitation (consommation d'énergie, besoins en accompagnement), d'entretien et de maintenance. Sur la base du tableau 6.1, on estime une élasticité du coût à la part de service réalisée avec des rames tractées, de l'ordre de 5 %. Sachant que la part de service réalisée avec un tel matériel varie fortement entre les régions, il paraît souhaitable de tenir compte de ce facteur. Il n'est pas possible de connaître le nombre de trains.kilomètres réalisés, par type de matériel roulant, dans chaque région. En revanche, nous pouvons introduire la variable *VK* qui indique l'importance du trafic et, partant de là, le recours nécessaire à du matériel roulant plus lourd.
- éléments liés aux contraintes de service public :
  - **Le nombre d'arrêts moyen par train.kilomètre** affecte les coûts d'exploitation des trains (dont la vitesse commerciale est réduite) et des gares (qui sont plus actives).
  - **La répartition des Tr.K entre heures pleines et heures creuses** influe également sur les prix (voir tableau 6.1). Une région qui développe l'offre en heure creuse bénéficie de prix moins élevés que si elle développe les TER en heure pleine. Il est délicat de définir une variable à même de rendre compte de ce phénomène. Cependant, il ne semble pas que la répartition du service, entre heures pleines et heures creuses, varie fortement d'une région à l'autre [5].

[5] À partir des horaires, on peut vérifier que la distribution des trains partant de la préfecture

- éléments sociaux :
  - **Le taux de délinquance régional** influe sur les charges au sol (personnel dans les gares), d'accompagnement et d'entretien (tant du matériel roulant que des gares). À partir de données portant sur le métro lyonnais, nous avons estimé une élasticité des coûts au taux de délinquance qui paraît assez faible (de l'ordre de 2 %). Cette valeur doit être considérée avec précaution, du fait des différences d'exploitation entre un métro et un train.
  - **Les mouvements sociaux** ont un impact pris en compte dans le modèle : la variable TrK correspond au nombre de trains.kilomètres ayant réellement circulé ; elle est donc nette des effets grèves. De même, la définition que nous avons donnée de la variable de coût C tient compte des perturbations du trafic pour cause de mouvement social. En effet, nous avons précisé que la variable C devait intégrer les pénalités versées par la SNCF en cas de non-réalisation de l'offre. Ces pénalités, négociées dans chaque convention régionale, sont plus ou moins élevées suivant les régions. Toutefois, elles demeurent bien inférieures au coût du train.kilomètre, de sorte que les coûts paraîtront d'autant plus élevés dans une région qu'il y aura eu d'importants mouvements sociaux.

Cette analyse des différents facteurs susceptibles de biaiser les comparaisons d'efficience achève de fixer notre modèle économique. Toutefois, dans l'esprit d'une mise en œuvre consensuelle de la concurrence par comparaison, il appartient aux opérateurs de proposer d'inclure d'autres variables environnementales. Si le régulateur estime que de telles variables sont susceptibles d'expliquer une part d'efficience, exogène aux opérateurs, il doit les ajouter aux autres variables environnementales. Nous présentons dans la section suivante le modèle économétrique qui permet d'estimer enfin la frontière et les scores d'efficience de nos opérateurs, en tenant compte des variables environnementales que nous venons d'étudier.

6.2 Le modèle économétrique

La démarche suivie à la section 5.2 nous a conduits à construire une frontière de coût, définissant le coût minimum à atteindre pour une production donnée. Dans cette section, nous exposons la méthodologie retenue afin d'estimer économétriquement les scores d'efficience. Il s'agit d'abord de définir la frontière de coût à estimer, puis d'évaluer l'efficience coût, corrigée des variables environnementales.

6.2.1 La frontière *stochastique* de coût

Les choix effectués dans les paragraphes qui suivent s'appuient sur les résultats de Lovell (1993) [110] et de Kumbhakar & Lovell (2000) [106]. Il s'agit de privilégier l'approche paramétrique, stochastique.

régionale au cours des six heures les plus chargées de la journée, est peu dispersée ($\sigma/\bar{x} = 8\%$). Cela semble suffisant pour garantir l'absence de distorsion de ce fait lors de l'estimation des scores d'effcience.

L'inefficience coût paramétrique

L'inefficience coût des opérateurs est évaluée par l'écart entre leur position et la frontière. La littérature distingue deux approches permettant d'évaluer l'efficience, à l'aide de frontières :

- **L'approche non-paramétrique :** elle consiste à construire la frontière, sans spécifier de fonction de coût pour le support. La frontière est établie point par point, à partir des observations, à l'aide des techniques de programmation linéaire [6].
- **L'approche paramétrique :** elle consiste à estimer la frontière économétriquement, à partir de la forme fonctionnelle spécifiée.

Le recours à une frontière paramétrique présente plusieurs avantages. Il permet notamment d'exploiter les propriétés statistiques des scores d'inefficience pour expliquer l'inefficience et, dans notre cas, la corriger des variables environnementales. Il permet également de disposer de certains résultats intermédiaires, comme les rendements de production (β_{TrK}). Enfin, il offre un gage de transparence à notre travail. Signalons aussi que la spécification de la fonction de coût ne présente pas de difficulté particulière, ainsi que nous l'avons vu à la section précédente (figure 6.1).

En incluant un paramètre d'inefficience, la fonction de coût prend la forme suivante :

$$C_{i,t} = f(TrK_{i,t}, LL_{i,t}).IC_{i,t}$$

où $IC_{i,t}$ est l'inefficience coût de l'opérateur i à la période t.
Ce terme est défini par le rapport entre le coût observé et le coût minimum pouvant être obtenu : $IC_{i,t} = C_{i,t}/f(TrK_{i,t}, LL_{i,t}) > 1$.

En log-linéarisant la relation, cette inefficience se trouve modélisée par un terme d'erreur asymétrique :

$$\begin{aligned} \ln C_{i,t} &= f(TrK_{i,t}, LL_{i,t}) + u_{i,t} \\ \text{avec } u_{i,t} &\geqslant 0 \end{aligned}$$

L'inefficience coût est alors : $IC_{i,t} = e^{u_{i,t}}$.

Parmi les différents traitements économétriques envisageables, nous choisissons de travailler sur des données de panel, compte tenu du faible nombre d'observations disponibles chaque année (20). Le travail sur plusieurs années permet d'affermir l'estimation de la frontière, tout en permettant de calculer des scores d'efficience pour chaque observation (individu ; année).

L'approche stochastique

L'écriture que nous avons adoptée jusque là laisse supposer que l'intégralité du terme d'erreur correspond à de l'inefficience, dans une approche déterministe. Afin de tenir

[6] Les principales techniques employées sont la méthode DEA et la méthode FDH.

compte des aléas autres que l'inefficience (chocs exogènes, erreurs de mesure...), nous estimons plutôt une frontière stochastique. L'analyse par frontière stochastique permet de distinguer, dans les écarts résiduels à la frontière, ce qui relève de l'inefficience et ce qui relève des perturbations exogènes. Le terme résiduel est alors composé d'un terme d'inefficience $u_{i,t}$ et d'un terme stochastique $v_{i,t}$. Sur la figure 6.3 ci-dessous, le surcoût de l'entreprise i est dû à une forte inefficience (u_i), partiellement compensée par un choc exogène qui lui a été favorable (v_i). Inversement, le surcoût de l'entreprise j résulte d'une faible inefficience (u_j), à laquelle s'ajoute un aléa défavorable (v_j). L'identification des deux termes u et v repose sur l'hypothèse de distribution des $u_{i,t}$ [7] ; les $v_{i,t}$ sont supposés indépendants et identiquement distribués, suivant une loi normale centrée.

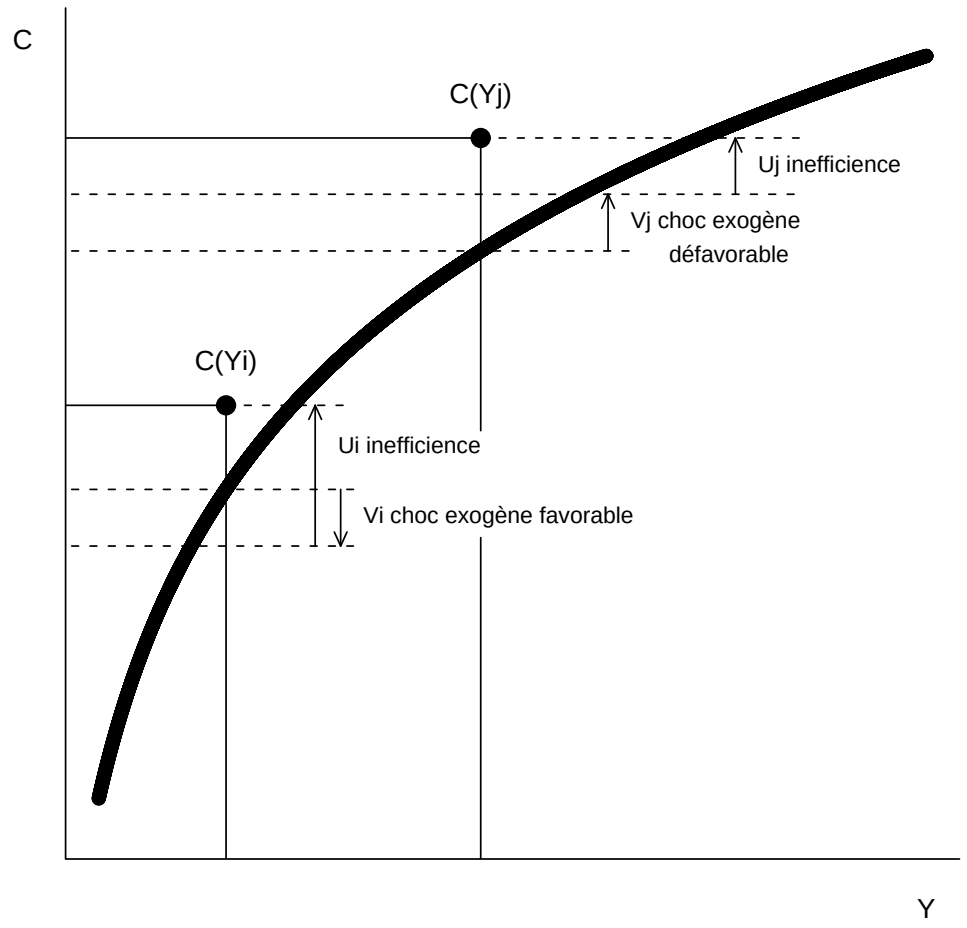

FIG. 6.3 – Frontière stochastique de coût : décomposition du terme résiduel.

[7] Les distributions fréquemment retenues sont les lois semi-normale, exponentielle ou normale tronquée. Voir Kumbhakar & Lovell (2000) [106] pour une présentation détaillée de ces lois ; voir aussi Chaffai (1997) [94] qui recommande plutôt l'usage d'une loi normale tronquée.

La frontière à estimer s'écrit alors :

$$\ln C_{i,t} \quad = \quad f(TrK_{i,t}, LL_{i,t}) + v_{i,t} + u_{i,t}$$

6.2.2 Intégration des variables environnementales

Le terme d'inefficience $u_{i,t}$ doit ensuite être régressé sur les variables environnementales retenues (les z_k) :

$$u_{i,t} = \delta_0 + \sum_k \delta_k z_{k_{i,t}} + w_{i,t}$$

où $\delta_.$ sont les coefficients linéaires,
$w_{i,t}$ est le résidu (égal au logarithme de l'efficience nette). Nous précisons dans ce qui suit le procédé retenu afin de passer de l'inefficience brute ($IC = e^u$) à l'inefficience nette des variables environnementales (e^w). Nous choisissons d'estimer simultanément la frontière et l'impact de ces variables. Cette estimation simultanée permet d'éviter la contradiction des modèles dits « à deux étapes » (voir Coelli, Perelman & Romano (1999) [97]). En effet, ces modèles qui estiment d'abord les coefficients de la frontière retournent alors des scores d'inefficience brute, supposés indépendants et identiquement distribués. Mais ces scores sont ensuite régressés sur les variables environnementales, ce qui signifie qu'ils ne sont pas identiquement distribués (à moins que tous les coefficients associés aux variables environnementales soient nuls).

D'un autre côté, l'estimation simultanée n'est plus recommandée dès lors qu'il manque des variables environnementales susceptibles d'expliquer de l'inefficience (voir Chaffai (1997) [94]). En effet, une erreur de spécification des variables environnementales biaisera non seulement les coefficients estimés, associés à ces variables, mais aussi ceux de la frontière.

Afin de mettre en œuvre ce procédé, nous recourons au logiciel Frontier 4.1 (voir Coelli (1996) [96], Battese & Coelli (1995) [91] et Coelli & *al.* (1998) [98]). Les coefficients de la frontière sont estimés par la méthode du maximum de vraisemblance. Après initialisation par les MCO, le programme procède aux itérations suivant la procédure de Davidon-Fletcher-Powell Quasi-Newton.

La frontière stochastique de coût estimée a la forme suivante :

$$\begin{aligned} \ln C_{i,t} &= \beta_0 + \beta_{TrK} \ln TrK_{i,t} + \beta_{LL} \ln LL_{i,t} + v_{i,t} + u_{i,t} \\ \text{avec } u_{i,t} &= \delta_0 + \delta_{\text{délinq}} \ln \mathit{délinq}_{i,t} + \delta_{VK} \ln VK_{i,t} + w_{i,t} \end{aligned}$$

où

- $v_{i,t}$ est le terme stochastique, correspondant au choc exogène qui affecte l'observation (i,t). Les $v_{i,t}$ sont supposés indépendants et identiquement distribués suivant $\mathcal{N}(0, \sigma_v^2)$ et indépendants des $u_{i,t}$.
- $u_{i,t}$ est le logarithme de l'inefficience brute. Les $u_{i,t}$ sont positifs et indépendamment distribués suivant $\mathcal{N}(m_{i,t}, \sigma_u^2)$ tronquée en zéro, avec :
  $m_{i,t} = \delta_0 + \delta_{\text{délinq}} \ln \mathit{délinq}_{i,t} + \delta_{VK} \ln VK_{i,t}$

- *délinq* et VK, respectivement le taux de délinquance régional et le trafic, sont les variables environnementales qui permettent de tenir compte de l'hétérogénéité externe.
- $w_{i,t}$ est le logarithme de l'inefficience nette des variables environnementales.

Le modèle étant définitivement précisé, il reste à en présenter les résultats et à interpréter ceux-ci économiquement.

6.3 Les résultats

À partir des données de la comptabilité régionale de la SNCF entre 1993 et 1998 (hors 1996), nous avons pu estimer une frontière stochastique de coût, et en déduire des scores d'efficience régionaux pour chaque année. Les statistiques descriptives de ces données sont présentées dans le tableau 6.2 :

Variable	Moyenne	Écart-type
Coûts (k€ 2004)	54525	31682
Trains.kilomètres (milliers)	5350	2910
Longueur de lignes (km)	1070	326
Taux de délinquance (%)	5,340	1,591
Trafic (milliers de VK)	347958	245646

Tab. 6.2 – Statistiques descriptives régionales.
Sources : SNCF, Ministère de l'Intérieur.

Ces statistiques renseignent ainsi sur le coût moyen du train.kilomètre (pour les seules charges C_1), proche de 10 euros. Le taux de remplissage des trains s'élève à 65 voyageurs en moyenne. Signalons également le cas de la région Rhône-Alpes, où circulent deux fois plus de trains que dans n'importe quelle autre région ; il faudra veiller à ce que l'estimation de son efficience ne soit pas perturbée par un phénomène d'hétéroscédasticité. Nous exposons d'abord les résultats économétriques obtenus ; nous discutons ensuite l'interprétation économique des scores d'efficience calculés.

6.3.1 Résultats économétriques

Les résultats d'estimation de la frontière sont présentés dans le tableau 6.3 page suivante ; l'ensemble du programme et de ses résultats figure en annexe, page 235.

Dans les paragraphes qui suivent, nous discutons les résultats de l'estimation concernant :

- les *outputs* ;
- les variables environnementales ;
- l'inefficience.

Variable à expliquer : $\ln C$			
Observations : 100			
Paramètre	Valeur	Erreur	t-test
β_0	2,687	0,191	14,1
β_{TrK}	1,002	0,025	40,7
β_{LL}	-0,08914	0,02541	-3,51
δ_0	0,005181	0,029783	0,174
$\delta_{\text{délinq}}$	2,687	0,380	7,08
δ_{VK}	0,00009059	0,00004580	1,98
σ^2	0,002585	0,000395	6,54
γ	0,9999	0,3305	3,03

TAB. 6.3 – Estimation de la frontière stochastique.

Les *outputs*

On vérifie aisément que le coefficient associé à la variable explicative $\ln TrK$ n'est pas significativement différent de 1. Cela signifie que les rendements de production sont constants. Le signe négatif du coefficient associé à la variable $\ln LL$ n'est pas très surprenant. En effet, ce quasi-*output* ne peut plus expliquer beaucoup de variance en plus de $\ln TrK$. Rappelons par ailleurs que notre mesure des coûts n'intègre pas les charges d'infrastructure. On peut interpréter ce signe par les surcoûts induits par la densité de circulations ferroviaires sur un réseau donné[8]. La courbe de frontière estimée, ainsi que les observations, sont représentées sur la figure 6.4 page suivante.

La figure 6.4 illustre la frontière de coût et les écarts qui la séparent des observations. Économétriquement, ces écarts s'interprètent par l'influence des variables environnementales, la présence d'inefficience et l'impact du terme d'erreur stochastique.

Les variables environnementales

Les variables environnementales mobilisées dans cette estimation sont donc le taux de délinquance et le trafic. Ces deux variables sont les deux facteurs explicatifs de l'inefficience qui ressortent des estimations auxquelles nous avons procédé. Le taux de délinquance régional, d'une valeur moyenne de 5,3 % sur la période, atteint 9,4 % en PACA et 6,9 % en Languedoc-Roussillon, régions dont l'inefficience brute est assez élevée. Il est vrai que face à la délinquance, l'opérateur ferroviaire a significativement augmenté le personnel en gare chargé de l'accueil-filtrage sur les quais.

[8] En effet, les résultats que nous obtenons indiquent qu'à nombre de trains.kilomètres égal, les coûts seront plus faibles sur un réseau étendu que sur un réseau de petite taille, très sollicité (toutes choses égales par ailleurs).

