

Université de Lille : Sciences et Technologies

Faculté des Sciences Économiques et Sociales

De l'incidence des infrastructures sur le secteur du transport de personnes en France : Contributions à la mesure semi-paramétrique de la productivité et de l'efficacité

Thèse pour obtenir le grade de
Docteur en Sciences Economiques

présentée et soutenue publiquement par

Linjia ZHANG

le 10 juillet 2013

devant le jury composé de

Jean-Philippe Boussemart :

Professeur à l'IESEG School of Management, Universités Charles de Gaulle-Lille 3

Hédi Essid :

Maître de conférences à l'Institut Supérieur de Gestion de Tunis, rapporteur

Hervé Leleu :

Directeur de Recherche au CNRS et IESEG Lille, Professeur à l'IESEG School of Management, Université Catholique de Lille

Pierre Ouелlette :

Professeur à l'Université du Québec à Montréal, Co-directeur de thèse

Nicolas Peypoch :

Maître de conférences à l'Université de Perpignan rapporteur

Stéphane Vigeant :

Professeur à l'Université des Sciences et Technologies de Lille, Directeur de thèse

L'Université de Lille : Sciences et Technologies n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Je tiens tout d'abord à remercier mon directeur de recherche, le Professeur Stéphane Vigeant. Cette thèse qui trouve son aboutissement n'aurait pas pu voir le jour sans lui, qui a dirigé mon travail de bout en bout. Il m'a encadré, encouragé et, surtout, consacré énormément de temps et de patience tout au long de ces années de doctorat. C'est un honneur d'être son étudiant. Stéphane a toujours su m'indiquer de bonnes directions de recherche quand il le fallait et a toujours été très patient quand venait le temps de m'expliquer les concepts techniques les plus difficiles. J'ai sincèrement apprécié travailler avec lui. Mes remerciements s'étendent aussi à mon co-directeur, le Professeur Pierre Ouellette qui m'a beaucoup aidé pour améliorer mes connaissances en modélisation et m'indiquer des avenues de recherche. Je suis très honoré d'avoir travaillé avec lui et ses commentaires et suggestions m'accompagneront toute ma vie.

Je remercie les membres du jury : MM. les Professeurs Hédi Essid et Nicolas Peypoch qui m'ont fait l'honneur d'être mes rapporteurs; MM. les Professeurs Hervé Leleu et Jean-Philippe Boussemart qui ont accepté d'y participer.

Je tiens à remercier également le Professeur Hubert Jayet qui m'a accueilli dans le laboratoire EQUIPPE et m'a aidé à trouver un financement. Je remercie aussi Monsieur Jean Dominique Blardone du Ministère des équipements et le Professeur Alain Ayong-le-Kama pour le contact INSÉE pour ma base de données. Je remercie particulièrement Madame Sandrine Maës, responsable de centre de documentation, qui m'a beaucoup aidé pour trouver certaines références difficiles à trouver ailleurs.

Je tiens à exprimer mes remerciements aussi aux doctorants de la « ruche » qui m'ont accueilli chaleureusement à mon arrivée au laboratoire et ont facilité mon intégration. Ces mots adressent notamment à Romain Noël qui est toujours disponible quand j'ai besoin d'aide pour les études et dans ma vie personnelle, Kassoum Ayouba que je

considère comme un ami et à qui je peux demander toutes les aides possibles, et Sylvain Petit qui m'a beaucoup aidé au début de ma vie comme doctorant et m'a donné plusieurs conseils utiles. Mes remerciements s'adressent aussi au Joseph Faye et Hélène Zwick.

Je remercie enfin ma chère femme, Yingchan Luo, et mes parents, Yiqiu Gu et Xiaoqiang Zhang, pour leur patience, leur soutien et leurs encouragements pendant ces très longues journées de rédaction. Ils sont la source de la force qui m'a permis de finir ce travail.