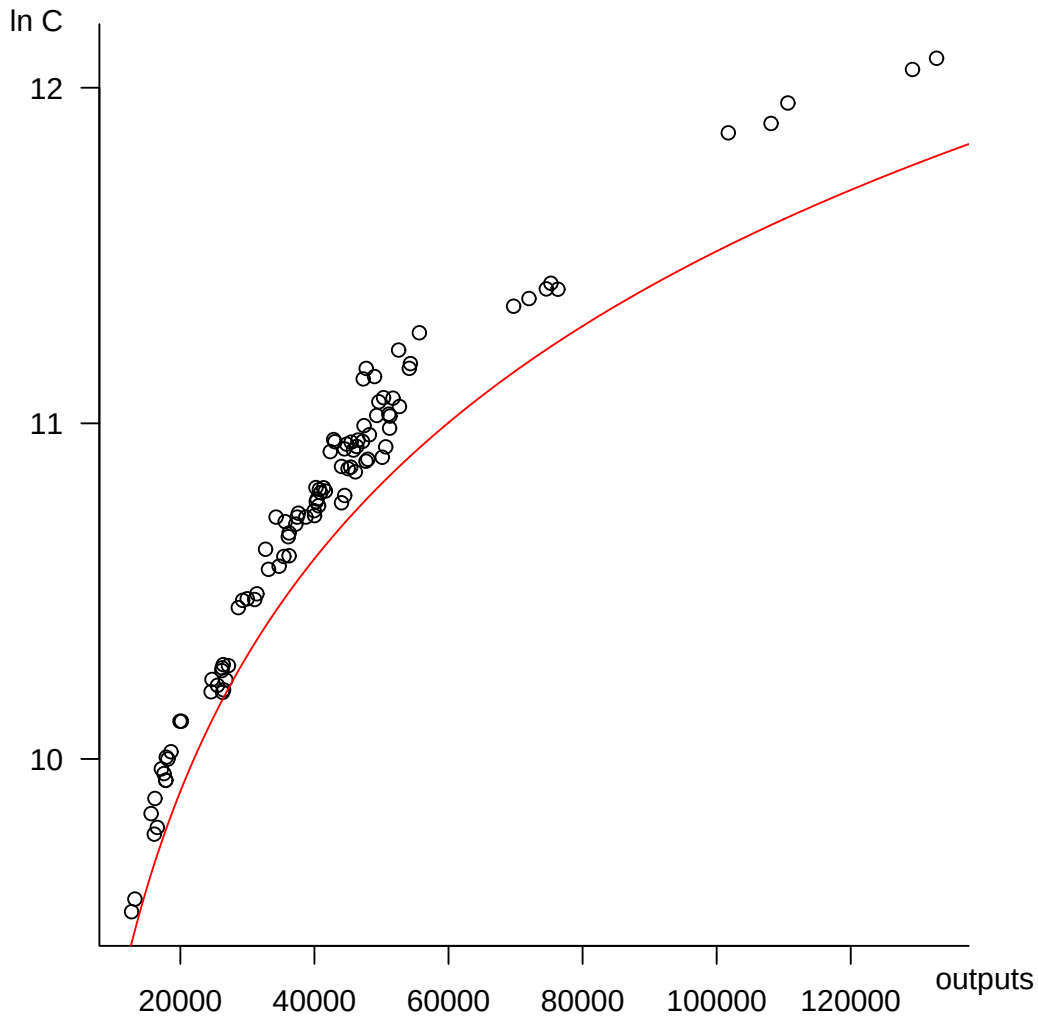

FIG. 6.4 – Frontière stochastique de coût : $\ln C$ en fonction de LL et TrK agrégés.

Néanmoins, lorsque l'on s'attarde sur la décomposition des charges dans ces régions, on constate également d'autres surcoûts qui ne sauraient être imputables à la délinquance (charges de conduite, par exemple). La variable du taux de délinquance explique donc aussi une part d'inefficience endogène à l'entreprise.

L'impact du trafic sur l'inefficience est justifié par plusieurs raisons, notamment par les déséconomies d'échelle que nous avons évaluées dans la première partie. Mais le trafic peut également être à l'origine de surcoûts en terme de charges d'entretien plus élevées pour les gares ou pour du matériel roulant plus lourd...

Par ailleurs, nous avons testé l'impact du capital, supposé améliorer l'efficience, mais le coefficient associé n'est pas du signe attendu [9]. Dans Lévêque (2004) [109] nous avons testé, sur un échantillon restreint aux années 1997 et 1998, l'impact de la vitesse de circulation et de la mise en œuvre de la phase d'expérimentation. Mais ces variables ne se sont pas révélées significatives.

L'inefficience

Le résidu de notre estimation comporte le terme d'erreur stochastique v et le terme d'inefficience u. Le coefficient γ indique la part de l'inefficience dans la variance totale :

$$\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$$

Dans notre estimation, γ est significativement différent de 0 ; cela signifie qu'il y a bel et bien présence d'inefficience. Un test du ratio de vraisemblance du terme asymétrique, u, atteste en outre ce résultat.
Appliqué à l'hypothèse H_0 : $\delta_0 = \delta_{\text{délinq}} = \delta_{VK} = \gamma = 0$, la statistique du ratio de vraisemblance suit une loi du χ^2 mixte, à 4 degrés de liberté [10].
On obtient $LR = 52,70 > \chi^2_{\text{mixte};4;0,001} = 17{,}61$.

On remarque également que γ n'est pas significativement différent de 1. Cela signifie que l'essentiel des fluctuations du terme résiduel est dû à l'ineficience (u_i), et que l'erreur stochastique (v_i) est quasi-nulle. Cela indique que notre frontière stochastique n'est pas significativement différente d'une frontière déterministe.

Les scores d'efficience brute $(u_{i,t})$ fournis par le programme permettent de calculer les scores d'efficience nette des variables de contrôle, pour chaque observation. Nous présentons dans le tableau 6.4 les scores d'efficience nette obtenus par la SNCF dans les différentes régions, en 1998. Ces scores sont normés de façon à ce que le meilleur opérateur ait 1, afin de faciliter les comparaisons relatives. Ils varient ainsi de 1

[9] On pourrait penser qu'un accroissement du capital, c'est-à-dire l'investissement en matériel roulant, augmenterait l'efficience technique de nos opérateurs ainsi que l'observent Dalen & Gómez-Lobo (2003) [100] pour les bus norvégiens. Malheureusement il n'en est rien... Cela signifie au mieux que l'investissement n'améliore pas significativement l'efficience et qu'il est corrélé à un autre effet inobservable.

[10] On trouve les valeurs tabulées de cette loi dans Kodde & Palm (1986) [105].

(Centre) à 1,200 (Basse-Normandie), avec une valeur moyenne de 1,106. Ces scores s'interprètent directement en terme de surcoût : un score de 1,200 correspond à un surcoût de 20,0 % dû à l'inefficience de l'opérateur dans cette région.

Régions	Scores d'efficience
Centre	1,000
Nord-Pas-de-Calais	1,017
Poitou-Charentes	1,056
Franche-Comté	1,056
Champagne-Ardennes	1,056
Limousin	1,057
Bretagne	1,063
Rhône-Alpes	1,087
Bourgogne	1,102
Pays de la Loire	1,109
Aquitaine	1,116
Midi-Pyrénées	1,122
Alsace	1,126
Haute-Normandie	1,134
Languedoc-Roussillon	1,144
Auvergne	1,152
PACA	1,155
Picardie	1,166
Lorraine	1,195
Basse-Normandie	1,200
moyenne	1,106

TAB. 6.4 – Scores d'efficience nette normés pour 1998 ($e^{w_{i,1998}}/e^{w_{\text{Centre},1998}}$).

Afin de mieux cerner l'origine de ces scores d'efficience nette, nous proposons de les comparer avec d'autres mesures de la performance. Il s'agit en l'occurence des scores d'efficience brute et du ratio des charges de production unitaires C/TrK. Le tableau 6.5 de la page suivante permet de vérifier que ces mesures de la performance, toutes normées à 1 pour le meilleur opérateur, sont relativement proches[11].

Les charges unitaires de production sont corrélées à hauteur de 71,1 % avec les scores d'efficience coût, nets de l'hétérogénéité. La prise en compte des variables environnementales correspond au passage de l'efficience brute à l'efficience nette, corrélées à hauteur de 67,5 %. On remarque que l'introduction de ces variables réduit effectivement l'inefficience dont le maximum passe de 1,320 (PACA) à 1,200 (Basse-Normandie). Le classement général des opérateurs n'est pas sensiblement modifié selon la mesure retenue. Ainsi que nous nous y attendions, la prise en compte du taux

[11] Pour une comparaison aisée des classements des opérateurs suivant les différentes mesures, voir les tableaux présentés en annexe, page 241, pour l'efficience brute et les charges de production unitaires.

région	efficience nette	efficience brute	C/TrK
Centre	1,000	1,026	1,000
Nord-Pas-de-Calais	1,017	1,124	1,123
Poitou-Charentes	1,056	1,017	1,008
Franche-Comté	1,056	1,029	1,060
Champagne-Ardennes	1,056	1,061	1,063
Limousin	1,057	1,000	1,012
Bretagne	1,063	1,038	1,035
Rhône-Alpes	1,087	1,246	1,172
Bourgogne	1,102	1,077	1,059
Pays de la Loire	1,109	1,101	1,106
Aquitaine	1,116	1,141	1,110
Midi-Pyrénées	1,122	1,124	1,092
Alsace	1,126	1,168	1,219
Haute-Normandie	1,134	1,151	1,203
Languedoc-Roussillon	1,144	1,230	1,231
Auvergne	1,152	1,112	1,104
PACA	1,155	1,320	1,330
Picardie	1,166	1,226	1,211
Lorraine	1,195	1,198	1,166
Basse-Normandie	1,200	1,160	1,202
coefficient de corrélation	1	0,675	0,711

TAB. 6.5 – Comparaison de trois mesures normées de la performance pour 1998.

de délinquance régional entraine une nette amélioration des scores de Languedoc-Roussillon et surtout de PACA (inefficience réduite de moitié). La variable environnementale VK conduit notamment à améliorer l'efficience apparente des régions Nord-Pas-de-Calais et Rhône-Alpes. Inversement, sa prise en compte réduit l'efficience de l'Auvergne et de la Basse-Normandie.

Cette inefficience (ce surcoût inexpliqué) estimée économétriquement doit à présent être discutée en termes économiques.

6.3.2 Interprétation économique

Inefficience et asymétrie d'information

Afin d'interpréter les scores d'efficience calculés, nous reprenons le modèle classique de Laffont & Tirole (1993) [76]. Ces auteurs proposent la fonction de coût suivante :

$$C = (\beta - e)q$$

Rappelons que β désigne le niveau de productivité exogène de l'entreprise ; cette caractéristique peut être améliorée par des investissements. Le régulateur n'observe

pas ce paramètre de productivité. En général, les entreprises profitent de cette asymétrie d'information pour dégager des rentes informationnelles, en annonçant une productivité $\tilde{\beta}$ inférieure à la réalité.
Rappelons aussi que l'effort e fourni par l'entreprise est endogène ; il permet de réduire les coûts. Le régulateur n'observe pas ce facteur ; l'entreprise peut donc fournir un effort sous-optimal. Toutefois, un audit permet au régulateur de connaître les coûts comptables de l'entreprise, plutôt que ses coûts annoncés.

En transposant à notre cas particulier la décomposition de la caractéristique de productivité proposée par Auriol & Laffont (1992) [56] [12], il est possible de décomposer nos scores d'efficience de la façon suivante :

$$e^{w_{i,t}} = \alpha b^t + (1-\alpha)\epsilon^{i,t} + (\tilde{\beta}^{i,t} - \beta^{i,t}) - e_{i,t}$$

Ces scores d'efficience se décomposent donc en quatre termes :

- **Le terme $\boldsymbol{\alpha b^t}$** correspond à la composante de la productivité commune à la SNCF à l'année t. Les comparaisons permettent de filtrer ce terme.
- **Le terme $\boldsymbol{(1-\alpha)\epsilon^{i,t}}$** désigne la composante de la productivité particulière à l'opérateur i à l'année t. En pratique, ce terme rend compte d'améliorations de la productivité régionale, suite à des investissements.
- **Le terme $\boldsymbol{(\tilde{\beta}^{i,t} - \beta^{i,t})}$** représente la rente informationnelle de l'opérateur i à l'année t. Ce terme-là n'apparaît pas dans notre analyse, puisque celle-ci est menée à partir de coûts comptables, et non de coûts « annoncés ». Il n'y a donc pas de rente informationnelle, $\tilde{\beta}^{i,t} = \beta^{i,t}$.
- **Le terme $\boldsymbol{e_{i,t}}$** est l'effort fourni par l'opérateur i à l'année t. Cet effort, qui varie dans le temps, est le résultat des actions quotidiennes de l'organisation. Ses variations peuvent être interprétées en terme de variations (en sens inverse) de l'excédent organisationnel [13]. Il peut s'agir d'optimiser les rotations du personnel, de réduire les circulations à vide ou encore d'introduire l'équipement à agent seul.

Ainsi, les écarts d'efficience entre nos opérateurs peuvent être justifiés de deux manières. Soit il s'agit de différences de productivité particulière ; mais la réglementation imposée dans toute la SNCF laisse peu de marges aux initiatives locales. Soit il s'agit de l'exercice d'un meilleur niveau d'effort, ce qui semble plus probable, compte tenu des possibilités de réduction de l'excédent organisationnel. Afin de mieux caractériser économiquement ces scores d'efficience, nous examinons leur variation chronologique sur la période 1993-1998.

L'apport d'information en dynamique

La figure 6.5 de la page suivante illustre la variabilité chronologique des scores d'efficience nette normés. On distingue plus particulièrement trois groupes d'opérateurs régionaux :

[12] Voir à la section 3.3, page 87 ; on a, pour un modèle statique, $\beta^i = \alpha b + (1-\alpha)\epsilon^i$.
[13] Voir à ce sujet page 53.

- **Ceux qui progressent :** il s'agit essentiellement des opérateurs en régions Rhône-Alpes, Champagne-Ardennes et Centre. Leurs scores d'efficience s'améliorent chaque année. Cette performance dynamique est attribuable soit à la mise en œuvre de plans de productivité, soit à un effort continu. Toutefois, nous n'avons pas connaissance d'investisements particuliers qui aient permis une amélioration de la productivité. Aussi, nous pensons plutôt qu'une réduction particulièrement forte de l'excédent organisationnel a été mise en œuvre dans ces régions.
- **Ceux qui régressent :** il s'agit notamment des opérateurs en régions Basse-Normandie, Languedoc-Roussillon, Lorraine. Leurs scores sont clairement orientés à la hausse, ce qui correspond à une diminution de leur efficience relative. Ce résultat a pour origine un niveau d'effort inférieur à ceux des autres opérateurs.
- **Les autres :** la majorité des opérateurs présentent un niveau d'efficience qui fluctue autour de sa valeur moyenne. Ces variations d'inefficience laissent penser qu'elles résultent de l'exercice d'un niveau d'effort moyen. Il n'y a *a priori* pas lieu de supposer que la productivité ait évolué significativement.

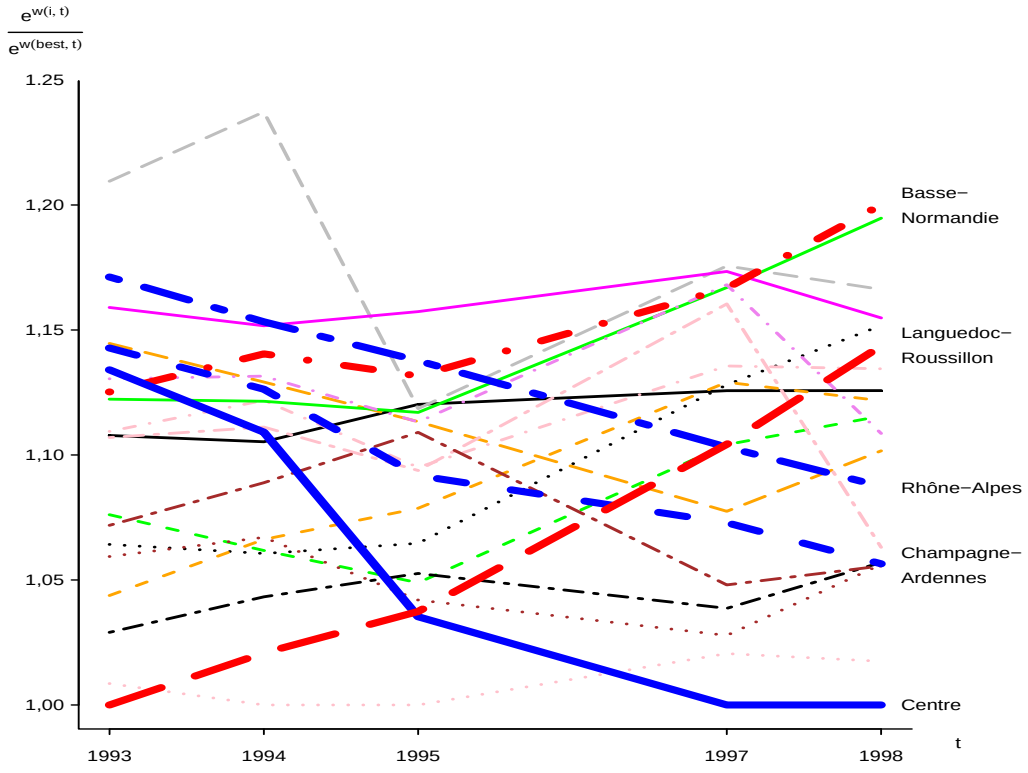

FIG. 6.5 – Variation des scores d'efficience nette normés :
– progression en Rhône-Alpes, Champagne-Ardennes et Centre ;
– régression en Basse-Normandie et Languedoc-Roussillon.

La variabilité observée nous incite à conclure que les niveaux d'efficience calculés reflètent davantage les efforts fournis par les opérateurs, que leur caractéristique de productivité intrinsèque. Ces résultats sont largement représentatifs d'une époque où le secteur était organisé par la SNCF dans un contexte relativement peu marchand. La régionalisation intervenue en 2002 devrait avoir sensiblement modifié certaines caractéristiques de nos résultats. Il est notamment fort probable que la variance des coûts, et donc celle de l'efficience associée, ait augmenté. En effet, le même exercice conduirait dorénavant à travailler non plus avec des coûts comptables, mais avec des coûts contractuellement négociés. Ceux-ci intègrent notamment une rémunération de l'opérateur (variable suivant les régions), et des rentes informationnelles. En effet, si l'audit initial a permis de fixer le montant de chaque forfait régional, la progression généreusement indicée de ce dernier et l'augmentation de la production ont permis aux opérateurs de dégager des rentes informationnelles. Tout cela concourt à renforcer la pertinence de telles mesures et, plus généralement, d'un mécanisme comparatif.

Conclusion : une opportunité de mise en œuvre

L'analyse développée dans ce chapitre nous a conduits à révéler des surcoûts inexpliqués, que nous attribuons à un effort sous-optimal. Cette inefficience coût, qui atteint jusqu'à 20 %, incite fortement à mettre en œuvre le mécanisme suggéré à la section 5.2. En effet, un système de *price cap*, dont les paramètres seraient fixés par comparaison, inciterait chaque opérateur à réduire son inefficience en fournissant un effort plus élevé.