TABLE DES MATIÈRES

Introduction.....	1
0-1 Le contexte de la recherche	1
0-2 L'importance des infrastructures en général.....	3
0-3 L'importance des infrastructures de transport	6
0-4 La problématique de la thèse	12
0-5 Le plan de la thèse	14
Première Partie.....	17
Une revue de la littérature théorique et empiriques portant sur la mesure de l'effet des infrastructures sur la croissance économique	17
Chapitre 1. Les infrastructures et la croissance économique.....	19
1-1 Le développement sur la définition de l'infrastructure	19
1-2 Le développement de la pensée économique	23
1-3 Revue de littérature sur des recherches empiriques.....	28
1-3-1 Utilisation d'une fonction de production au niveau macro-économique	29
1-3-2 Utilisation d'une fonction de production au niveau micro-économique	32
1-3-3 Par la fonction de coût agrégée	37
1-3-4 Les autres approches	41
Conclusion	43
Chapitre 2. Mesure de l'efficacité des entreprises : une revue et une synthèse méthodologique.....	45

2-1 La productivité et l'efficacité	45
2-1-1 Les notations de base.....	46
2-1-2 L'efficacité au sens de Farrell.....	52
2-2 La revue méthodologique	55
2-2-1 L'approche paramétrique.....	55
2-2-1-1 La méthode de moindres carrés corrigés	57
2-2-1-2 Les frontières stochastiques	57
2-2-2 L'approche non paramétrique.....	61
2-2-2-1 Le modèle DEA classique.....	62
2-2-3 L'approche semi-paramétrique : le modèle d'Aigner et Chu	64
2-2-3-1 Modèle de base	65
2-2-3-2 La littérature autour de modèle d'Aigner et Chu	67
Conclusion	71
Deuxième partie.....	73
Modélisation : introduction des inputs quasi-fixes et de l'indicateur d'accessibilité dans le modèle de Aigner et Chu.	73
Introduction.....	75
Chapitre 3. Modélisation.....	77
3-1 Modélisation du modèle non paramétrique	77
3-1-1 Modèle DEA avec inputs quasi-fixes	77
3-1-2 L'efficacité d'échelle avec DEA	79
3-2 Modélisation du modèle semi-paramétrique	82

3-2-1 Le modèle semi-paramétrique primal	82
3-2-2 Le modèle semi-paramétrique dual	84
Conclusion	89
Chapitre 4. La convergence des estimateurs et l'inférence statistique du modèle semi-paramétrique	91
4-1 Sur la convergence du modèle de frontière de Aigner et Chu	91
4-1-1 Le modèle linéaire/exponentiel	93
4-1-1-1 Estimation par maximum vraisemblance.....	94
4-1-1-2 Convergence de l'estimateur.....	99
4-1-2 Le modèle quadratique /semi-normal.....	111
4-1-2-1 Estimation par maximum vraisemblance.....	111
4-1-2-2 Démonstration de la convergence	117
4-2 Inférence statistique du modèle semi-paramétrique	128
4-2-1 Les méthodes ré-échantillonnages	130
4-2-1-1 Concepts et notations	131
4-2-1-2 Principes du ré-échantillonnage.....	133
4-2-1-3 Sous-échantillonnage	136
4-2-2 La correction du biais.....	139
4-2-3 Estimation des intervalles de confiance	139
4-2-3-1 La définition de l'intervalle de confiance	139
4-2-3-2 La méthode basée sur l'écart-type	140
4-2-3-3 La méthode basée sur le percentile simple	140

4-2-3-4 La méthode de calcul des intervalles de confiance	141
4-2-4 Algorithme pour calculer l'inférence statistique	143
Conclusion	144
Chapitre 5. L'indicateur d'infrastructure	145
5-1 Une mesure de la disponibilité des infrastructures	146
5-2 Les fondements de l'accessibilité.....	148
5-2-1 Les indicateurs simples	148
5-2-2 La mesure basée sur l'activité	149
5-3 Une mesure des infrastructures.....	151
5-3-1 La forme de l'indicateur	152
5-3-2 Déflateur à la distance et déflateur au temps.....	154
5-3-2-1 Modèle gravitaire simple et ses limites.....	154
5-3-2-2 Modèle gravitaire double contrainte et son ajustement	155
Conclusion	160
Troisième partie :	161
Impact des infrastructures sur la productivité des entreprises de transport routier de voyageurs en France	161
Introduction.....	163
Chapitre 6. Les données	167
6-1 Source des données.....	167
6-2 Sélection des variables.....	170
6-2-1 Inputs variables	171