En 1998, l'inefficience moyenne, estimée à plus de 10 %, correspond à plus de 115 M€ 2004. À l'heure où les besoins d'investissement en infrastructure se font sentir, et où les TER périurbains se développent, il faut donner au système ferroviaire français les moyens d'être plus productif. Mais la participation des voyageurs plafonne péniblement à 30 % du coût total des TER, tandis que les contraintes budgétaires du financement public sont saturées. Il faudra donc rapidement se résoudre à mettre à contribution les acteurs mêmes du système, à commencer par le principal d'entre eux.

Bien sûr, il faudrait aussi améliorer la productivité des TER en réaffectant de nombreux services de transport sur des lignes qui génèrent davantage de surplus. Mais il serait absurde, économiquement, de renoncer aux bénéfices collectifs qu'un mécanisme comparatif permettrait de retirer.

Conclusion : De la théorie à la pratique

Cette dernière partie de notre travail, consacrée aux mises en œuvre ferroviaires des mécanismes comparatifs, nous permet de conclure les résultats suivants :

- Les mécanismes comparatifs sont bien adaptés pour réguler divers marchés ferroviaires.
- Sur les marchés des TER français, on observe de l'inefficience-coût dans les différentes régions.
- Appliqué par l'ensemble des régulateurs régionaux, un mécanisme comparatif adéquat doit inciter les activités régionales de la SNCF à améliorer leur efficience.

À partir d'un panel de données sur les TER français, un traitement approprié de l'échantillon nous a conduit à identifier des termes économétriques d'inefficience, et à les interpréter économiquement. Ces derniers résultats de notre travail permettent de relier d'un point de vue *pratique* l'inefficience au sens de la théorie économétrique des frontières et l'asymétrie d'information de la nouvelle économie de la réglementation. D'un point de vue *théorique*, Laffont (1995) [33] souligne en effet que l'introduction de l'asymétrie d'information par la nouvelle économie de la réglementation a permis « de modéliser les inefficacités si longtemps occultées par les termes d'erreurs comme ceux postulés dans l'économétrie des frontières ».

En pratique, les comparaisons de ce type permettent donc de réduire l'asymétrie d'information entre les entreprises et le régulateur. Elles nécessitent néanmoins un minimum de transparence dans la production des entreprises ; celles-ci doivent au moins disposer d'une comptabilité analytique. À charge pour le régulateur d'en intégrer les données au mécanisme comparatif et les résultats des comparaisons à un mécanisme incitatif. C'est à cette étape que le praticien doit prendre des initiatives pertinentes. En effet, les modèles normatifs idéaux de l'économie publique ne peuvent pas être directement mis en œuvre sur les marchés que nous connaissons. En pratique, d'autres contraintes que celles étudiées, notamment les contraintes politiques, pèsent sur la régulation.

Toutefois, la variété des mises en œuvre de mécanismes comparatifs illustre la souplesse de cette forme de régulation. Dans les différentes industries de réseau qu'ils contribuent à réguler, les mécanismes comparatifs sont parvenus à intégrer

les contraintes propres à chaque marché. Dans notre cas, la contrainte essentielle, dont il faut tenir compte, est de nature politique. Il s'agit du maintien du monopole légal de la SNCF, dont nous avons montré qu'un mécanisme comparatif adéquat permet de s'affranchir.

Nous concluons ce travail en détaillant, le plus précisément possible, tel un mode d'emploi, le mécanisme de régulation comparative que nous proposons pour les TER.

Conclusion générale : Réguler les TER par comparaison : mode d'emploi

> « Il y a deux méthodes possibles pour réformer. La première découle de la réalisation heureuse d'une majorité politique qui exclut ces groupes d'intérêt et qui a suffisamment de pouvoir de marchandage pour imposer des pertes de bien-être (au moins momentanées) aux groupes d'intérêt et mettre en place une économie plus efficace. [...]
>
> Si la majorité politique contient les groupes d'intérêt opposés à la réforme, il faut organiser un échange favorable qui maintient ou accroît les avantages des groupes d'intérêt mais réalise la réforme souhaitée. [...]
>
> Mais que faire d'une entreprise inefficace qui aurait été capturée par ses employés qui, de plus, ont réussi à prendre le contrôle du recrutement qui s'effectue largement au bénéfice de leurs enfants ? Il faudrait compenser ceux-ci non seulement pour leurs pertes personnelles mais pour la valeur actualisée des pertes de toutes leurs dynasties. La démarche suggérée par certains est alors d'effectuer un petit pas qui recueille une majorité politique et transforme l'électorat qui devient alors favorable à un autre petit pas. »
>
> Jean-Jacques Laffont (1999),
> « Étapes vers un État moderne : une analyse économique »,
> *rapport au Conseil d'Analyse Économique.*

À l'issue de ce travail de recherche, il paraît nécessaire de conclure notre « proposition d'application de la concurrence par comparaison » de façon pratique. En attendant la libéralisation des marchés ferroviaires régionaux, cette proposition pourrait être un des « petits pas » évoqués par Jean-Jacques Laffont.
Nous avons montré que la concurrence par comparaison pouvait améliorer la régulation des chemins de fer, notamment des TER français. Concernant ce cas particulier, nous avons calculé que la seule inefficience interne à l'opérateur historique engendrait un surcoût pour les régions de plus de 100 M€ par an. Dans cette conclusion générale, nous détaillons, comme dans un mode d'emploi, le dispositif de régulation

par comparaison proposé. Nous rappelons d'abord le contexte dans lequel s'inscrit notre proposition, ainsi que ses objectifs. Dans un deuxième temps, nous explicitons la technique de comparaison à mettre en œuvre. Enfin, nous indiquons comment intégrer le résultat de ces comparaisons à la régulation des TER.

Contexte et objectifs

Rappelons d'abord, au titre d'avertissement, la condition nécessaire à l'efficience du mécanisme proposé : les régulateurs régionaux doivent s'entendre, afin de mettre en commun l'information dont ils disposent. Notre proposition s'intègre au cadre des conventions de délégation de service public, conclues entre chaque région et la SNCF. Plus précisément, il s'agit d'améliorer l'actuelle régulation des coûts, en tenant compte du monopole légal de la SNCF. En pratique, notre travail consiste à proposer le mécanisme incitatif qui fixe le forfait de charges, compensé à chaque opérateur. Ce mécanisme de *price cap* (prix plafond) est défini par :

- un plafond initial ;
- des gains de productivité annuels.

L'équation financière qui régit les relations commerciales consommateurs - SNCF est de la forme suivante :

$$C = C_{\text{tot}} - C_2 = C_{1\text{pub}} + R$$

où C est le coût de revient à la collectivité et aux voyageurs du service TER, net des charges exogènes C_2.
$C_{1\text{pub}}$ désigne la contribution publique [1] et R, les recettes commerciales.

Les actuelles conventions régions-SNCF, établies sur une base auditée, fixent d'une part des prévisions de recettes et, d'autre part, une évolution très généreuse du forfait $C_{1\text{pub}}$. La négociation des recettes prévisionnelles est propre à chaque région, compte tenu de la dynamique observée localement. Nous proposons donc ici une relation permettant d'établir l'évolution du forfait $C_{1\text{pub}}$ dans les prochaines conventions. Le « petit pas » consisterait donc à passer d'un régime de *cost plus*, déguisé en *price cap*, à un mécanisme de prix plafond, véritablement incitatif.

Technique de comparaison

Ainsi que nous l'avons justifié dans la troisième partie, il est tout-à-fait pertinent de comparer la performance des activités régionales de la SNCF, sur la base des coûts $C = C_{1\text{pub}} + R$. Afin de fixer les idées, nous prenons l'exemple du renouvellement d'une convention qui s'achève en 2006 [2]. De façon générique, l'année d'échéance est notée 0 :

[1] Ce terme correspond au $S - C_2$ du modèle établi au chapitre 6 ; nous le rebaptisons ici, afin de clarifier l'exposé.

[2] Toutes les conventions débutent, parfois rétroactivement, en 2002. Les plus courtes d'entre elles arrivent à échéance fin 2006.

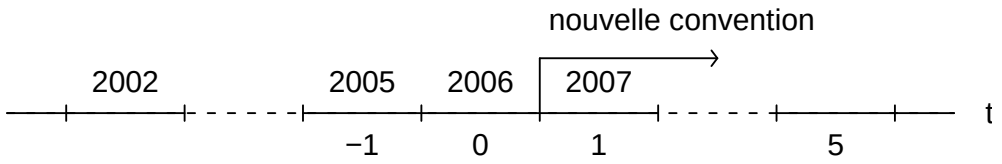

FIG. CG.1 – Déroulement chronologique.

Nous proposons de reprendre le travail de comparaison réalisé dans le chapitre 6. Nous conseillons donc au régulateur commun de construire une nouvelle frontière de coût, à partir des données des années 2002 à -1 (c'est-à-dire 2005 dans notre exemple)[3]. L'estimation de la frontière requiert :

- **Les coûts observés :** Les termes $C_{i,t} = C_{1\text{pub},i,t} + R_{i,t}$ sont évalués par la somme des recettes et des contributions versées forfaitairement par les collectivités locales. Cette dernière part inclut, essentiellement, la contribution versée par la région. Il convient néanmoins d'y ajouter les contributions versées par les autres collectivités locales (communes, agglomérations, départements, *etc.*). Rappelons également qu'il convient de soustraire à ce terme les pénalités perçues par la région pour non réalisation du service.
- **La production réalisée :** Face au terme de coût précédent, il faut mettre en regard le service réalisé. Il s'agit surtout du nombre de trains.kilomètres ayant circulé en service commercial, ainsi que nous l'avons justifié au chapitre 5. De même que nous l'avons fait, il est possible d'ajouter le quasi-produit qu'est la longueur de lignes du réseau exploité.
- **Les variables environnementales :** Dans notre estimation, nous avons retenu le trafic et le taux de délinquance. Nous avons toutefois précisé que la liste des variables à inclure pouvait être adaptée à partir des propositions des opérateurs. Concrètement, il appartient à la SNCF de justifier pourquoi telle variable environnementale biaise son efficience et doit être prise en compte. Lorsqu'il s'agit de variables endogènes, l'opérateur peut fournir les données à condition que le régulateur commun puisse les vérifier. Si l'impact des variables est significatif et intervient dans le sens attendu, le régulateur ne peut s'opposer à leur prise en compte.

Une fois la frontière estimée, le régulateur peut calculer des scores d'efficience nette, pour la dernière année dont il dispose (année -1). Ces scores sont valides s'ils ne font pas apparaître d'hétéroscédasticité, c'est-à-dire si une région ne devient pas subitement bien meilleure ou bien moins efficiente, lors du calcul de l'efficience nette. Les scores normés obtenus sont interprétables, quantitativement, comme un surcoût injustifié que l'opérateur doit réduire durant la prochaine convention.

[3] Pour une nouvelle convention en vigueur à partir de 2007, cela correspond à 80 observations. Cela suffit pour estimer une frontière solide. Pour les conventions qui échoient ultérieurement, une frontière avec davantage d'observations pourra être établie.

Mécanisme de régulation

L'introduction du résultat des comparaisons dans le mécanisme de régulation permet d'inciter de façon pertinente l'opérateur à réduire ses coûts, là où ils sont trop élevés. Quelle sera la distribution de ces futurs scores d'efficience ? Il est difficile de répondre à cette question, bien que deux éléments de réponse (de sens contraire) puissent être avancés :

- Initialement, l'efficience brute présentera sans doute davantage de variance que celle estimée au chapitre 6. Cela est dû au fait que les coûts ultérieurs à 2002 ne seront plus issus de la comptabilité, mais négocié contractuellement avec plus ou moins de succès [14]. En plus des disparités d'effort que nous avons observées sur 1993-1998, vient s'ajouter le développement de rentes informationnelles. Ces rentes sont apparues de façons inégales, selon les indexations de forfait et les nouveaux services négociés.
- Finalement, avec la prise en compte de variables environnementales proposées par la SNCF, l'efficience nette pourrait bien ne pas être trop dispersée.

A priori, au vu des variations d'effort mises en évidence à la fin du dernier chapitre, il est raisonnable d'exiger de chaque opérateur qu'il réduise ses surcoûts au terme de cinq années. Ainsi, à moyen terme, chaque opérateur devra avoir renoncé à sa rente informationnelle et mis en œuvre un niveau d'effort optimal, relativement aux autres.

Nous reprenons la relation générale d'indexation *RPI-X* des mécanismes *price cap* britanniques. Rappelons que *RPI* (*retail price index*) est le taux d'inflation et que X désigne les gains de productivité. En imposant une réduction régulière des coûts sur les cinq premières années de la convention, on a alors :

$$C_{i,t} = C_{i,t-1} \times \frac{RPI_t}{RPI_{t-1}} \times (1 - X_i)$$

La condition de résorption des surcoûts, après cinq ans, s'écrit :

$$C_{i,5} = C_{i,0} \times \frac{RPI_5}{RPI_0} \times \frac{e^{w_{\text{best},-1}}}{e^{w_{i,-1}}}$$

où le dernier terme correspond à l'inverse du score d'efficience nette normé (par rapport au meilleur - *best* - opérateur), estimé à l'étape précédente. Il s'agit donc bien, compte tenu des gisements de productivité à exploiter, d'aligner les performances des opérateurs sur celle du meilleur d'entre eux, plutôt que sur la performance moyenne de l'industrie.

Le gain de productivité annuel requis pour l'opérateur i s'écrit donc :

$$X_i = 1 - \sqrt[5]{\frac{e^{w_{\text{best},-1}}}{e^{w_{i,-1}}}}$$

Numériquement, la table CG.1 fournit la valeur des gains de productivité annuels en fonction des scores normés d'efficience nette.

[14] En travaillant ainsi à partir de données ayant davantage de signification économique et en intégrant les prix unitaires du travail, l'estimation des coûts se trouvera, de plus, renforcée.

Scores d'efficience	Gains annuels de productivité
1,00	0,00 %
1,05	0,97 %
1,10	1,89 %
1,15	2,76 %
1,20	3,58 %
1,25	4,36 %
1,30	5,11 %

TAB. CG.1 – Gains de productivité annuels en fonction du score d'efficience nette normé.

La contribution publique versée à l'opérateur i pour l'année $t \in [1;5]$, compte tenu des recettes prévisionnelles $R_{i,t}$, s'écrit donc :

$$C_{1\text{pub},i,t} = C_{i,0} \times \frac{RPI_t}{RPI_0} \times \sqrt[5]{\frac{e^{w_{\text{best},-1}}}{e^{w_{i,-1}}}}^{t} - R_{i,t}$$

À première vue, le mécanisme que nous proposons ressemble au modèle théorique de Shleifer. Cependant, le niveau de coût atteint à terme est celui du meilleur opérateur à l'année -1 ; il n'y a aucune raison *a priori* pour qu'il s'agisse de l'optimum. Ensuite, il est prévisible que les opérateurs les moins efficients tentent de négocier un mécanisme moins contraignant. L'opérateur doit alors comprendre que tout relâchement d'effort est assimilable à un comportement opportuniste. Un tel comportement nuit à sa réputation et sera sanctionné à l'issue de la convention par la perte de marchés soumis à appel d'offre. La libéralisation prochaine des marchés offre ainsi aux régulateurs régionaux un moyen de pression sur la SNCF dont ils peuvent déjà faire usage. Toutefois, il peut être envisagé qu'un opérateur justifie exogènement sa mauvaise performance ; il faudrait dans ce cas en tenir compte. C'est tout l'intérêt de ce type de mécanismes incitatifs : ils renversent la charge de la preuve du régulateur à l'opérateur, contraint de justifier - à défaut de la reconnaître - son inefficience.

Au delà...

Le « petit pas » que nous proposons pour la première révision des conventions régions-SNCF doit être suivi par d'autres. Ultérieurement, la libéralisation du secteur devrait permettre de relâcher la contrainte politique du monopole de la SNCF sur ces marchés. Des appels d'offre seront lancés pour l'exploitation de sous-réseaux, et la concurrence *pour* le marché contribuera à améliorer davantage l'efficience des opérateurs. Ainsi que nous l'avons souligné au chapitre 2, cette libéralisation ne dispense pas, au contraire, de réguler les marchés. Avec l'allotissement, les comparaisons devront porter sur des performances désagrégées. Cela nécessitera de mieux corriger l'hétérogénéité environnementale qui devient importante, à l'échelle des sous-réseaux.

Mais l'apparition de nouveaux entrants sur le marché va bouleverser la donne avec laquelle nous sommes actuellement contraints de travailler. Les futurs opérateurs, *a priori* plus efficients, vont agir comme des révélateurs de la productivité β de la SNCF. Leur entrée sur le marché introduira de l'hétérogénéité interne, et donc, de la sélection adverse. Si les nouveaux entrants pourront en profiter pour dégager des rentes informationnelles, la concurrence par comparaison devrait réduire l'ampleur de celles-ci, sur le modèle proposé par Auriol et Laffont.

Toutefois, la régulation des chemins de fer concerne de nombreux marchés différents, dont certains ne devraient pas être libéralisés. L'entretien et la maintenance de l'infrastructure par la SNCF en tant que gestionnaire délégué devraient rapidement être régulés, de façon plus contraignante, par des comparaisons. Reste la question des services voyageurs à longue distance, pour lesquels un mécanisme incitatif reste à proposer.

Ce travail de recherche sur la régulation des chemins de fer et la concurrence par comparaison nous a conduits à proposer la mise en œuvre pratique d'un mécanisme comparatif. À terme, l'augmentation des efforts et la réduction de l'inefficience doivent permettre d'économiser plus de 100 M€ par an aux régions. Alors qu'il y a une occasion réellement favorable d'appliquer à la régulation certains résultats de la théorie des incitations, nous espérons que cette thèse contribuera à améliorer la régulation des chemins de fer.

Bibliographie

Bibliographie de la partie 1

[1] Armstrong M., Cowan S. & Vickers J. (1994), *Regulatory Reform*, Cambridge MA : MIT Press, 392 p.

[2] Baumol W., Panzar J. & Willig R. (1982), *Contestable Markets and The Theory of Industry Structure*, New York : Harcourt Brace Jovanovitch, 510 p.

[3] Baumstark L. & Bonnafous A. (1998), « France », *Les redevances d'usage des infrastructures ferroviaires* / ed. par CEMT, Paris : Éditions de l'OCDE, p. 51-108.

[4] Baumstark L. & Bonnafous A. (2000), « La relecture théorique de Jules Dupuit par Maurice Allais face à la question du service public », *Les traditions économiques françaises 1848-1939* / ed. par Dockès P., Frobert L., Klotz G. *et al.*, Paris : CNRS Éditions, p. 397-412.

[5] Bonnafous A. (1989), *Le siècle des ténèbres de l'économie*, Paris : Economica, 184 p.

[6] Bonnafous A. (1992), « Transport et environnement. Comment valoriser et maîtriser les effets externes ? », *Économie et Statistique*, n° 258-259, p. 121-128.