Main d'œuvre.....	172
Énergie	172
Entretien et réparation.....	173
Matériels et fournitures	173
6-2-2 Inputs quasi-fixes	177
Véhicules.....	178
Les infrastructures.....	178
6-2-3 Outputs et coût variable.....	185
6-3 Synthèse des données.	185
6-4 Ajustement des variables de prix pour la fonction de coût.....	190
Chapitre 7 : Application au transport routier de voyageurs avec un modèle non paramétrique DEA standard.....	193
7-1 Introduction	193
7-2 Applications DEA au secteur du transport.....	193
7-3 Modèle et résultats.....	195
Conclusion	203
Chapitre 8. Application empirique avec le modèle semi-paramétrique.....	205
8-1 La spécification du modèle : forme fonctionnelle.....	205
8-2 Ajustement du modèle	208
8-3 Mesure de la technologie.....	211
8-3-1 Fonction de production.....	211
Efficacités technique	211

Élasticité des inputs.....	212
Rendements d'échelle	212
8-3-2 Fonction de coût	213
Efficacités de coût.....	213
Impact des inputs quasi-fixes sur les inputs variables	214
Récupération des rendements d'échelle à partir de la fonction de coût variable.....	214
8-3-3 Les mesures de la technologie dans le cas Translog	215
8-4 Résultats du modèle et analyse des résultats	217
8-4-1 Résultats de la fonction de production	218
8-4-1-1 Coefficients estimés et le résultat de simulation.....	218
Elasticité des inputs.....	223
Les rendements d'échelle.....	226
Efficacité technique	233
8-4-2 Résultat de la fonction de coût	239
Coefficients estimés	239
Elasticités	244
Rendements d'échelle	246
Efficacité	252
8-4-3 Comparaison entre les modèles primal et dual de Aigner et Chu et le modèle DEA.....	256
8-4-4 Résultats sans indicateur d'infrastructures	257
Conclusion	266

Conclusion Générale.....	267
BIBLIOGRAPHIE.....	269
Annexes.....	283
Annexe A. Les hypothèses de Wald (1949)	283
Annexe B. Les indices de prix utilisé pour calculer l'indice de Laspeyres	285
Annexe C. Les résultats pour la fonction production avec des inputs agrégés..	287

LISTE DES TABLEAUX

Tableau 1-1 : Études empiriques utilisant la fonction de production au niveau national.....	33
Tableau 1-2 : Étude empiriques utilisant la fonction de production au niveau régional	35
Tableau 1-3 : Études empiriques utilisant la fonction de coût au niveau macroéconomique	39
Tableau 1-4 : Études empiriques utilisant la fonction de coût au niveau microéconomique.....	40
Tableau 6-1 : L'indice des prix des matériaux	176
Tableau 6-2 : Les différentes entrées utilisées pour construire l'indice de Laspeyres	176
Tableau 6-3 : Déflateurs de distance et du temps perdus à cause de la congestion	184
Tableau 6-4 : Sources et descriptions des variables.....	186
Tableau 6-5 : Statistiques descriptives des variables du modèle primal	188
Tableau 6-6 : Variables utilisées dans le modèle paramétrique et la fonction primal du modèle semi-paramétrique	190
Tableau 6-7 : Variables utilisées dans la fonction duale du modèle semi-paramétrique	191
Tableau 6-8 : Statistiques descriptives des variables utilisées dans le modèle dual	191
Tableau 7-1 : Statistique des coefficients d'efficacité pondéré par le coût (CCR)	196