[7] Bouf D., Crozet Y. & Lévêque J. (2005), « Vertical separation, disputes resolution and competition in railway industry », *9th International Conference on Competition and Ownership in Land Passenger Transport*, Lisbonne, 5-9 septembre, 13 p. Disponible sur : http://www.let.fr/fr/annuaire/auteurs /affichage.php?id=96

[8] Braeutigam R.R. (1989), « Optimal Policies for Natural Monopolies », *Handbook of Industrial Organization*, vol. 2 / ed. par Scmalensee R. & Willig R.D., Amsterdam : North-Holland, p. 1290-1347.

[9] Braeutigam R.R. (1999), « Learning about Transport Costs », *Essays in transportation economics and policy : A handbook in honor of John R. Meyer* / ed. par Gómez-Ibáñez J.A., Tye W.B. & Winston C., Washington, D.C. : Brookings Institution, p. 57-97.

[10] Cantos Sánchez P. (2001), « Vertical relationships for the European railway industry », *Transport Policy*, vol. 8, n° 2, p. 77-83.

[11] Caron F. (1997), *Histoire des chemins de fer en France (1740-1883)*, Paris : Fayard, 700 p.

[12] Caves D.W., Christensen L.R. & Tretheway M.W. (1984), « Economies of density versus economies of scale : why trunk and local service airline costs differ », *Rand Journal of Economics*, vol. 15, n° 4, p. 471-489.

[13] Caves D.W., Christensen L.R., Tretheway M.W. & Windle R.J. (1985), « Network effects and the measurement of returns to scale and density for U.S. railroads », *Analytical Studies in Transport Economics* / ed. par Daugherty A., Cambridge : Cambridge University Press, p. 97-120.

[14] Creti A. & Perrot A. (1997), « Les entreprises en réseau », *Réglementation et Concurrence* / ed. par Perrot A., Paris : Economica, p. 49-74.

[15] Curien N. (2000), *Économie des réseaux*, Paris : La Découverte, 121 p.

[16] Curien N. & Dupuy G. (1996), *Réseaux de communication. Marchés et territoires*, Paris : Presses de l'école nationale des Ponts et chaussées, 176 p.

[17] Cyert R. & March J. (1963), *A Behavioral Theory of the Firm*, Cambridge MA : Blackwell, 252 p.

[18] Demsetz H. (1968), « Why regulate utilities ? », *Journal of Law and Economics*, vol. 11, n° 1, p. 55-65.

[19] Dodgson J. (1997), « La situation actuelle des chemins de fer britanniques et les enseignements de la privatisation de British Rail », *Revue d'histoire des chemins de fer*, n° 16-17, p. 303-313.

[20] Dupuy G. (1999), *La Dépendance automobile : symptômes, analyse, diagnostic, traitements*, Paris : Economica, 160 p.

[21] Economides N. (1996), « The Economics of networks », *International Journal of Industrial Organization*, vol. 14, n° 6, p. 673-699.

[22] Estache A., Guasch J.-L. & Trujillo L. (2003), « Price Caps, Efficiency Payoffs and Infrastructure Contracts Renegotiation in Latin America », *Policy Research Working Paper Series*, n° 3129, Banque Mondiale. Disponible sur : http://wdsbeta.worldbank.org/external/default/WDSContentServer/IW3P/IB/2003/09/23/000094946_03091104060249/Rendered/PDF/multi0page.pdf

[23] Fremdling R. (1997), « Discussions prussiennes et hollandaises, solutions pour la réglementation des chemins de fer et leur nationalisation au XIXe siècle », *Revue d'histoire des chemins de fer*, n° 16-17, p. 230-264.

[24] Gabszewicz J.J., Shaked A., Sutton J. & Thisse J.-F. (1986), « Segmenting the Market : The Monopolist's Optimal Product Mix », *Journal of Economic Theory*, vol. 39, n° 2, p. 273-289.

[25] Garcia O. (2001), *La réforme ferroviaire en France : une évaluation comptable*, Mémoire pour le DEA d'Économie des Transports de l'Université Lyon 2, 123 p.

[26] Gasmi F., Laffont J.-J. & Sharkey W.W. (2002), « The natural monopoly test reconsidered : an engineering process-based approach to empirical analysis in telecommunications », *International Journal of Industrial Organization*, vol. 20, n° 4, p. 435-459.

[27] Gauthier-Lescop L. & Lévêque J. (2005), « De la régionalisation à la concurrence régulée. Analyse économique et juridique de la future organisation du transport ferroviaire régional de voyageurs », *Politique et management public*, vol. 23, n° 4 (*à paraître*).

[28] Guasch J.-L., Laffont J.-J. & Straub S. (2003), « Renegotiation of Concession Contracts in Latin America », *Policy Research Working Paper Series*, n° 3011, Banque Mondiale.

[29] Hart O.D. (1983), « The market mechanism as an incentive scheme », *The Bell Journal of Economics*, vol. 10, n° 1, p. 366-382.

[30] Hausman J.A. & Taylor W.E. (1981), « Panel Data and Unobservable Individual Effects », *Econometrica*, vol. 49, n° 6, p. 1377-1398.

[31] Ida T. & Suda M. (2004), « The cost structure of the Japanese railway industry : The economies of network density and of scope and the cost gap between Japan's regional railways after privatization », *International Journal of Transport Economics*, vol. 31, n° 1, p. 23-37.

[32] Katz M.L. & Shapiro C. (1985), « Network Externalities, Competition, and Compatibility », *American Economic Review*, vol. 75, n° 3, p. 424-440.

[33] Laffont J.-J. (1995), « La nouvelle économie de la réglementation dix ans après », *Revue d'Économie Industrielle*, n° spécial : Développements récents, p. 331-366.

[34] Laffont J.-J. & Tirole J. (1987), « Auctioning Incentive Contracts », *Journal of Political Economy*, vol. 95, n° 5, p. 921-937.

[35] Laffont J.-J. & Tirole J. (1988), « Repeated auctions of incentives contracts, investments and bidding parity with an application to takeovers », *Rand Journal of Economics*, vol. 19, n° 4, p. 516-537.

[36] Leibenstein H. (1966), « Allocative Efficiency vs. "X-Efficiency" », *American Economic Review*, vol. 56, n° 3, p. 392-415.

[37] Lévêque F. (1998), *Économie de la réglementation*, Paris : La Découverte, 125 p.

[38] Merger M. (1997), « Les conventions de 1885 en Italie : un exemple à ne pas suivre ? », *Revue d'histoire des chemins de fer*, n° 16-17, p. 190-213.

[39] Michard A.-S. (2000), « L'excédent organisationnel », *L'économie industrielle des stratégies d'entreprises* / ed. par Chevalier J.-M., Paris : Montchrestien, p. 93-108.

[40] Ortunez Goicolea P.P. (1997), « Le processus de nationalisation des chemins de fer espagnols (1913-1941) », *Revue d'histoire des chemins de fer*, n° 16-17, p. 165-189.

[41] Page W.H. & Lopatka J.E. (2000), « Network Eternalities », *Encyclopaedia of Law and Economics*, vol. 1, *The History and Methodology of Law and Economics* / ed. par Bouckaert B. & de Geest G., Cheltenham : Edward Elgar, p. 952-980.

[42] Pels E. & Rietveld P. (2000), « Cost Fuctions in Tranport », *Handbook of Transport Modelling* / ed. par Hensher D.A. & Button K.J., Oxford : Pergamon, p. 321-333.

[43] Peltzman S. (1976), « Toward a More General Theory of Regulation », *Journal of Law and Economics*, vol. 19, n° 2, p. 211-240.

[44] Posner R.A. (1974), « Theories of economic regulation », *The Bell Journal of Economics and Management Science*, vol. 5, n° 2, p. 335-358.

[45] Ribeill G. (1993), *La révolution ferroviaire : La formation des compagnies de chemins de fer en France (1823-1870)*, Paris : Belin, 480 p.

[46] Ribeill G. (1997), « Le principe du « libre parcours » sur les premiers chemins de fer concédés français. Fondements théoriques et obstacles pratiques », *Revue d'histoire des chemins de fer*, n° 16-17, p. 29-46.

[47] Riordan M.H. & Sappington D. (1987), « Awarding Monopoly Franchises », *American Economic Review*, vol. 77, n° 3, p. 375-387.

[48] Seabright P. (2003), « The Economics of Passenger Rail Transport : A Survey », *IDEI Report*, 122 p. Disponible sur : http://idei.fr/doc/wp/2003/rapport_db_1.pdf

[49] Sevestre P. (2002), *Économétrie des données de panel*, Paris : Dunod, 211 p.

[50] Stigler G. (1971), « The theory of economic regulation », *The Bell Journal of Economics and Management Science*, vol. 2, n° 1, p. 3-21.

[51] Toutain J.-C. (1967), « Les transports en France de 1830 à 1965 », *Cahiers de l'ISEA*, Histoire quantitative de l'économie française, n° 9, p. 1-306.

[52] Varian H.R. (1992), *Introduction à la microéconomie*, Bruxelles : De Boeck Université, 509 p.

[53] Vickers J. & Yarrow G. (1988), *Privatization : An Economic Analysis*, Cambridge MA : MIT Press, 454 p.

[54] de Villemeur E., Ivaldi M. & Pouyet J. (2003), « Entry in the passenger rail industry : A theoretical investigation », *IDEI Report*, 61 p. Disponible sur : http://idei.fr/doc/wp/2003/rapport2_db_2.pdf

[55] Wauthy X. (1996), « Quality choice in models of vertical differentiation », *The Journal of Industrial Economics*, vol. 44, n° 3, p. 345-353.

Bibliographie de la partie 2

[56] Auriol E. & Laffont J.-J. (1992), « Regulation by Duopoly », *Journal of Economics & Management Strategy*, vol. 1, n° 3, p. 507-533.

[57] Auriol E. (1993), « Monopole ou duopole : l'effet de comparaison », *Annales d'économie et de statistique*, n° 31, p. 1-31.

[58] Auriol E. (1998), « Deregulation and quality », *International Journal of Industrial Organization*, vol. 16, n° 2, p. 169-194.

[59] Auriol E. (2000), « Concurrence par comparaison : Un point de vue normatif », *Revue économique*, vol. 51, n° 3, p. 621-634.

[60] Baiman S. & Demski J.S. (1980), « Economically Optimal Performance Evaluation and Control Systems », *Journal of Accounting Research*, vol. 18, Supplément, p. 184-220.

[61] Baron D.P. & Myerson R.B. (1982), « Regulating a Monopolist with Unknown Costs », *Econometrica*, vol. 50, n° 4, p. 911-930.

[62] Bivand R. & Szymanski S. (1997), « Spatial dependence through local yardstick competition : theory and testing », *Economics Letters*, vol. 55, n° 2, p. 257-265.

[63] Boyer M. & Laffont J.-J. (2003), « Competition and the reform of incentive schemes in the regulated sector », *Journal of Public Economics*, vol. 87, n° 10, p. 2369-2396.

[64] Choné P. & Lesur R. (2001), « A note on yardstick competition under adverse selection », *INSEE papers*, 12 p. Disponible sur : http://www.crest.fr/pageperso/lei/chone/yardjre.pdf

[65] Contreras E.M. & Rickman N. (2004), « Regulation by duopoly under political constraints », *Royal Economic Society Annual Conference 2004*, Swansea, 5-7 avril, 33 p. Disponible sur : http://repec.org/res2004/MendozaRickman.pdf

[66] Crémer J. & McLean R.P. (1988), « Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions », *Econometrica*, vol. 56, n° 6, p. 1247-1257.

[67] Dalen D.M. (1998), « Yardstick Competition and Investment Incentives », *Journal of Economics & Management Strategy*, vol. 7, n° 1, p. 105-126.

[68] Demski J. S. & Sappington D. (1984), « Optimal Incentive Contracts with Multiple Agents », *Journal of Economic Theory*, vol. 33, n° 1, p. 152-171.

[69] Faure-Grimaud A. & Reiche S. (2003), « Dynamic Yardstick Regulation », *STICERD Theoretical Economics Paper Series*, n° 459, 15 p. Disponible sur : http://sticerd.lse.ac.uk/dps/te/te459.pdf

[70] Green J.R. & Stokey N.L. (1983), « A Comparison of Tournaments and Contracts », *Journal of Political Economy*, vol. 91, n° 3, p. 349-364.

[71] Holmström B. (1982), « Moral hazard in teams », *The Bell Journal of Economics*, vol. 13, n° 2, p. 324-340.

[72] Ivaldi M., Jullien B., Rey P., Seabright P. & Tirole J. (2003), « The Economics of Tacit Collusion », *IDEI Report*, 75 p. Disponible sur : http://idei.fr/doc/wp/2003/tacit_collusion.pdf

[73] Laffont J.-J. & Martimort D. (2000), « Mechanism design with collusion and correlation », *Econometrica*, vol. 68, n° 2, p. 309-342.

[74] Laffont J.-J. & Martimort D. (2002), *The Theory of Incentives - The Principal-Agent Model*, Princeton NJ : Princeton University Press, 421 p.

[75] Laffont J.-J. & Tirole J. (1986), « Using Cost Observation to Regulate Firms », *Journal of Political Economy*, vol. 94, n° 3, p. 614-640.

[76] Laffont J.-J. & Tirole J. (1993), *The Theory of Incentives in Procurement and Regulation*, Cambridge MA : MIT Press, 705 p.

[77] Lazear E.P. & Rosen S. (1981), « Rank-Order Tournaments as Optimum Labor Contracts », *Journal of Political Economy*, vol. 89, n° 5, p. 841-864.

[78] Linnemer L. & Souam S. (1997), « Cartels et collusion tacite », *Réglementation et concurrence* / ed. par Perrot A., Paris : Economica, p. 75-98.

[79] McAfee R.P. & Reny P.J. (1992), « Correlated Information and Mechanism Design », *Econometrica*, vol. 60, n° 2, p. 395-421.

[80] Meyer M.A. & Vickers J. (1997), « Performance Comparisons and Dynamic Incentives », *Journal of Political Economy*, vol. 105, n° 3, p. 547-581.

[81] Mookherjee D. (1984), « Optimal Incentive Schemes with Many Agents », *Review of Economic Studies*, vol. 51, n° 3, p. 433-446.

[82] Nalebuff B.J. & Stiglitz J.E. (1983), « Information, Competition, and Markets », *American Economic Review*, vol. 73, n° 2, p. 278-283.

[83] Nalebuff B.J. & Stiglitz J.E. (1983), « Prizes and incentives : towards a general theory of compensation and competition », *The Bell Journal of Economics*, vol. 14, n° 1, p. 21-43.

[84] Potters J., Rockenbach B., Sadrieh A. & van Damme E. (2004), « Collusion under yardstick competition : an experimental study », *International Journal of Industrial Organization*, vol. 22, n° 7, p. 1017-1038.

[85] Pouyet J. (2002), « Collusion Under Asymmetric Information : The Role of the Correlation », *Journal of Public Economic Theory*, vol. 4, n° 4, p. 543-572.

[86] Powell K. & Szymanski S. (1997), « Regulation through comparative performance evaluation », *Utilities Policy*, vol. 6, n° 4, p. 293-301.

[87] Shleifer A. (1985), « A theory of yardstick competition », *RAND Journal of Economics*, vol. 16, n° 3, p. 319-327.

[88] Sobel J. (1999), « A Reexamination of Yardstick Competition », *Journal of Economics & Management Strategy*, vol. 8, n° 1, p. 33-60.

[89] Tangerås T.P. (2002), « Collusion-proof yardstick competition », *Journal of Public Economics*, vol. 83, n° 2, p. 231-254.

[90] Tangerås T.P. (2002), « Regulation of cost and quality under yardstick competition », *IUI working papers*, n° 573, 41 p. Disponible sur : http://www.iui.se/wp/wp573/iuiwp573.pdf

Bibliographie de la partie 3

[91] Battese G.E. & Coelli T. (1995), « A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data », *Empirical Economics*, vol. 20, n° 2, p. 325-332.

[92] Bouf D. & Lévêque J. (2005), « Yardstick Competition for Transport Infrastructure Services », *The Limits of (De-)regulation of Transport Infrastructure Services* / ed. par CEMT, Paris : Éditions de l'OCDE, p. 63-108.

[93] Bouf D. & Péguy P.-Y. (2001), « Is yardstick competition desirable for western European railways ? », *International Journal of Transport Economics*, vol. 28, n° 2, p. 205-227.

[94] Chaffai M.E. (1997), « Estimation de frontières d'efficience : un survol des développements récents de la littérature », *Revue d'économie du développement*, n° 3, p. 33-67.

[95] Chaigneau E. (1997) *La Desserte ferrée des zones urbaines et périurbaines*, tome 4, *L'utilisation des emprises ferroviaires - possibilité techniques et critères de choix*, Paris : CERTU, 126 p.

[96] Coelli T. (1996), « A Guide to FRONTIER Version 4.1 : A Computer Program for Stochastic Frontier Production and Cost Function Estimation », *CEPA Working Papers*, n° 96/07, 33 p.

[97] Coelli T., Perelman S. & Romano E. (1999), « Accounting for Environmental Influences in Stochastic Frontier Models : With Application to International Airlines », *Journal of Productivity Analysis*, vol. 11, n° 3, p. 251-273.

[98] Coelli T., Prasada Rao D.S. & Battese G.E. (1998), *An Introduction to Efficiency and Productivity Analysis*, Dordrecht : Kluwer Academic Publishers, 275 p.

[99] Cowan S. (1997), « Competition in the Water Industry », *Oxford Review of Economic Policy*, vol. 13, n° 1, p. 83-92.

[100] Dalen D.M. & Gómez-Lobo A. (2003), « Yardsticks on the road : Regulatory contracts and cost efficiency in the Norwegian bus industry », *Transportation*, vol. 30, n° 4, p. 371-386.

[101] Estache A., González M. & Trujillo L. (2002), « Efficiency Gains from Port Reform and the Potential for Yardstick Competition : Lessons from Mexico », *World development*, vol. 30, n° 4, p. 545-560.

[102] Estache A., Rossi M.A. & Ruzzier C.A. (2004), « The Case for International Coordination of Electricity Regulation : Evidence from the Measurement of Efficiency in South America », *Journal of Regulatory Economics*, vol. 25, n° 3, p. 271-295.

[103] Farell M.J. (1957), « The measurement of productive efficiency », *Journal of the Royal Statistical Society*, Series A (General), vol. 120, n° 3, p. 253-281.

[104] Kennedy J. & Smith A.S.J. (2004), « Assessing the Efficient Cost of Sustaining Britain's Rail Network : Perspectives based on Zonal Comparisons », *Journal of Transport Economics and Policy*, vol. 38, n° 2, p. 157-190.

[105] Kodde D.A. & Palm F.C. (1986), « Wald Criteria for Jointly testing equality and inequality restrictions », *Econometrica*, vol. 54, n° 5, p. 1243-1248.

[106] Kumbhakar S.C. & Lovell C.A. (2000), *Stochastic Frontier Analysis*, Cambridge : Cambridge University Press, 333 p.

[107] Lawrence D., Houghton J. & George A. (1997), « International Comparisons of Australia's Infrastructure Performance », *Journal of Productivity Analysis*, vol. 8, n° 4, p. 361-378.

[108] Lévêque J. (2003), « Application de la concurrence par comparaison aux marchés régionaux du système ferroviaire français », *Séminaire d'Études et de Statistiques Appliquées à la Modélisation en Économie*, Caen, 8-10 septembre, 19 p. Disponible sur : http://www.let.fr/fr/annuaire/auteurs/affichage.php?id=89

[109] Lévêque J. (2004) « An application proposal of yardstick competition to the regional markets of the French railway system », *European Transport Conference*, Strasbourg, 4-6 octobre, 17 p. Disponible sur : http://www.let.fr/fr/annuaire/auteurs/affichage.php?id=86

[110] Lovell C.A. (1993), « Production frontiers and productive efficiency », *The measurement of productive efficiency. Techniques and applications.* / ed. par Fried H.O., Lovell C.A. & Schmidt S.S., New York : Oxford University Press, p. 3-67.

[111] Mizutani F. (1997), « Empirical analysis of yardstick competition in the Japanese railway industry », *International Journal of Transport Economics*, vol. 24, n° 3, p. 367-392.

[112] Nash C. (2000), « Modelling performance : Rail », *Handbook of Transport Modelling* / ed. par Hensher D.A. & Button K.J., Oxford : Pergamon, p. 565-575.

[113] Nash C. & Shires J. (2000), « Analyse comparative des chemins de fer européens - Évaluation des données actuelles et des indicateurs recommandés », *Méthodes d'analyses comparatives dans les transports* / ed. par CEMT, Paris : Éditions de l'OCDE, p. 127-147.

[114] Okabe M. (2004), « New Passenger Railway Fares », *Japan Railway & Transport Review*, n° 37, p. 4-15.

[115] Oum T.H., Waters II W.G. & Yu C. (1999), « A Survey of Productivity and Efficiency Measurement in Rail Transport », *Journal of Transport Economics and Policy*, vol. 33, n° 1, p. 9-42.

[116] Perelman S. & Pestieau P. (1994), « A Comparative Performance Study of Postal Services : A Productive Efficiency Approach », *Annales d'Économie et de Statistique*, n° 33, p. 187-202.

[117] Riordan M.H. & Sappington D.E.M. (1989), « Second sourcing », *RAND Journal of Economics*, vol. 20, n° 1, p. 41-58.

[118] Shoji K. (2001), « Lessons from Japanese Experiences of Roles of Public and Private Sectors in Urban Transport », *Japan Railway & Transport Review*, n° 29, p. 12-18.

[119] Taylor L.D. & Weisman D.L. (1996), « A note on Price Cap Regulation and Competition », *Review of Industrial Organization*, vol. 11, n° 4,p. 459-471.

[120] Terada K. (2001), « Railways in Japan - Public & Private Sectors », *Japan Railway & Transport Review*, n° 27, p. 48-55.

[121] Waters II W.G. (2000), « Productivity measurement », *Handbook of Transport Modelling* / ed. par Hensher D.A. & Button K.J., Oxford : Pergamon, p. 335-351.

Annexes

Annexe A

Rappels de méthodologie en économétrie des données de panel

L'économétrie des données de panel, développée depuis les années 1970, propose de nombreux modèles adaptés aux différents échantillons à analyser. L'objectif de ces modèles consiste à inclure dans les estimateurs, des effets propres aux différents individus [1]. Pour une présentation approfondie des éléments théoriques de l'économétrie des données de panel, voir Sevestre (2002) [49].

Dans notre cas d'analyse, la nature des données (le fait que la variable longueur de lignes, LL, soit constante dans le temps) contraint le choix des modèles et estimateurs à notre disposition. En effet, prenons par exemple le plus simple de ces modèles : le modèle à effets fixes. Il consiste à introduire une série de constantes individuelles, α_n, identifiant chacune les effets propres à l'individu n. Ce modèle présente l'intérêt de pouvoir être prédictif dès que l'échantillon intègre la totalité de la population [2]. Le modèle à estimer s'écrit alors :

$$\begin{aligned} \ln C_{n,t} &= \alpha_0 + \alpha_n + \beta_{VK} \ln VK_{n,t} + \beta_{LL} \ln LL_{n,t} \\ &\quad + \gamma_{VK} (\ln VK)^2_{n,t} + \gamma_{LL} (\ln LL)^2_{n,t} + \delta \ln VK_{n,t} \ln LL_{n,t} + \epsilon_{n,t} \end{aligned}$$

Dans la mesure où les variables $\ln LL$ et $(\ln LL)^2$ ne varient pas dans le temps, elles sont colinéaires avec les α_n. Il n'est alors plus possible d'en estimer les coefficients β_{LL} et γ_{LL} séparément des α_n. Ce modèle à effets fixes, aux effets spécifiques constants dans le temps (α_n), ne nous permet donc pas d'évaluer correctement les rendements d'échelle.

[1] Nous ne présentons ici que les modèles à effets individuels et non temporels, dans la mesure où la présence d'effets temporels dans nos observations n'est pas avérée.

[2] En revanche, il présente l'inconvénient de ne pas être directement estimable (il faut introduire une contrainte identifiante). Il est par ailleurs biaisé lorsque l'amplitude de la dimension temporelle est réduite.

Le modèle à erreurs composées suppose, lui, que l'hétérogénéité inobservable entre les individus est aléatoire. Le modèle à estimer s'écrit alors :

$$\begin{aligned}
\ln C_{n,t} &= \alpha_0 + \beta_{VK} \ln VK_{n,t} + \beta_{LL} \ln LL_{n,t} \\
&\quad + \gamma_{VK} (\ln VK)^2_{n,t} + \gamma_{LL} (\ln LL)^2_{n,t} + \delta \ln VK_{n,t} \ln LL_{n,t} + \epsilon_{n,t} \\
\epsilon_{n,t} &= u_n + w_{n,t}
\end{aligned}$$

La perturbation $\epsilon_{n,t}$ est donc composée de deux éléments, u_n et $w_{n,t}$. Le premier terme représente l'effet individuel, rendant compte de l'influence sur $\ln C$ des variables non prises en compte, dès lors que celles-ci sont stables dans le temps. Le second terme représente l'influence des autres variables omises, variant dans les deux dimensions. Ces deux éléments vérifient les hypothèses suivantes :

$$\begin{aligned}
\forall n, \mathrm{E}(u_n) &= 0 \\
\forall (n,t), \mathrm{E}(w_{n,t}) &= 0 \\
\forall (n,n'), \mathrm{E}(u_n u_{n'}) &= \delta_{n,n'} \sigma_u^2 \\
\forall (n,n',t,t'), \mathrm{E}(w_{n,t} w_{n',t'}) &= \delta_{n,n'} \delta_{t,t'} \sigma_w^2 \\
\forall (n,n',t), \mathrm{E}(u_n w_{n',t}) &= 0
\end{aligned}$$

Mais ce modèle fait apparaître une autocorrélation temporelle des perturbations $\epsilon_{n,t}$. Cela conduit à recourir à d'autres estimateurs que celui des Moindres Carrés Ordinaires (MCO). Parmi eux, l'estimateur intra-individuel (appelé également estimateur *within*) correspond à l'application des MCO au modèle écrit en écarts aux moyennes individuelles :

$$\begin{aligned}
\ln C_{n,t} - \ln C_{n,.} &= \beta_{VK} (\ln VK_{n,t} - \ln VK_{n,.}) + \beta_{LL} (\ln LL_{n,t} - \ln LL_{n,.}) \\
&\quad + \gamma_{VK} [(\ln VK)^2_{n,t} - (\ln VK)^2_{n,.}] + \gamma_{LL} [(\ln LL)^2_{n,t} - (\ln LL)^2_{n,.}] \\
&\quad + \delta (\ln VK_{n,t} \ln LL_{n,t} - \ln VK_{n,.} \ln LL_{n,.}) + w_{n,t} - w_{n,.}
\end{aligned}$$

Là encore, les variables $\ln LL$ et $(\ln LL)^2$ qui ne varient pas dans le temps vont poser problème puisqu'elles vont disparaître du modèle à estimer. Le second inconvénient de cette méthode d'estimation est qu'elle n'utilise, pour estimer les coefficients, que la variabilité intra-individuelle des observations, excluant toute la variabilité attribuable aux différences permanentes entre individus. Or, c'est dans cette dimension de coupe transversale que nos données présentent l'essentiel de la variance.

L'estimateur le plus souvent retenu dans les modèles à erreur composée est l'estimateur des Moindres Carrés Quasi-Généralisés (MCQG). Celui-ci combine des éléments issus de la régression intra-individuelle (voir ci-dessus) et d'autres, issus de la régression inter-individuelle (application des MCO aux moyennes individuelles). L'inconvénient majeur de cet estimateur est qu'il ne converge qu'à condition que les régresseurs soient exogènes. Cela signifie que les effets spécifiques aléatoires ne

doivent pas être corrélés aux autres variables exogènes du modèle. Dans ce cas, la distribution asymptotique de cet estimateur est la même que celle de l'estimateur intra-individuel. Aussi, afin de tester la convergence de l'estimateur des MCQG, Hausman & Taylor (1981) [30] ont proposé un test, dit « de Hausman », qui compare cet estimateur et sa matrice des variances-covariances aux caractéristiques de l'estimateur *within*. Compte tenu de notre variable LL constante dans le temps [3], le test ne peut porter que sur les estimations des coefficients associés aux variables faisant intervenir VK. C'est ce que proposent Hausman et Taylor.

En cas d'échec de ce test, Hausman et Taylor, toujours, proposent d'estimer le modèle par la méthode des variables instrumentales pour laquelle plusieurs auteurs ont proposé des estimateurs. Mais compte tenu du faible nombre de variables dont nous disposons et des conditions d'exogénéité, ces modèles semblent délicats à mettre en œuvre sur nos données. Nous détaillons dans les paragraphes qui suivent, les résultats obtenus sur notre échantillon régional, d'abord, puis sur l'échantillon des sous-réseaux.

[3] On a vu que l'estimateur *within* ne permettait pas d'en estimer les coefficients.

Annexe B

Programmes d'estimation des fonctions de coût

B.1 Données régionales

B.1.1 Résultats du programme

```
--> REGRESS;Lhs=LNC;Rhs=LNVK2,LNLL2;Panel;Str=STRAT;Random;PrintVC$

+----------------------------------------------------------------------+
| OLS Without Group Dummy Variables                                    |
| Ordinary    least squares regression    Weighting variable = none    |
| Dep. var. = LNC      Mean=   17.99880704    , S.D.=   .5250588850    |
| Model size: Observations =     120, Parameters =   3, Deg.Fr.=    117 |
| Residuals:  Sum of squares= 1.933857854    , Std.Dev.=         .12856 |
| Fit:        R-squared=  .941053, Adjusted R-squared =          .94005 |
| Model test: F[  2,    117] =  933.92,    Prob value =          .00000 |
| Diagnostic: Log-L =     77.4059, Restricted(b=0) Log-L =     -92.4611 |
|             LogAmemiyaPrCrt.=   -4.078, Akaike Info. Crt.=     -1.240 |
| Panel Data Analysis of LNC        [ONE way]                          |
|           Unconditional ANOVA (No regressors)                         |
| Source      Variation        Deg. Free.     Mean Square               |
| Between       32.4011               19.         1.70532               |
| Residual      .405664              100.         .405664E-02           |
| Total         32.8067              119.         .275687               |
+----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 LNVK2     .2869972406E-01  .85094098E-03   33.727   .0000  157.85297
 LNLL2     .1353618871E-01  .34052284E-02    3.975   .0001  48.128954
 Constant  12.81698788      .14417509       88.899   .0000
```

```
+-----------------------------------------------------+
| Random Effects Model: v(i,t) = e(i,t) + u(i)        |
| Estimates:  Var[e]              =   .143170D-02     |
|             Var[u]              =   .150970D-01     |
|             Corr[v(i,t),v(i,s)] =   .913381         |
| Lagrange Multiplier Test vs. Model (3) =  179.66    |
| ( 1 df, prob value =  .000000)                      |
| (High values of LM favor FEM/REM over CR model.)    |
| Reestimated using GLS coefficients:                 |
| Estimates:  Var[e]              =   .239351D-02     |
|             Var[u]              =   .354267D-01     |
|             Sum of Squares          .388876D+01     |
|             R-squared               .941053D+00     |
+-----------------------------------------------------+

+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 LNVK2     .1944544413E-01  .11217936E-02   17.334   .0000  157.85297
 LNLL2     .3397138190E-01  .71466527E-02    4.753   .0000  48.128954
 Constant  13.29428130      .32901142       40.407   .0000


Matrix Cov.Mat. has  3 rows and  3 columns.

               1             2             3
        +-------------------------------------------
       1|  .1258421D-05 -.2776261D-05 -.6502693D-04
       2| -.2776261D-05  .5107464D-04 -.2019928D-02
       3| -.6502693D-04 -.2019928D-02  .1082485D+00
```

B.1.2 Test d'Hausman

```
> setwd("D:/thèse_partie 1/112éco_échelle/reg")
> res=read.table("reg.txt",sep="",dec=",",header=TRUE)
> names(res)
[1] "strat"    "t"        "lnC"      "LNVK"     "lnLL"     "LNVK2"    "lnLL2"
[8] "LNLLlnVK"
> attach(res)
>
> #calcul de l'estimateur within
> x=c(LNVK2)
> X=matrix(x,nrow=120,ncol=1)
>
```

```
> m2=tapply(res$LNVK2,res$strat,mean)
>
> Xn=matrix(nrow=120,ncol=1)
> WnX=matrix(nrow=120,ncol=1)
>
> for (i in 1:120) {
+ j<-res$strat[i]
+ Xn[i]<-m2[j]
+ }
> WnX=X-Xn
>
> m=tapply(res$lnC,res$strat,mean)
>
> yn=matrix(nrow=120,ncol=1)
> Wny=matrix(nrow=120,ncol=1)
>
> for (i in 1:120) {
+ j<-res$strat[i]
+ yn[i]<-m[j]
+ }
> Wny=res$lnC-yn
>
> a=t(X)%*%WnX
> inv=solve(a,diag(1))
> bw=inv%*%t(X)%*%Wny
> bw
           [,1]
[1,] 0.01525306
>
> # calcul de la matrice de variances-covariances du within
> M=Wny-WnX%*%bw
> s2w=1/(120-21)*t(M)%*%M
> s2w
            [,1]
[1,] 0.002229021
> d=matrix(0,nr=1,nc=1)
> diag(d)<-s2w
> Vbw=d%*%inv
> Vbw
             [,1]
[1,] 2.803147e-06
>
```

```
> # test d'Hausman a partir des valeurs fournies par Limdep pour bmcqg
> bmcqg=c(0.01944544413)
> Vbmcqg=matrix(c(0.000001258421),nrow=1,ncol=1)
> Vinv=solve(Vbw-Vbmcqg,diag(1))
> Qh=t(bw-bmcqg)%*%Vinv%*%(bw-bmcqg)
> Qh
         [,1]
[1,] 11.37810
```

B.2 Données sous-réseaux

B.2.1 Résultats du programme

```
--> REGRESS;Lhs=LNC;Rhs=LNVK,LNLL,LNVK2,LNLL2;Panel;Str=STRAT;Random;PrintVC$

+----------------------------------------------------------------------+
| OLS Without Group Dummy Variables                                    |
| Ordinary    least squares regression    Weighting variable = none    |
| Dep. var. = LNC      Mean=   15.94950036    , S.D.=   .9281873263    |
| Model size: Observations =     169, Parameters =   5, Deg.Fr.=    164 |
| Residuals:  Sum of squares= 8.488951844    , Std.Dev.=         .22751 |
| Fit:        R-squared=  .941349, Adjusted R-squared =          .93992 |
| Model test: F[  4,    164] =  658.05,    Prob value =          .00000 |
| Diagnostic: Log-L =     12.9501, Restricted(b=0) Log-L =    -226.7050 |
|             LogAmemiyaPrCrt.=   -2.932, Akaike Info. Crt.=      -.094 |
| Panel Data Analysis of LNC        [ONE way]                          |
|           Unconditional ANOVA (No regressors)                        |
| Source      Variation        Deg. Free.     Mean Square              |
| Between       143.387               32.         4.48085              |
| Residual      1.35007              136.         .992697E-02          |
| Total         144.737              168.         .861532              |
+----------------------------------------------------------------------+


+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 LNVK     -1.184179239      .57260328       -2.068   .0402  17.120078
 LNLL      .8894247013      .22661049        3.925   .0001  4.7407878
 LNVK2     .5202306248E-01  .16485422E-01    3.156   .0019  294.64037
 LNLL2    -.7325022525E-01  .24454306E-01   -2.995   .0032  23.186508
 Constant  18.37649005      4.5463462        4.042   .0001
```

```
            +-----------------------------------------------+
            | Random Effects Model: v(i,t) = e(i,t) + u(i)     |
            | Estimates:  Var[e]              =   .189039D-02  |
            |             Var[u]              =   .498715D-01  |
            |             Corr[v(i,t),v(i,s)] =   .963479      |
            | Lagrange Multiplier Test vs. Model (3) =  259.40 |
            | ( 1 df, prob value =  .000000)                   |
            | (High values of LM favor FEM/REM over CR model.) |
            | Reestimated using GLS coefficients:              |
            | Estimates:  Var[e]              =   .695423D-02  |
            |             Var[u]              =   .871933D-01  |
            |             Sum of Squares          .142360D+02  |
            |             R-squared               .941349D+00  |
            +-----------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 LNVK     -1.074256192      .31006021       -3.465   .0005  17.120078
 LNLL      1.092334346      .27994504        3.902   .0001  4.7407878
 LNVK2     .4312147205E-01  .90031430E-02    4.790   .0000  294.64037
 LNLL2    -.7801666915E-01  .29488296E-01   -2.646   .0082  23.186508
 Constant  18.27399889      2.5788534        7.086   .0000

Matrix Cov.Mat. has  5 rows and  5 columns.
               1             2             3             4             5
        +-----------------------------------------------------------------------
       1|  .9613734D-01 -.1968572D-01 -.2785174D-02  .1916754D-02 -.7763675D+00
       2| -.1968572D-01  .7836923D-01  .5358668D-03 -.8186059D-02 -.2510363D-02
       3| -.2785174D-02  .5358668D-03  .8105658D-04 -.5262963D-04  .2247947D-01
       4|  .1916754D-02 -.8186059D-02 -.5262963D-04  .8695596D-03  .1330666D-02
       5| -.7763675D+00 -.2510363D-02  .2247947D-01  .1330666D-02  .6650485D+01
```