Tableau 7-2 : Statistique des coefficients d'efficacité pondéré par le coût (BCC)	197
Tableau 7-3 : Statistique des coefficients d'efficacité pondéré par le coût par année	198
Tableau 7-4 : Efficacité et les rendements d'échelles selon la taille des entreprises	202
Tableau 7-5 : Comparaison des efficacités moyennes	202
Tableau 8-1 : Les modèles empiriques.....	208
Tableau 8-2 : Résultats de simulation des coefficients pour la fonction de production sous forme Cobb-Douglas	219
Tableau 8-3 : Résultats de simulation des coefficients pour la fonction de production sous forme Translog.....	220
Tableau 8-4 : Élasticité de l'output par rapport aux inputs	224
Tableau 8-5 : Résultats de simulation pour les rendements d'échelles en utilisant une fonction de production	231
Tableau 8-6 : Efficacité technique, modèle semi-paramétrique, primal	233
Tableau 8-7 : Coefficients d'efficacité par région, modèle semi-paramétrique, primal	238
Tableau 8-8 : Résultats de simulation des coefficients pour la fonction de coût Cobb-Douglas	240
Tableau 8-9 : Résultats de simulation des coefficients pour la fonction de coût Translog	241
Tableau 8-10 : Elasticité des inputs pour la fonction de coût	244
Tableau 8-11 : Effet externe des infrastructures sur les autres inputs.....	246

Tableau 8-12 Résultats de simulation des rendements d'échelles pour la fonction de coût.....	250
Tableau 8-13 : Efficacité des entreprises par la fonction de coût	252
Tableau 8-14 : Coefficients de l'efficacité-coût par région, modèle semi-paramétrique, forme Translog.....	255
Tableau 8-15 : Coefficients estimés avec la forme Cobb-Douglas	258
Tableau 8-16 : Coefficients estimés de la fonction de production Translog avec et sans infrastructures.....	259
Tableau 8-17 : Coefficients estimés de la fonction de coût Translog avec et sans infrastructures	260
Tableau 8-18 : Efficacité technique avec et sans infrastructure, modèle semi-paramétrique de Aigner et Chu	261
Tableau 8-19 : Élasticité des inputs et rendements d'échelle par la fonction de production	261
Tableau 8-20 : Élasticité des inputs et rendements d'échelle avec la fonction de coût.....	262
Tableau 8-21 : Comparaison des mesures d'efficacité et des rendements d'échelle selon la méthode, la taille des entreprises et la présence ou non des infrastructures	265
Tableau Annexe B-1 : Tableau entrées-sorties pour le transport routier de voyageur.....	285
Tableau Annexe B-2 : Indices de prix utilisés	286
Tableau Annexe C-1: Coefficients estimés de la fonction de production avec inputs agrégés sous forme Cobb-Douglas.....	287

Tableau Annexe C-2: Coefficients estimés de la fonction de production avec inputs agrégés sous forme Translog.....	288
Tableau Annexe C-4 : Élasticités de l'output par rapport aux inputs et les rendements d'échelles de la fonction de production avec inputs agrégés .	291
Tableau Annexe C-5 : Intervalle de confiance pour les rendements d'échelles de la fonction de production avec inputs agrégés.....	292

LISTE DES FIGURES

Figure 0-1 : L'investissement en infrastructures de transport terrestre en pourcentage du PIB en prix courants	7
Figure 0-2 : La relation entre les infrastructures de transport et la croissance économique	8
Figure 2-1 : Frontière de production et efficacité technique.....	48
Figure 2-2 : Efficacité technique de Farrell	53
Figure 2-3 : Efficacité allocative de Farrell	54
Figure 3-1 : Modèle BCC avec les inputs quasi-fixes	78
Figure 3-2 : Frontière d'efficacité estimée par CCR et BCC.....	79
Figure 6-1 : La carte de la France au niveau des zones d'emplois	182
Figure 6-2 : La carte de centroïdes au niveau zone d'emploi	183
Figure 6-3 : La carte croisée (les lignes rouges représentent les routes principales).....	183
Figure 6-4 : La distribution d'indicateur des infrastructures par zone d'emploi	184
Figure 7-1 : Différences entre les coefficients des modèles BCC et CCR	200
Figure 7-2 : Différence d'efficacité du modèle BCC avec et sans indicateur classé par ordre croissant de l'indicateur d'infrastructures.....	201
Figure 8-1 : Histogramme des rendements d'échelles avec et sans les infrastructures pour la fonction de production Translog avec le critère de minimisation de la somme des erreurs (cas PL)	229
Figure 8-2 : Histogramme des rendements d'échelles avec et sans les infrastructures, pour la fonction de production Translog avec le critère de minimisation de la somme du carré des erreurs (cas PQ).....	230