B.2.2 Test d'Hausman

```
> setwd("D:/thèse_partie 1/112éco_échelle/sres")
> res=read.table("sres.txt",sep="",dec=",",header=TRUE)
> names(res)
 [1] "sres"     "année"    "strat"    "t"        "lnC"      "LNVK"
 [7] "lnLL"     "LNVK2"    "lnLL2"    "LNLLlnVK"
> attach(res)
>
> # calcul de l'estimateur within
> x=c(LNVK,LNVK2)
> X=matrix(x,nrow=169,ncol=2)
```

```
>
> m1=tapply(res$LNVK,res$strat,mean)
> m2=tapply(res$LNVK2,res$strat,mean)
>
> Xn=matrix(nrow=169,ncol=2)
> WnX=matrix(nrow=169,ncol=2)
>
> for (i in 1:169) {
+ j<-res$strat[i]
+ Xn[i,1]<-m1[j]
+ Xn[i,2]<-m2[j]
+ }
> WnX=X-Xn
>
> m=tapply(res$lnC,res$strat,mean)
>
> yn=matrix(nrow=169,ncol=1)
> Wny=matrix(nrow=169,ncol=1)
>

> for (i in 1:169) {
+ j<-res$strat[i]
+ yn[i]<-m[j]
+ }
> Wny=res$lnC-yn
>
> a=t(X)%*%WnX
>
> inv=solve(a,diag(2))
>
> bw=inv%*%t(X)%*%Wny
> bw
            [,1]
[1,] -0.77469687
[2,]  0.03265775
>
> # calcul de la matrice de variances-covariances du within
> M=Wny-WnX%*%bw
> s2w=1/(169-35)*t(M)%*%M
> s2w
            [,1]
[1,] 0.007164184
>
> d=matrix(0,nr=2,nc=2)
```

```
> diag(d)<-s2w
> Vbw=d%*%inv
> Vbw
            [,1]          [,2]
[1,]  0.40006075 -0.0116561331
[2,] -0.01165613  0.0003415336
>
> # test d'Hausman a partir des valeurs fournies par Limdep pour bmcqg
> bmcqg=c(-1.074256192,0.04312147205)
> Vbmcqg=matrix(c(0.09613734,-0.002785174,-0.002785174,0.00008105658),nrow=2,ncol=2)
> Vinv=solve(Vbw-Vbmcqg,diag(2))
> Qh=t(bw-bmcqg)%*%Vinv%*%(bw-bmcqg)
> Qh
         [,1]
[1,] 2.204003
```

Annexe C

Résultats estimés des rendements d'échelle et de densité

Les tableaux suivants rapportent les résultats des calculs de rendements d'échelle (*RTS*) et de densité (*RTD*) pour les réseaux régionaux, puis pour les sous-réseaux. Les individus sont identifiés par leur nom administratif ou géographique et l'année d'observation. Pour les réseaux régionaux, nous indiquons également les résultats de la simulation d'un doublement des charges d'infrastructure, ainsi que la valeur des rendements d'échelle au seuil de confiance de 97,5 %.

C.1 Résultats régionaux

individus	*RTS*	*RTS* 97,5 %	*RTS* infra×2	*RTS* 97,5 % infra×2	*RTD*	*RTD* infra×2
Alsace 92	1,091	0,997	1,082	0,990	1,347	1,350
Alsace 93	1,091	0,997	1,082	0,990	1,347	1,350
Alsace 94	1,095	1,000	1,086	0,994	1,353	1,356
Alsace 95	1,101	1,005	1,092	0,998	1,362	1,365
Alsace 97	1,099	1,003	1,089	0,996	1,358	1,361
Alsace 98	1,095	1,000	1,086	0,993	1,353	1,356
Aquitaine 92	1,068	0,968	1,058	0,961	1,350	1,353
Aquitaine 93	1,069	0,968	1,059	0,961	1,351	1,354
Aquitaine 94	1,070	0,970	1,060	0,962	1,353	1,357
Aquitaine 95	1,082	0,979	1,072	0,972	1,373	1,376
Aquitaine 97	1,080	0,977	1,070	0,970	1,369	1,372
Aquitaine 98	1,079	0,977	1,069	0,970	1,368	1,371

individus	*RTS*	*RTS* 97,5 %	*RTS* infra×2	*RTS* 97,5 % infra×2	*RTD*	*RTD* infra×2
Auvergne 92	1,112	1,006	1,101	0,999	1,408	1,411
Auvergne 93	1,115	1,009	1,105	1,002	1,414	1,417
Auvergne 94	1,117	1,011	1,107	1,003	1,417	1,420
Auvergne 95	1,123	1,016	1,112	1,008	1,427	1,430
Auvergne 97	1,115	1,009	1,104	1,001	1,413	1,416
Auvergne 98	1,114	1,008	1,103	1,001	1,412	1,415
Basse-Normandie 92	1,168	1,060	1,157	1,052	1,469	1,472
Basse-Normandie 93	1,172	1,063	1,161	1,055	1,476	1,479
Basse-Normandie 94	1,178	1,068	1,167	1,060	1,485	1,489
Basse-Normandie 95	1,189	1,077	1,178	1,069	1,503	1,506
Basse-Normandie 97	1,179	1,069	1,168	1,061	1,487	1,491
Basse-Normandie 98	1,174	1,065	1,164	1,057	1,480	1,483
Bourgogne 92	1,110	1,003	1,099	0,996	1,411	1,414
Bourgogne 93	1,114	1,007	1,103	0,999	1,418	1,421
Bourgogne 94	1,110	1,003	1,099	0,996	1,411	1,414
Bourgogne 95	1,103	0,997	1,092	0,990	1,399	1,402
Bourgogne 97	1,095	0,991	1,085	0,984	1,387	1,390
Bourgogne 98	1,096	0,992	1,086	0,985	1,389	1,392
Bretagne 92	1,107	1,003	1,096	0,996	1,397	1,400
Bretagne 93	1,098	0,996	1,088	0,989	1,383	1,387
Bretagne 94	1,103	1,000	1,092	0,992	1,390	1,393
Bretagne 95	1,113	1,009	1,103	1,001	1,407	1,411
Bretagne 97	1,113	1,008	1,102	1,000	1,406	1,410
Bretagne 98	1,110	1,005	1,099	0,998	1,401	1,405
Centre 92	1,077	0,975	1,066	0,967	1,363	1,366
Centre 93	1,078	0,976	1,068	0,969	1,365	1,368
Centre 94	1,075	0,974	1,065	0,966	1,360	1,364
Centre 95	1,065	0,965	1,055	0,958	1,344	1,347
Centre 97	1,056	0,958	1,046	0,951	1,330	1,333
Centre 98	1,048	0,952	1,039	0,945	1,318	1,321
Champagne-Ardennes 92	1,114	1,010	1,104	1,003	1,407	1,410
Champagne-Ardennes 93	1,120	1,015	1,109	1,007	1,415	1,419
Champagne-Ardennes 94	1,117	1,013	1,107	1,005	1,411	1,415
Champagne-Ardennes 95	1,133	1,025	1,122	1,017	1,436	1,439
Champagne-Ardennes 97	1,123	1,017	1,113	1,010	1,421	1,424
Champagne-Ardennes 98	1,124	1,018	1,114	1,011	1,422	1,426

individus	*RTS*	*RTS* 97,5 %	*RTS* infra×2	*RTS* 97,5 % infra×2	*RTD*	*RTD* infra×2
Franche-Comté 92	1,141	1,037	1,131	1,029	1,432	1,435
Franche-Comté 93	1,149	1,043	1,138	1,035	1,443	1,447
Franche-Comté 94	1,150	1,044	1,139	1,036	1,446	1,449
Franche-Comté 95	1,164	1,055	1,153	1,048	1,468	1,471
Franche-Comté 97	1,163	1,054	1,152	1,047	1,466	1,469
Franche-Comté 98	1,164	1,056	1,153	1,048	1,468	1,471
Haute-Normandie 92	1,157	1,052	1,147	1,045	1,447	1,450
Haute-Normandie 93	1,171	1,064	1,161	1,056	1,469	1,473
Haute-Normandie 94	1,177	1,068	1,166	1,061	1,478	1,481
Haute-Normandie 95	1,174	1,066	1,163	1,058	1,473	1,476
Haute-Normandie 97	1,170	1,063	1,159	1,055	1,466	1,470
Haute-Normandie 98	1,166	1,060	1,156	1,052	1,462	1,465
Languedoc-Roussillon 92	1,107	1,004	1,097	0,996	1,395	1,399
Languedoc-Roussillon 93	1,111	1,007	1,100	0,999	1,401	1,404
Languedoc-Roussillon 94	1,112	1,008	1,102	1,000	1,403	1,406
Languedoc-Roussillon 95	1,119	1,014	1,108	1,006	1,414	1,417
Languedoc-Roussillon 97	1,115	1,011	1,105	1,003	1,408	1,412
Languedoc-Roussillon 98	1,115	1,010	1,104	1,003	1,408	1,411
Limousin 92	1,195	1,077	1,183	1,069	1,529	1,533
Limousin 93	1,200	1,082	1,188	1,074	1,539	1,542
Limousin 94	1,193	1,076	1,181	1,067	1,526	1,530
Limousin 95	1,195	1,078	1,183	1,069	1,530	1,533
Limousin 97	1,189	1,073	1,177	1,064	1,520	1,523
Limousin 98	1,189	1,073	1,178	1,065	1,521	1,524
Lorraine 92	1,067	0,967	1,057	0,960	1,349	1,352
Lorraine 93	1,066	0,966	1,056	0,959	1,346	1,349
Lorraine 94	1,067	0,966	1,056	0,959	1,347	1,350
Lorraine 95	1,080	0,978	1,070	0,970	1,369	1,372
Lorraine 97	1,076	0,974	1,065	0,967	1,362	1,365
Lorraine 98	1,077	0,975	1,066	0,967	1,363	1,366
Midi-Pyrénées 92	1,064	0,964	1,054	0,957	1,344	1,347
Midi-Pyrénées 93	1,064	0,964	1,054	0,957	1,344	1,347
Midi-Pyrénées 94	1,077	0,975	1,066	0,967	1,364	1,367
Midi-Pyrénées 95	1,084	0,981	1,074	0,973	1,376	1,379
Midi-Pyrénées 97	1,075	0,973	1,065	0,966	1,362	1,365
Midi-Pyrénées 98	1,074	0,972	1,064	0,965	1,360	1,363

individus	*RTS*	*RTS* 97,5 %	*RTS* infra×2	*RTS* 97,5 % infra×2	*RTD*	*RTD* infra×2
Nord-Pas-de-Calais 92	1,042	0,950	1,033	0,944	1,296	1,299
Nord-Pas-de-Calais 93	1,031	0,940	1,022	0,934	1,278	1,281
Nord-Pas-de-Calais 94	1,029	0,939	1,021	0,933	1,276	1,279
Nord-Pas-de-Calais 95	1,034	0,943	1,025	0,937	1,283	1,286
Nord-Pas-de-Calais 97	1,030	0,940	1,021	0,933	1,277	1,280
Nord-Pas-de-Calais 98	1,031	0,941	1,022	0,934	1,278	1,281
Pays de la Loire 92	1,092	0,992	1,082	0,985	1,370	1,374
Pays de la Loire 93	1,090	0,990	1,080	0,983	1,366	1,370
Pays de la Loire 94	1,087	0,988	1,078	0,981	1,363	1,366
Pays de la Loire 95	1,101	1,000	1,091	0,992	1,385	1,388
Pays de la Loire 97	1,094	0,994	1,084	0,986	1,373	1,376
Pays de la Loire 98	1,090	0,990	1,080	0,983	1,367	1,370
Picardie 92	1,072	0,973	1,062	0,966	1,348	1,351
Picardie 93	1,074	0,975	1,064	0,968	1,352	1,355
Picardie 94	1,075	0,976	1,065	0,968	1,353	1,356
Picardie 95	1,059	0,962	1,049	0,955	1,327	1,330
Picardie 97	1,045	0,950	1,035	0,944	1,305	1,308
Picardie 98	1,043	0,949	1,034	0,942	1,303	1,306
Poitou-Charentes 92	1,185	1,066	1,172	1,057	1,527	1,530
Poitou-Charentes 93	1,208	1,085	1,195	1,076	1,566	1,569
Poitou-Charentes 94	1,216	1,092	1,203	1,082	1,580	1,584
Poitou-Charentes 95	1,189	1,070	1,177	1,061	1,535	1,539
Poitou-Charentes 97	1,180	1,062	1,167	1,053	1,519	1,522
Poitou-Charentes 98	1,180	1,062	1,167	1,053	1,519	1,522
PACA 92	1,063	0,969	1,054	0,962	1,323	1,327
PACA 93	1,070	0,974	1,061	0,967	1,334	1,337
PACA 94	1,065	0,970	1,056	0,964	1,327	1,330
PACA 95	1,063	0,968	1,054	0,962	1,323	1,326
PACA 97	1,064	0,969	1,055	0,962	1,324	1,327
PACA 98	1,064	0,969	1,055	0,963	1,325	1,328
Rhône-Alpes 92	0,997	0,905	0,988	0,898	1,255	1,258
Rhône-Alpes 93	1,000	0,907	0,990	0,900	1,259	1,262
Rhône-Alpes 94	1,001	0,908	0,992	0,902	1,262	1,265
Rhône-Alpes 95	1,001	0,908	0,992	0,902	1,262	1,265
Rhône-Alpes 97	0,991	0,900	0,982	0,893	1,246	1,248
Rhône-Alpes 98	0,988	0,897	0,979	0,891	1,241	1,244
Moyenne	1,104	1,001	1,094	0,994	1,392	1,396

C.2 Résultats des sous-réseaux

individus	*RTS*	*RTD*
Strasbourg N 93	1,309	2,097
Strasbourg N 94	1,312	2,105
Strasbourg N 95	1,333	2,160
Strasbourg N 97	1,319	2,124
Strasbourg N 98	1,317	2,119
Strasbourg SO 93	1,206	2,234
Strasbourg SO 94	1,216	2,267
Strasbourg SO 95	1,225	2,301
Strasbourg SO 97	1,216	2,270
Strasbourg SO 98	1,216	2,267
Strasbourg - Mulhouse 93	1,131	1,771
Strasbourg - Mulhouse 94	1,137	1,785
Strasbourg - Mulhouse 95	1,145	1,804
Strasbourg - Mulhouse 97	1,133	1,774
Strasbourg - Mulhouse 98	1,123	1,750
Strasbourg E 93	1,354	2,002
Strasbourg E 94	1,363	2,021
Strasbourg E 95	1,383	2,066
Strasbourg E 97	1,418	2,145
Strasbourg E 98	1,406	2,118
Mulhouse - Kruth 93	1,178	3,090
Mulhouse - Kruth 94	1,182	3,113
Mulhouse - Kruth 95	1,191	3,180
Mulhouse - Kruth 97	1,173	3,053
Mulhouse - Kruth 98	1,177	3,082
Colmar - Metzeral 93	1,230	4,596
Colmar - Metzeral 94	1,242	4,756
Colmar - Metzeral 95	1,248	4,845
Colmar - Metzeral 97	1,233	4,629
Colmar - Metzeral 98	1,242	4,755
Strasbourg - Lauterbourg 93	1,247	2,974
Strasbourg - Lauterbourg 94	1,275	3,135
Strasbourg - Lauterbourg 95	1,257	3,029
Strasbourg - Lauterbourg 97	1,259	3,045
Strasbourg - Lauterbourg 98	1,280	3,169
Mulhouse - Belfort 93	1,243	3,158
Mulhouse - Belfort 94	1,260	3,270
Mulhouse - Belfort 95	1,283	3,427
Mulhouse - Belfort 97	1,283	3,424
Mulhouse - Belfort 98	1,289	3,473
Strasbourg - Kehl 93	1,022	4,741
Strasbourg - Kehl 94	1,057	5,601

individus	*RTS*	*RTD*
Strasbourg - Kehl 95	1,061	5,739
Strasbourg - Kehl 97	1,093	6,817
Strasbourg - Kehl 98	1,069	5,965
Nevers NS 93	1,553	3,371
Nevers NS 94	1,545	3,335
Nevers NS 95	1,551	3,362
Nevers NS 97	1,454	2,939
Nevers NS 98	1,449	2,920
Nevers E 93	1,483	2,327
Nevers E 94	1,470	2,295
Nevers E 95	1,484	2,329
Nevers E 97	1,476	2,309
Nevers E 98	1,479	2,316
Dijon S 93	1,290	1,983
Dijon S 94	1,294	1,992
Dijon S 95	1,313	2,037
Dijon S 97	1,281	1,961
Dijon S 98	1,279	1,956
Paray-le-Monial 93	1,702	3,323
Paray-le-Monial 94	1,700	3,315
Paray-le-Monial 95	1,767	3,579
Paray-le-Monial 97	1,759	3,547
Paray-le-Monial 98	1,756	3,535
Auxerre 93	1,710	2,862
Auxerre 94	1,773	3,045
Auxerre 95	1,745	3,050
Auxerre 97	1,751	3,067
Auxerre 98	1,767	3,116
Dijon - Montereau 93	1,460	2,249
Dijon - Montereau 94	1,409	2,128
Dijon - Montereau 95	1,393	2,093
Dijon - Montereau 97	1,377	2,057
Dijon - Montereau 98	1,387	2,080
Dijon - Is-sur-Tille 93	1,246	3,433
Dijon - Is-sur-Tille 94	1,262	3,555
Dijon - Is-sur-Tille 95	1,276	3,672
Dijon - Is-sur-Tille 97	1,270	3,621
Dijon - Is-sur-Tille 98	1,281	3,711
Dijon - Saint-Amour 93	1,533	3,371
Dijon - Saint-Amour 94	1,502	3,223
Dijon - Saint-Amour 95	1,516	3,290
Dijon - Saint-Amour 97	1,521	3,310
Dijon - Saint-Amour 98	1,523	3,319