Figure 8-3 : Relation entre l'efficacité et le chiffre d'affaires des firmes en utilisant une fonction de production Cobb-Douglas et en minimisant la somme des erreurs (cas PL)	235
Figure 8-4 : Relation entre l'efficacité et le chiffre d'affaires des firmes en utilisant une fonction de production Cobb-Douglas et en minimisant la somme du carré des erreurs (cas PQ).....	235
Figure 8-5 : Relation entre l'efficacité et le chiffre d'affaire des firmes en utilisant une fonction de production Translog et en minimisant la somme des erreurs (cas PL).....	236
Figure 8-6 : Relation entre l'efficacité et le chiffre d'affaire des firmes en utilisant une fonction de production Translog et en minimisant la somme du carré des erreurs (cas PQ)	236
Figure 8-7 : Histogramme des rendements d'échelles avec et sans infrastructures, pour la fonction de coût Translog en minimisant la somme des erreurs (cas PL).....	248
Figure 8-8 : Histogramme des rendements d'échelles avec et sans infrastructures, pour la fonction de coût Translog en minimisant la somme du carré des erreurs (cas PQ).....	249
Figure 8-9 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Cobb-Douglas en minimisant la somme des erreurs	253
Figure 8-10 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Cobb-Douglas en minimisant la somme du carré des erreurs	253
Figure 8-11 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Translog en minimisant la somme des erreurs	254
Figure 8-12 : Relation entre la performance et le chiffre d'affaire avec la fonction	

de coût Translog en minimisant la somme du carré des erreurs254

LISTE DES ACRONYMES

BCC : Banker – Charnes – Cooper pour le modèle DEA avec les rendements d'échelles Variables

CCR : Charnes – Cooper – Rhodes pour le modèle DEA avec les rendements d'échelles constants

CGDD : Commissariat général au développement durable

DEA : *Data Envelopment Analysis*

DMU : *Decision Making Units*

DOM : Départements d'Outre-Mer

EAE : Enquête Annuelle des Entreprises

FDH : *Free Disposal Hull*

FDC : Fonction de Probabilité Cumulative

INSEE : Institut National de la Statistique et des Études Économiques

MCO : Moindres Carrés Ordinaires

MCC : Moindres Carrés Corrigé

MCOC : Moindre Carré Ordinaires Corrigé

MIT : Moyenne de l'Inefficacité Technique

NAF : Nomenclature d'Activités Française

PGD : Processus Générateur des Données

PL : Programme Linéaire

PQ : Programme Quadratique

REC : Rendements d'Échelle Constants

RED : Rendements d'Échelle Décroissants

REI : Rendements d'Échelle Croissants

REV : Rendements d'Échelle Variables

SOC : *Social Overhead Capital*

TES : Tableau Entrées-Sorties

Introduction

0-1 Le contexte de la recherche

Le développement des revenus et le niveau de vie d'une société dépend de sa capacité à transformer les ressources disponibles en biens et services. Cette capacité dépend de sa productivité et de l'efficacité de l'appareil productif. Plusieurs économistes ont cherché les sources de la croissance économique. Le modèle néoclassique proposé par Solow (1957) explique la croissance économique par une théorie exogène ; dans ce modèle, le taux de croissance de long terme d'une économie est déterminé de façon exogène par le progrès technologique et la dynamique démographique. Avec la théorie de la croissance endogène, Romer (1986) et Lucas (1988) ont essayé d'expliquer la croissance des écarts de richesse entre les pays développés et les pays pauvres. Dans leur modèle, l'écart économique entre les pays peut toujours exister, et le progrès technique est déterminé par les investissements. Peu importe le modèle ou les arguments théoriques retenus, tous cherchaient à expliquer la croissance de la productivité car elle constitue l'une des principales sources de la croissance économique.