individus	*RTS*	*RTD*
Belfort - Dijon 92	1,238	1,881
Belfort - Dijon 93	1,241	1,886
Belfort - Dijon 94	1,248	1,902
Belfort - Dijon 95	1,276	1,969
Belfort - Dijon 97	1,271	1,958
Belfort - Dijon 98	1,274	1,963
Belfort E 92	1,570	2,636
Belfort E 93	1,653	2,879
Belfort E 94	1,664	2,911
Belfort E 95	1,692	2,998
Belfort E 97	1,711	3,058
Belfort E 98	1,692	2,998
Besançon S 92	1,578	2,541
Besançon S 93	1,650	2,733
Besançon S 94	1,616	2,643
Besançon S 95	1,668	2,783
Besançon S 97	1,669	2,786
Besançon S 98	1,679	2,814
Besançon - Le Locle 92	1,380	3,240
Besançon - Le Locle 93	1,398	3,343
Besançon - Le Locle 94	1,379	3,235
Besançon - Le Locle 95	1,403	3,369
Besançon - Le Locle 97	1,396	3,334
Besançon - Le Locle 98	1,402	3,369
Lyon - Grenoble/Chambéry 93	1,180	1,787
Lyon - Grenoble/Chambéry 94	1,182	1,792
Lyon - Grenoble/Chambéry 95	1,176	1,778
Lyon - Grenoble/Chambéry 97	1,142	1,701
Lyon - Grenoble/Chambéry 98	1,136	1,687
Lyon - Genève - sillon alpin 93	1,416	1,603
Lyon - Genève - sillon alpin 94	1,428	1,618
Lyon - Genève - sillon alpin 95	1,421	1,609
Lyon - Genève - sillon alpin 97	1,406	1,590
Lyon - Genève - sillon alpin 98	1,399	1,581
Lyon S 93	1,377	2,071
Lyon S 94	1,362	2,039
Lyon S 95	1,333	1,973
Lyon S 97	1,210	1,827
Lyon S 98	1,204	1,814
Saint-Étienne NO 93	1,352	2,281
Saint-Étienne NO 94	1,348	2,270
Saint-Étienne NO 95	1,365	2,319
Saint-Étienne NO 97	1,343	2,248

individus	*RTS*	*RTD*
Saint-Étienne NO 98	1,333	2,221
Lyon - Roanne 93	1,231	2,310
Lyon - Roanne 94	1,282	2,246
Lyon - Roanne 95	1,291	2,274
Lyon - Roanne 97	1,376	2,155
Lyon - Roanne 98	1,367	2,133
Ouest lyonnais 93	1,308	2,721
Ouest lyonnais 94	1,313	2,744
Ouest lyonnais 95	1,339	2,858
Ouest lyonnais 97	1,310	2,729
Ouest lyonnais 98	1,295	2,666
Valence/Grenoble - Veynes-Dévoluy 93	1,351	2,627
Valence/Grenoble - Veynes-Dévoluy 94	1,466	3,099
Valence/Grenoble - Veynes-Dévoluy 95	1,471	3,118
Valence/Grenoble - Veynes-Dévoluy 97	1,437	2,969
Valence/Grenoble - Veynes-Dévoluy 98	1,629	2,868
Chambéry E 93	1,295	2,045
Chambéry E 94	1,318	2,103
Chambéry E 95	1,327	2,126
Chambéry E 97	1,298	2,054
Chambéry E 98	1,306	2,074
Bourg-en-Bresse 93	1,419	2,171
Bourg-en-Bresse 94	1,427	2,189
Bourg-en-Bresse 95	1,454	2,254
Bourg-en-Bresse 97	1,430	2,197
Bourg-en-Bresse 98	1,414	2,158
Lyon - Macon 93	1,171	2,340
Lyon - Macon 94	1,163	2,308
Lyon - Macon 95	1,160	2,297
Lyon - Macon 97	1,145	2,240
Lyon - Macon 98	1,131	2,186
Lyon - Saint-Étienne 93	1,018	1,873
Lyon - Saint-Étienne 94	1,020	1,878
Lyon - Saint-Étienne 95	1,031	1,917
Lyon - Saint-Étienne 97	1,025	1,896
Lyon - Saint-Étienne 98	1,021	1,881
Saint-Gervais - Vallorcine 93	1,160	3,120
Saint-Gervais - Vallorcine 94	1,177	3,251
Saint-Gervais - Vallorcine 95	1,163	3,145
Saint-Gervais - Vallorcine 97	1,147	3,028
Saint-Gervais - Vallorcine 98	1,149	3,039
moyenne	1,348	2,701

Annexe D

Chronologies des régimes concurrentiels

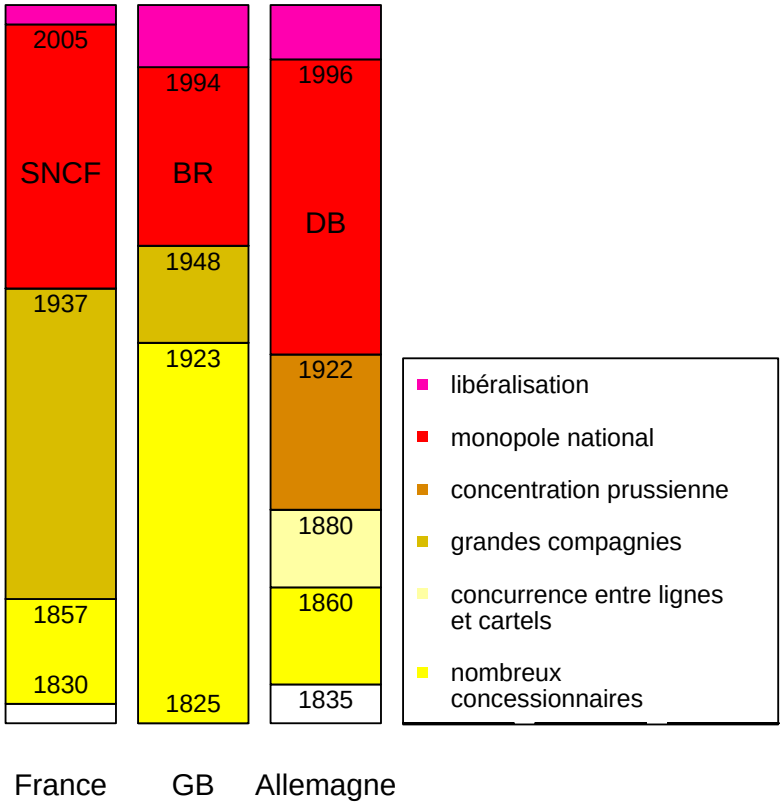

FIG. D.1 – Intégration horizontale en France, Grande-Bretagne et Allemagne. Sources : SNCF, Dodgson (1997) [19], DB, Fremdling (1997) [23].

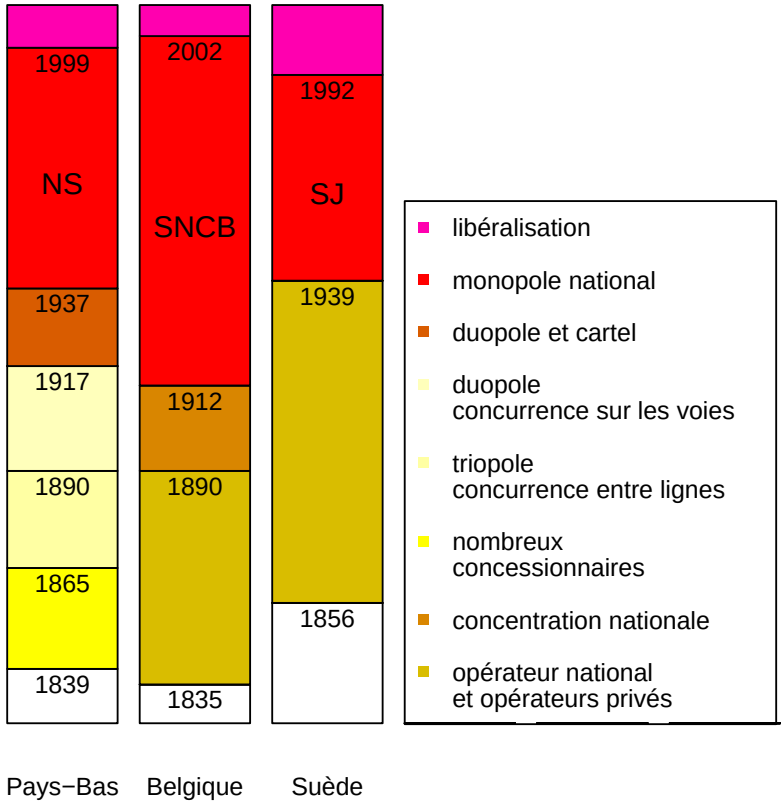

FIG. D.2 – Intégration horizontale en Hollande, Belgique et Suède.
Sources : Fremdling (1997) [23], SNCB, SJ.

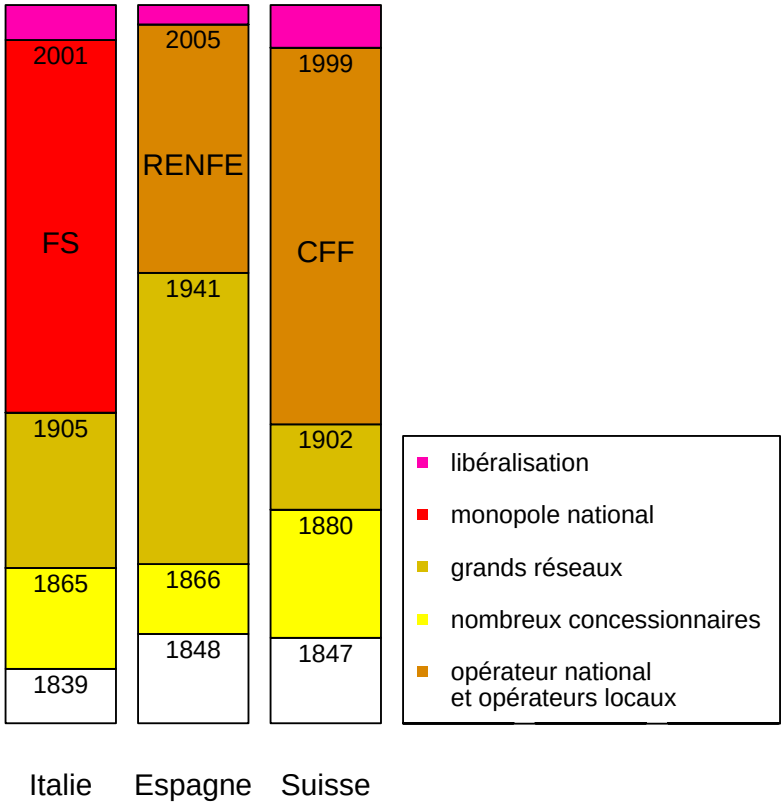

FIG. D.3 – Intégration horizontale en Italie, Espagne et Suisse.
Sources : Merger (1997) [38], Ortunez Goicolea [40], CFF.

Annexe E

Formulation et résolution du modèle économique complet

E.1 Forme et spécification des fonctions

Nous détaillons dans les paragraphes qui suivent la forme et la spécification envisagées pour les différentes fonctions intervenant dans le modèle complet, ébauché à la section 6.1 : fonctions objectif, fonctions de demande et fonctions de coût.

- **Fonction objectif de l'opérateur i :** maximiser son profit. Cette fonction objectif s'inscrit dans le cadre d'un jeu à somme nulle où la coordination centralisée des agents serait donc sans effet et où seul joue l'effet de réputation des gestionnaires d'activités régionales.

$$\underset{s_i}{\text{Max}} \quad P_i - C_1\big[TrK(s_i)\big]$$

$$\underset{s_i}{\text{Max}} \quad R_i + S_i - (C_1 + C_2)\big[TrK(s_i)\big]$$

$$\underset{s_i}{\text{Max}} \quad r\,VK\big[TrK(s_i)\big] + s_i\,TrK(s_i) - (C_1 + C_2)\big[TrK(s_i)\big]$$

- **Fonction objectif du régulateur i :** maximiser son propre surplus. Afin de tenir compte du poids de l'opérateur national dans les négociations, et de l'attitude raisonnable (voire bienveillante) des régulateurs régionaux à son égard, il semble toutefois plus réaliste d'introduire dans la fonction objectif du régulateur le profit de l'opérateur pondéré positivement. Il s'agit donc pour le régulateur régional de maximiser un surplus collectif :

$$\text{Surplus}(s_i) = \int_{s_i}^{\infty} TrK(s)\,\mathrm{d}s + \alpha\Big(r\,VK\big[TrK(s_i)\big] + s_i\,TrK(s_i) - (C_1+C_2)\big[TrK(s_i)\big]\Big)$$

- **Fonction de demande *TrK*.** Nous retenons pour chacune de nos deux fonctions de demande un modèle de type gravitaire ; ces modèles - ainsi nommés par analogie avec la loi de la gravitation de Newton - sont de la forme suivante :

$$T_{j,k} = kC_{j,k}^{E} p_j p_k d_{j,k}^{-1}$$

où
- $T_{j,k}$ correspond à la demande de transport entre les villes j et k ;
- $C_{j,k}$ est un indicateur du coût correspondant ;
- E est l'élasticité-prix de la demande ;
- p_l est la population de la ville l ;
- et $d_{j,k}$ est la distance entre les deux villes.

Nous conservons l'expression de cette fonction sous la forme d'un modèle gravitaire élémentaire, tout en sachant qu'elle pourra être enrichie au fil de l'analyse. Il s'avérera sans doute nécessaire de prendre en compte la concurrence routière et celle des trains grandes lignes, à l'aide d'un modèle logit multinomial. En agrégeant pour chaque région et en introduisant les variables définies dans la partie précédente, on obtient pour la fonction de demande contractuelle :

$$\begin{aligned} TrK &= ks^{E_{TrK/s}} \sum_{(j,k)} p_j p_k d_{j,k}^{-1} \\ \text{soit : } \ln TrK &= K + E_{TrK/s} \ln s + \ln \left(\sum_{(j,k)} p_j p_k d_{j,k}^{-1} \right) \end{aligned}$$

Nous avons par ailleurs testé l'apport d'une variable indicatrice du revenu régional (le revenu moyen des foyers fiscaux). Mais cette variable ne s'est pas révélée significative.

- **Fonction de demande *VK*.** Pour la fonction de trafic, un modèle gravitaire élémentaire devrait suffire pour obtenir une estimation très correcte de VK. Aussi, nous estimerons les paramètres de la fonction suivante :

$$\begin{aligned} VK &= a\, TrK^{\eta} CT^{\beta} \sum_{(j,k)} p_j p_k d_{j,k}^{-1} \\ \text{soit : } \ln VK &= A + \eta \ln TrK + \beta \ln CT + \ln \left(\sum_{(j,k)} p_j p_k d_{j,k}^{-1} \right) \end{aligned}$$

- **Fonction de coût C_1.** Voir page 165 pour les justifications du choix de la spécification : $\ln C_1 = \beta_0 + \beta_{TrK} \ln TrK + \beta_{LL} \ln LL$
- **Fonction de coût C_2.** Ainsi que nous l'avons signalé page 168, il s'agit d'une des principales difficultés de modélisation. Certes, une modélisation log-linéaire de la forme $\ln C_2 = \zeta_0 + \zeta_{TrK} \ln TrK + \zeta_{VK} \ln VK$ explique plus de 90 % de la variance de la variable endogène. Mais elle engendre des résidus parfois très élevés - qu'ils soient négatifs (Picardie, Midi-Pyrénées) ou positifs (Haute-Normandie) - qu'il semble très difficile de réduire.

E.2 Résolution du problème de maximisation

Le problème de maximisation du surplus du régulateur s'écrit donc :

$$\mathrm{Max}_{s_i} \int_{s_i}^{\infty} TrK(s)\,\mathrm{d}s + \alpha\Big(r\,VK\big[TrK(s_i)\big] + s_i\,TrK(s_i) - (C_1 + C_2)\big[TrK(s_i)\big]\Big)$$

$$\begin{aligned}
\Rightarrow\quad & \frac{\mathrm{dSurplus}}{\mathrm{d}s}(s_i) = 0 \\
\Rightarrow\quad & 0 = -\,TrK_i \\
& \quad + \alpha\Big(r\frac{\partial VK}{\partial TrK}(TrK_i)\frac{\partial TrK}{\partial s}(s_i) + TrK_i + s_i\frac{\partial TrK}{\partial s}(s_i) - \frac{\partial(C_1+C_2)}{\partial TrK}(TrK_i)\frac{\partial TrK}{\partial s}(s_i)\Big) \\
\Rightarrow\quad & TrK_i(1-\alpha) = \alpha\frac{\partial TrK}{\partial s}(s_i)\Big(r\frac{\partial VK}{\partial TrK}(TrK_i) + s_i - \frac{\partial C_1}{\partial TrK}(TrK_i) - \frac{\partial C_2}{\partial TrK}(TrK_i)\Big) \\
\Rightarrow\quad & s_i\frac{1-\alpha}{E_{TrK/s}} = \alpha\Big(r\frac{\partial VK}{\partial TrK}(TrK_i) + s_i - \frac{\partial C_1}{\partial TrK}(TrK_i) - \frac{\partial C_2}{\partial TrK}(TrK_i)\Big) \\
\Rightarrow\quad & s_i\left(\frac{1-\alpha}{E_{TrK/s}} - \alpha\right) = \alpha\Big(r\frac{\partial VK}{\partial TrK}(TrK_i) - \frac{\partial C_1}{\partial TrK}(TrK_i) - \frac{\partial C_2}{\partial TrK}(TrK_i)\Big) \\
\Rightarrow\quad & s_i = \frac{r\frac{\partial VK}{\partial TrK}(TrK_i)) - \frac{\partial C_1}{\partial TrK}(TrK_i) - \frac{\partial C_2}{\partial TrK}(TrK_i)}{\left(\frac{1}{\alpha} - 1\right)\frac{1}{E_{TrK/s}} - 1}
\end{aligned}$$

La subvention optimale qui émerge de ce programme se rapproche dans sa forme générale des solutions classiques de Ramsey-Boiteux, ce qui est normal, compte tenu de l'écriture du programme. Les différences qui apparaissent avec le cas standard proviennent du fait que nous considérons une subvention nette des recettes commerciales en guise de prix.