La définition de la productivité est tout simplement le rapport entre la production et une combinaison des inputs réels utilisés pour réaliser cette production sur la même période. Deux types de mesure de la productivité ont été utilisés : la productivité partielle et la productivité multifactorielle souvent appelé productivité totale des facteurs (PTF). La productivité partielle est mesurée par le rapport entre une production et un facteur de production, presque toujours le travail. L'évaluation de ce rapport, entre deux périodes, caractérise l'évolution de la productivité du facteur considéré. Ce type de mesure a été critiqué parce qu'il reporte sur un facteur (souvent le travail) la contribution des autres inputs. Pour pallier à ce défaut, la PTF, développée par Solow (1957), mesure le rapport entre la production et l'ensemble des

inputs utilisés. La contribution de Solow a été de définir une mesure agrégée du taux de croissance des inputs en accord avec la théorie économique. Les premières données sur la productivité ont été rassemblées aux États-Unis par le Bureau de la Statistique du Travail en 1891, Par la suite, plusieurs économistes se sont intéressés à mesurer le niveau de la productivité. Par exemple, Tinbergen (1942) a estimé la productivité totale des facteurs à l'aide d'une fonction de production Cobb-Douglas avec des pondérations fixes. Mais jusqu'à la fin des années 50, la mesure la plus fréquemment utilisée était la productivité du travail, qui est définie comme le rapport entre la production et la quantité du facteur travail utilisée, soit une productivité partielle.

Mesurer l'impact des infrastructures sur la croissance de la productivité a été au cœur des préoccupations des économistes. Barro (1990) intègre la dépense publique dans le modèle AK , nommé ainsi à cause de la forme fonctionnelle de la fonction de production : $Y = AK$, où Y représente l'output, A représenté le niveau de technologie, K représente à la fois le capital physique et le capital humain. Ce dernier facteur explique l'impact des infrastructures sur la croissance de la productivité par leurs externalités positives sur le capital privé. Pour Barro, les infrastructures facilitent la circulation des biens, des personnes et de l'information et diminuent le coût de production pour les entreprises privées, etc. D'autres recherches montrent que le ralentissement de la productivité peut, dans une large mesure être expliqué par la réduction des taux d'investissement dans les infrastructures publiques. Parmi les principales contributions, on note : Aschauer (1989), Berndt et Hansson (1991), Gramlich (1994), De Haan, Sturm et Sikken (1996) et Seitz (2001). Cela signifie que la diminution de l'investissement public dans les infrastructures entraîne une externalité négative pour la croissance économique. Réduire l'investissement public amoindrit donc les performances des entreprises, car elles ne profitent plus d'un environnement économique propice permettant leur développement. Toutefois, bien

que le lien entre la croissance économique et les investissements dans les infrastructures soient acceptés, il est important de souligner qu'il existe des opinions divergentes sur l'évaluation quantitative de ce lien. Holtz-Eakin (1994) et Hulten et Schwab (1997) suggèrent qu'il n'y a pas de contributions significatives des infrastructures sur la croissance économique. Bien que cet aspect ne soit pas au centre des préoccupations de notre thèse, il importe de mentionner que la présence du gouvernement ne se fait pas sentir que par ses seules dépenses en investissements. Cette présence est aussi ressentie par la réglementation que le gouvernement a imposée au marché et aux entreprises (Viscusi *et al.*, 2000).