La fonction trafic étant de la forme :

$$VK(TrK) = a\,TrK^{\eta}\sum_{(j,k)}\left(p_j p_k d_{j,k}^{-1}\right)CT^{\beta}$$

D'où :

$$\frac{\partial VK}{\partial TrK} = a\eta\,TrK^{\eta-1}\sum_{(j,k)}\left(p_j p_k d_{j,k}^{-1}\right)CT^{\beta}$$

Les fonctions de coût ayant pour expression :

$$C_1 = \beta_0\,TrK^{\beta_{TrK}}\,LL^{\beta_{LL}}$$

$$C_2 = \zeta_0\,TrK^{\zeta_{TrK}}\,VK^{\zeta_{VK}}$$

D'où :

$$\frac{\partial C_1}{\partial TrK} = \beta_0\beta_{TrK}\,TrK^{\beta_{TrK}-1}\,LL^{\beta_{LL}}$$

$$\frac{\partial C_2}{\partial TrK} = \zeta_0 \zeta_{TrK}\, TrK^{\zeta_{TrK}-1}\, VK^{\zeta_{VK}}$$

On a donc à estimer le système de 4 équations suivant :

$$\begin{aligned}
s_i &= \frac{1}{\left(\frac{1}{\alpha}-1\right)\frac{1}{E_{TrK/s}}-1}\Bigg[ra\eta\, TrK_i^{\eta-1}\sum_{(j,k)}\left(p_j p_k d_{j,k}^{-1}\right)CT_i^{\beta} \\
&\quad -\beta_0\beta_{TrK}\, TrK_i^{\beta_{TrK}-1}LL_i^{\beta_{LL}} - \zeta_0\zeta_{TrK}\, TrK_i^{\zeta_{TrK}-1}LL_i^{\zeta_{LL}}\Bigg] + \epsilon_s \\
\ln TrK_i &= K + E_{TrK/s}\ln s_i + \ln\left(\sum_{(j,k)} p_j p_k d_{j,k}^{-1}\right) + \epsilon_{TrK} \\
\ln VK_i &= \ln a + \eta \ln TrK_i + \ln\left(\sum_{(j,k)} p_j p_k d_{j,k}^{-1}\right) + \beta \ln CT_i + \epsilon_{VK} \\
R_i &= r\, VK_i + \epsilon_R
\end{aligned}$$

Les deux premières équations de ce système sont simultanées ; la première équation étant par ailleurs non linéaire et non identifiable en l'état. Le système peut être estimé par l'introduction de variables instrumentales ou par les doubles moindres carrés. Afin de distinguer les rendements de production, β_{TrK} et ζ_{TrK} qui semblent constants (égaux à 1) d'après les regressions avec les MCO de $\ln C_1$ sur $\ln TrK$ et $\ln LL$ et de $\ln C_2$ sur $\ln TrK$ et $\ln VK$, nous proposons de travailler avec des simulations numériques. En faisant varier les valeurs de ces coefficients autour de 1, un test statistique du ratio de vraisemblance permettra de retenir la valeur la plus appropriée.
Enfin, on obtiendra le coût du service en calculant $C_i = s_i\, TrK_i + r\, VK_i - C_{2,i}$.

Annexe F

Estimation des élasticités du coût aux facteurs d'hétérogénéité

Les élasticités que nous calculons sont des élasticités variables, fonction de la valeur du facteur d'hétérogénéité. Néanmoins, nous procédons à l'application numérique pour la seule valeur moyenne de ce paramètre, afin d'obtenir un simple ordre de grandeur.

Élasticités du coût à la vitesse commerciale : Chaigneau (1997) [95] propose la fonction de coût suivante pour les charges de conduite et d'accompagnement (personnel roulant) :

$$\begin{aligned}
C_{\text{cond, acc}} &= 700t + 13\,TrK \\
C_{\text{cond, acc}} &= TrK\left(\frac{700}{V_{\text{com}}} + 13\right) \\
C_{\text{cond, acc}} = 0,37C_1 \Rightarrow \acute{e}l &= -0,37\frac{700\,TrK}{V_{\text{com}}^2}\frac{1}{\frac{TrK}{V_{\text{com}}}\left(\frac{700}{V_{\text{com}}} + 13\right)} \\
\acute{e}l &= -\frac{0,37}{1 + \frac{13}{700}V_{\text{com}}} \\
\acute{e}l &\sim -16\%
\end{aligned}$$

Élasticités du coût à la part du service réalisée en traction thermique : Compte tenu du barème de prix communiqué dans la convention de la région Pays de la Loire (voir page 170), en supposant en première approximation l'équirépartition des circulations entre heures pleines et creuses, rame tractée et automoteur, on peut

calculer :

$$\begin{aligned}
C_{\text{éner, ent}} = 0,30C_1 \Rightarrow C_{\text{éner, ent}} &= 0,30\Big(5,4[1-\text{part}_{\text{diesel}}] + \big[5,4+0,3\big]\text{part}_{\text{diesel}}\Big)\,TrK \\
\textit{él} &= 0,30\frac{0,3\text{part}_{\text{diesel}}}{5,4[1-\text{part}_{\text{diesel}}] + \big[5,4+0,3\big]\text{part}_{\text{diesel}}} \\
\textit{él} &= 0,30\frac{0,3\text{part}_{\text{diesel}}}{5,4+0,3\text{part}_{\text{diesel}}} \\
\textit{él} &\sim 0,8\%
\end{aligned}$$

Élasticité du coût à la part de kilomètres réalisés à vide :

$$\begin{aligned}
C_{\text{circu-infra}} &= \rho TrK_{\text{tot}} \\
C_{\text{circu-infra}} &= \rho(\text{part}_{\text{charge}} + \text{part}_{\text{vide}})\,TrK_{\text{tot}} \\
C_{\text{circu-infra}} = 0,67C_1 \Rightarrow \textit{él} &= 0,67\frac{\rho TrK_{\text{tot}}\,\text{part}_{\text{vide}}}{\rho(\text{part}_{\text{charge}} + \text{part}_{vide})\,TrK_{tot}} \\
\textit{él} &= 0,67\text{part}_{\text{vide}} \\
\textit{él} &\sim 2,8\%
\end{aligned}$$

Élasticité du coût à la part du service réalisée en rame tractée : Compte tenu du barème de prix communiqué dans la convention de la région Pays de la Loire, en supposant en première approximation l'équirépartition des circulations entre heures pleines et creuses, énergie thermique et électrique, on peut calculer :

$$\begin{aligned}
C_{\text{acc,éner,ent}} = 0,44C_1 \Rightarrow C_{\text{acc,éner,ent}} &= 0,44\Big(4,7[1-\text{part}_{\text{tractée}}] + \big[4,7+1,8\big]\text{part}_{\text{tractée}}\Big)\,TrK \\
\textit{él} &= 0,44\frac{1,8\text{part}_{\text{tractée}}}{4,7[1-\text{part}_{\text{tractée}}] + \big[4,7+1,8\big]\text{part}_{\text{tractée}}} \\
\textit{él} &= 0,44\frac{1,8\text{part}_{\text{tractée}}}{4,7+1,8\text{part}_{\text{tractée}}} \\
\textit{él} &\sim 5\%
\end{aligned}$$

Annexe G

Programme d'estimation de la frontière de coût

```
Output from the program FRONTIER (Version 4.1c)

instruction file = ad.ins
data file =        3823.dta

 Tech. Eff. Effects Frontier (see B&C 1993)
 The model is a cost function
 The dependent variable is logged

the ols estimates are :
                 coefficient     standard-error    t-ratio
  beta 0         0.25152145E+01  0.15470211E+00  0.16258437E+02
  beta 1         0.11039882E+01  0.18367402E-01  0.60105845E+02
  beta 2        -0.16277637E+00  0.30325487E-01 -0.53676425E+01
  sigma-squared  0.43793436E-02

log likelihood function =   0.13117193E+03

the estimates after the grid search were :
  beta 0         0.24625834E+01
  beta 1         0.11039882E+01
  beta 2        -0.16277637E+00
  delta 0        0.00000000E+00
  delta 1        0.00000000E+00
  delta 2        0.00000000E+00
  sigma-squared  0.70179966E-02
  gamma          0.62000000E+00
```

```
 iteration =     0  func evals =     20  llf =  0.13153415E+03
     0.24625834E+01 0.11039882E+01-0.16277637E+00 0.00000000E+00 0.00000000E+00
     0.00000000E+00 0.70179966E-02 0.62000000E+00
 gradient step
 iteration =     5  func evals =     49  llf =  0.13358116E+03
     0.24848169E+01 0.10772621E+01-0.13283659E+00-0.57290776E-01 0.20720241E-01
     0.15330469E-03 0.70719438E-02 0.62506156E+00
 iteration =    10  func evals =     80  llf =  0.14692578E+03
     0.24584934E+01 0.10668227E+01-0.12411413E+00-0.10349293E+00 0.33722607E+01
     0.64661503E-04 0.41739944E-02 0.83420575E+00
 iteration =    15  func evals =    165  llf =  0.15672272E+03
     0.27316861E+01 0.10057945E+01-0.95032908E-01-0.35477335E-01 0.27532914E+01
     0.94356064E-04 0.25784673E-02 0.78922027E+00
 iteration =    20  func evals =    268  llf =  0.15744189E+03
     0.26791586E+01 0.10016667E+01-0.88522034E-01 0.95120847E-02 0.27190963E+01
     0.93057171E-04 0.25444647E-02 0.94753339E+00
 pt better than entering pt cannot be found
 iteration =    23  func evals =    315  llf =  0.15752270E+03
     0.26870829E+01 0.10018924E+01-0.89140593E-01 0.51811107E-02 0.26866059E+01
     0.90592493E-04 0.25847636E-02 0.99999999E+00

the final mle estimates are :

                 coefficient     standard-error    t-ratio
  beta 0         0.26870829E+01  0.19080614E+00  0.14082790E+02
  beta 1         0.10018924E+01  0.24606322E-01  0.40716869E+02
  beta 2        -0.89140593E-01  0.25414551E-01 -0.35074629E+01
  delta 0        0.51811107E-02  0.29782875E-01  0.17396274E+00
  delta 1        0.26866059E+01  0.37962917E+00  0.70769216E+01
  delta 2        0.90592493E-04  0.45796506E-04  0.19781529E+01
  sigma-squared  0.25847636E-02  0.39497209E-03  0.65441678E+01
  gamma          0.99999999E+00  0.33050422E+00  0.30256800E+01

log likelihood function =   0.15752270E+03

LR test of the one-sided error =   0.52701536E+02
with number of restrictions = 4
 [note that this statistic has a mixed chi-square distribution]

number of iterations =     23
(maximum number of iterations set at :   100)

number of cross-sections =     20
```

```
number of time periods =      6

total number of observations =    100

thus there are:     20  obsns not in the panel

covariance matrix :

  0.36406984E-01 -0.28413335E-02 -0.18815444E-02 -0.73124865E-03 -0.17779831E-01
  0.66201968E-05 -0.18947407E-05 -0.98139250E-02
 -0.28413335E-02  0.60547110E-03 -0.30640980E-03  0.17844766E-03 -0.81532165E-03
 -0.83063712E-06  0.28957384E-06 -0.64715077E-03
 -0.18815444E-02 -0.30640980E-03  0.64589939E-03 -0.18915623E-03  0.33732940E-02
 -0.16498883E-07  0.25202852E-06  0.16976878E-02
 -0.73124865E-03  0.17844766E-03 -0.18915623E-03  0.88701966E-03 -0.46593652E-02
  0.98697111E-07 -0.23663792E-05  0.20980428E-02
 -0.17779831E-01 -0.81532165E-03  0.33732940E-02 -0.46593652E-02  0.14411831E+00
 -0.53003999E-05  0.20137590E-05  0.28143623E-01
  0.66201968E-05 -0.83063712E-06 -0.16498883E-07  0.98697111E-07 -0.53003999E-05
  0.20973200E-08 -0.16772740E-08 -0.11284489E-05
 -0.18947407E-05  0.28957384E-06  0.25202852E-06 -0.23663792E-05  0.20137590E-05
 -0.16772740E-08  0.15600295E-06  0.14856419E-04
 -0.98139250E-02 -0.64715077E-03  0.16976878E-02  0.20980428E-02  0.28143623E-01
 -0.11284489E-05  0.14856419E-04  0.10923304E+00

cost efficiency estimates :

     firm  year             eff.-est.
       1     1           0.12557365E+01
       2     1           0.11935543E+01
       3     1           0.10849624E+01
       4     1           0.11598418E+01
       5     1           0.11775061E+01
       6     1           0.11364109E+01
       7     1           0.12100145E+01
       8     1           0.12073099E+01
       9     1           0.11010091E+01
      10     1           0.12099790E+01
      11     1           0.11914005E+01
      12     1           0.10236174E+01
      13     1           0.12045575E+01
      14     1           0.11215248E+01
      15     1           0.12062047E+01
      16     1           0.11898636E+01
```

17	1	0.13167153E+01
18	1	0.10982369E+01
19	1	0.14457498E+01
20	1	0.13986345E+01
1	2	0.12474482E+01
2	2	0.11525311E+01
3	2	0.10729699E+01
4	2	0.11605255E+01
5	2	0.11550490E+01
6	2	0.11264457E+01
7	2	0.11694443E+01
8	2	0.11846385E+01
9	2	0.10949750E+01
10	2	0.12141549E+01
11	2	0.11831905E+01
12	2	0.10214199E+01
13	2	0.11949707E+01
14	2	0.11217123E+01
15	2	0.11700785E+01
16	2	0.11755053E+01
17	2	0.13311290E+01
18	2	0.10977853E+01
19	2	0.14057224E+01
20	2	0.13533199E+01
1	3	0.12992265E+01
2	3	0.12068337E+01
3	3	0.11365412E+01
4	3	0.12080639E+01
5	3	0.12071352E+01
6	3	0.11692862E+01
7	3	0.11620354E+01
8	3	0.12064602E+01
9	3	0.11291804E+01
10	3	0.12375469E+01
11	3	0.12580034E+01
12	3	0.10944467E+01
13	3	0.12435940E+01
14	3	0.11955822E+01
15	3	0.12177159E+01
16	3	0.12183657E+01
17	3	0.12847307E+01
18	3	0.11972487E+01
19	3	0.14779131E+01
20	3	0.14058999E+01

```
      1     5          0.12450127E+01
      2     5          0.12009589E+01
      3     5          0.11533854E+01
      4     5          0.12009869E+01
      5     5          0.11238935E+01
      6     5          0.11926541E+01
      7     5          0.10799803E+01
      8     5          0.11468691E+01
      9     5          0.10705678E+01
     10     5          0.12227161E+01
     11     5          0.12535941E+01
     12     5          0.10408153E+01
     13     5          0.12481664E+01
     14     5          0.12023700E+01
     15     5          0.11980332E+01
     16     5          0.12295007E+01
     17     5          0.13026366E+01
     18     5          0.10849657E+01
     19     5          0.14092185E+01
     20     5          0.13299704E+01
      1     6          0.12545949E+01
      2     6          0.12252173E+01
      3     6          0.11946556E+01
      4     6          0.12458049E+01
      5     6          0.11572931E+01
      6     6          0.11144705E+01
      7     6          0.11026250E+01
      8     6          0.11395193E+01
      9     6          0.11052692E+01
     10     6          0.12362415E+01
     11     6          0.13212554E+01
     12     6          0.10741598E+01
     13     6          0.12868981E+01
     14     6          0.12071573E+01
     15     6          0.12073927E+01
     16     6          0.11829598E+01
     17     6          0.13170751E+01
     18     6          0.10927016E+01
     19     6          0.14173879E+01
     20     6          0.13380537E+01
mean efficiency =   0.12027868E+01
```

```
summary of panel of observations:

(1 = observed, 0 = not observed)

   n t:1   2   3   4   5   6
   1   1   1   1   0   1   1   5
   2   1   1   1   0   1   1   5
   3   1   1   1   0   1   1   5
   4   1   1   1   0   1   1   5
   5   1   1   1   0   1   1   5
   6   1   1   1   0   1   1   5
   7   1   1   1   0   1   1   5
   8   1   1   1   0   1   1   5
   9   1   1   1   0   1   1   5
  10   1   1   1   0   1   1   5
  11   1   1   1   0   1   1   5
  12   1   1   1   0   1   1   5
  13   1   1   1   0   1   1   5
  14   1   1   1   0   1   1   5
  15   1   1   1   0   1   1   5
  16   1   1   1   0   1   1   5
  17   1   1   1   0   1   1   5
  18   1   1   1   0   1   1   5
  19   1   1   1   0   1   1   5
  20   1   1   1   0   1   1   5
      20  20  20   0  20  20 100
```

Annexe H

Autres classements des opérateurs régionaux

H.1 Classement suivant l'efficience brute

région	efficience brute
Limousin	1,000
Poitou-Charentes	1,017
Centre	1,026
Franche-Comté	1,029
Bretagne	1,038
Champagne-Ardennes	1,061
Bourgogne	1,077
Pays de la Loire	1,101
Auvergne	1,112
Midi-Pyrénées	1,124
Nord-Pas-de-Calais	1,124
Aquitaine	1,141
Haute-Normandie	1,151
Basse-Normandie	1,160
Alsace	1,168
Lorraine	1,198
Picardie	1,227
Languedoc-Roussillon	1,230
Rhône-Alpes	1,246
PACA	1,320

TAB. H.1 – Scores d'efficience brute normés pour 1998 ($e^{u_{i,1998}}/e^{u_{\text{Limousin},1998}}$).

H.2 Classement suivant les coûts unitaires

région	C/TrK
Centre	1,000
Poitou-Charentes	1,008
Limousin	1,012
Bretagne	1,035
Bourgogne	1,059
Franche-Comté	1,060
Champagne-Ardennes	1,063
Midi-Pyrénées	1,092
Auvergne	1,104
Pays de la Loire	1,106
Aquitaine	1,110
Nord-Pas-de-Calais	1,123
Lorraine	1,166
Rhône-Alpes	1,172
Basse-Normandie	1,202
Haute-Normandie	1,203
Picardie	1,211
Alsace	1,219
Languedoc-Roussillon	1,231
PACA	1,330

TAB. H.2 – Coûts unitaires normés pour 1998.

Réguler les chemins de fer
sur une proposition de la nouvelle économie de la réglementation :
la « concurrence par comparaison » (*yardstick competition*)

Résumé

La libéralisation des marchés impose aux chemins de fer une révolution économique. Elle remet en cause l'organisation monopolistique au profit du développement de nouveaux marchés. L'analyse des défaillances des marchés (notamment l'évaluation des économies d'échelle) et des différentes formes de réglementation justifie le recours à une régulation incitative des marchés ferroviaires. À ce titre, la concurrence par comparaison paraît intéressante : la comparaison des performances d'entreprises semblables permet de réduire les contraintes informationnelles de la régulation classique. Malgré quelques limites théoriques, les mises en œuvre de mécanismes comparatifs sont nombreuses, comme par exemple dans les chemins de fer japonais. L'application proposée de la concurrence par comparaison aux TER français doit en améliorer la performance. En effet, l'estimation d'une frontière de coût révèle d'importantes disparités d'efforts entre les régions (représentant plus de 100 M€ par an), que le mécanisme comparatif proposé peut réduire.

Mots-clefs

concurrence par comparaison ; *yardstick competition* ; régulation ; réglementation ; chemins de fer ; transport ferroviaire ; SNCF ; TER.

Regulating railways at a proposition of the new economics of regulation: "yardstick competition"

Abstract

The market liberalization leads railways to an economic revolution. It favours the development of new markets, instead of the monopolistic organization. The analysis of the market failures (in particular, the assessment of economies of scale) and of the different forms of competition justifies the implementation of an incentive regulation of the railway markets. Therefore, yardstick competition looks interesting: the comparison of similar firms' performances leads to reduce the informational constraints of the usual monopoly regulation. In spite of some theoretical limits, there are many implementations of comparison mechanisms, such as in the Japanese railways. The proposal of yardstick competition implementation to the French regional trains should increase their performances. The estimation of a cost frontier reveals large effort differences between the regions (corresponding to more than 100 M€ yearly). The comparison mechanism we propose can reduce that overcost.

Keywords

yardstick competition; yardstick regulation; regulation; railways; SNCF; regional trains.

Laboratoire d'Économie des Transports, ISH, 14 avenue Berthelot, 69363 Lyon Cedex 07