0-2 L'importance des infrastructures en général

Le développement économique se trouve toujours au cœur de la politique industrielle. La politique industrielle vise à mieux promouvoir le développement économique. Afin de parvenir à ce but, nous avons souvent besoin d'identifier les secteurs clés de l'industrie et favoriser le développement de ces secteurs pour leur donner une certaine priorité, comme un avantage politique ou un financement spécifique. Les infrastructures peuvent favoriser le développement de l'industrie. Un investissement trop faible aura un effet limitant sur ce développement alors qu'un surinvestissement représentera un gaspillage de ressource. Bien entendu, comme il s'agit d'un investissement, il est nécessaire de comparer les bénéfices sociaux des infrastructures avec le coût public de ces investissements.

L'approvisionnement des infrastructures peut être l'un des outils les plus utiles pour un gouvernement qui souhaite promouvoir la croissance économique. Non seulement les infrastructures publiques augmenteront directement la production nationale, ne serait-ce que par le jeu de définitions comptables, mais plus important, les infrastructures peuvent aussi avoir un effet d'externalité ou de complémentarité sur l'ensemble des secteurs productifs.

Nous pouvons étudier les différents mécanismes par lesquels les infrastructures peuvent agir sur une économie. Bhatta et Drennan (2003) explique qu'un accroissement des infrastructures sur la croissance économique à long terme aura comme conséquences une augmentation de la production et de la productivité, une augmentation du taux d'emploi et du salaire réel, une augmentation de la qualité de vie, etc.

À notre avis, le mécanisme liant infrastructures et croissance économique peut être résumé par un mécanisme direct et un mécanisme indirect. Parmi ces deux mécanismes, le mécanisme direct se réfère principalement au fait que nous utilisons directement les infrastructures comme un facteur direct dans la production (Aschauer, 1989b ; Holtz-Eakin, 1994 ; Yilmaz *et al.*, 2001). Par exemple, le développement des infrastructures dans les zones reculées va réduire le coût des facteurs intermédiaires de production (les matériaux, l'énergie, la main d'œuvre, l'information, etc.). Il facilite également la transmission des outputs et des services, réduit le coût de production et augmente la productivité des entreprises qui se trouvent dans cette zone (Gannon et Liu, 1997).

Le mécanisme indirect peut être résumé dans les aspects suivants : premièrement, les infrastructures peuvent améliorer la productivité des autres facteurs de production. Donc, elle peut augmenter indirectement l'output total d'une région. Par exemple, l'investissement dans les infrastructures va améliorer les conditions de transport, promouvoir l'output (de la main d'œuvre, des terres, des capitaux physiques) et ainsi accroître la croissance économique dans la région. Deuxièmement, les infrastructures peuvent attirer des ressources d'autres régions pour promouvoir le développement économique local. Une amélioration des infrastructures locales va attirer la main d'œuvre, les capitaux et les technologies afflueront vers cette région. Les infrastructures montrent alors un effet d'agglomération (Hulten *et al.*, 2006). En plus, un grand nombre d'études théoriques et empiriques ont confirmé que les

infrastructures constituaient un facteur très important pour les pays en développement dans la mesure où elles attiraient des investissements directs étrangers (Sun *et al.*, 2002). En même temps, le développement du commerce international dépend également du réseau de transport et des infrastructures de communication. Button (1998) croit que la croissance économique dépend principalement de la réallocation des types de main-d'œuvre et des ressources dans une région. Il peut être intéressant de noter que les nouvelles infrastructures peuvent conduire à une croissance économique dans les zones avoisinantes.

En outre, les infrastructures peuvent permettre de tirer profit des économies d'échelle (Gannon et Liu, 1997). Un réseau de transport de qualité peut aider les entreprises à réduire le prix de leur production en réduisant les coûts de transport, les coûts de production, en améliorant la qualité du transport et en permettant la communication avec des marchés auparavant inaccessibles. Ainsi l'entreprise dans la même région peut profiter des économies d'échelle potentielles pour développer d'autres industries, promouvoir la spécialisation dans l'industrie et augmenter le nombre d'emplois, etc.

Mis à part les recherches des économistes du développement, cela fait seulement une vingtaine d'années que les économistes s'intéressent activement aux impacts des infrastructures sur le développement économique. Aschauer (1989) essaie d'analyser la relation entre la baisse de la productivité et le ralentissement des investissements dans les infrastructures aux États-Unis pendant années 60. Son travail précurseur incite une nouvelle vague de recherche analysant la relation entre infrastructures et développement économique. Par contre, son analyse est concentrée au niveau macroéconomique sans regarder la performance des firmes. La vaste littérature sur le sujet sera présentée dans le chapitre 1, En un mot, l'importance des infrastructures pour une économie peut s'exercer de différentes manières :

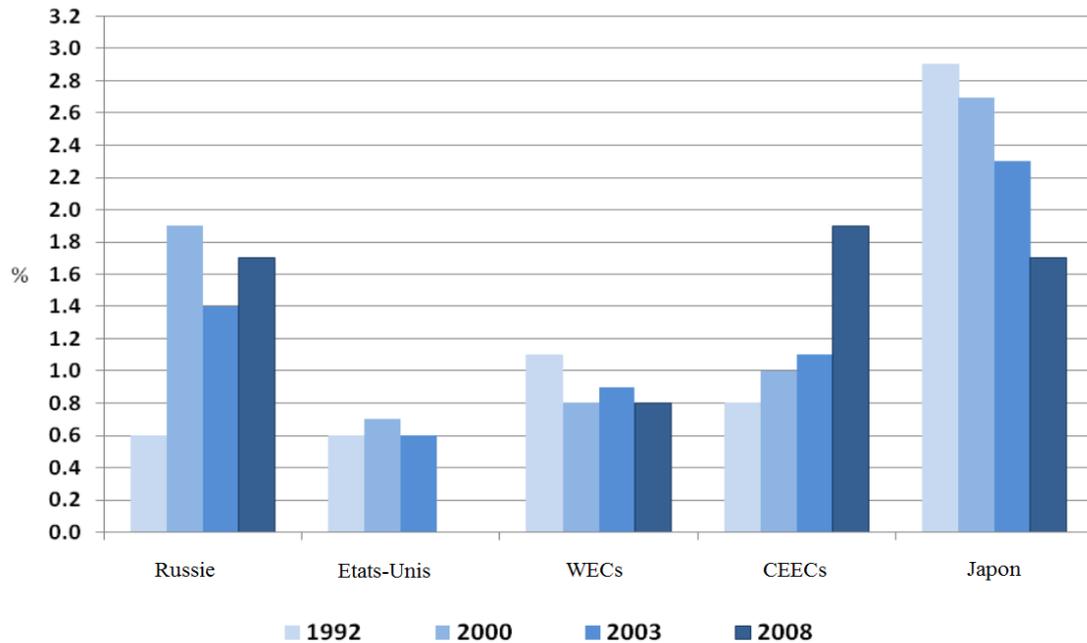
1. L'investissement en infrastructures peut augmenter la demande intérieure.
2. Les infrastructures peuvent être utilisées comme un facteur de production utilisé par les secteurs marchands et participe directement à la réalisation de la production.
3. Les infrastructures peuvent servir à la réduction des coûts de production et rendre ainsi la production plus efficace.

Prenons les infrastructures de transport comme un exemple. L'amélioration des réseaux routiers exige un investissement direct en infrastructures de transport ce qui augmente la demande intérieure et cette amélioration peut contribuer à réduire l'usure des véhicules, à allonger leur durée de vie et à augmenter leur vitesse de déplacement.

Au total, parce que les infrastructures sont inscrites dans l'espace régional, parce qu'elles constituent des éléments indispensables à la fonction des sociétés et des activités, parce qu'elles sont susceptibles d'être utilisées par l'ensemble de la collectivité à des fins diversifiées, les infrastructures peuvent être considérées comme un facteur de développement économique d'un état ou d'une région, aux côtés des conditions naturelles, de la population, de la situation géographique (périphérique ou centrale) et de la structure urbaine.

0-3 L'importance des infrastructures de transport

En reconnaissant l'importance des infrastructures de transport pour le développement économique national, les pays principaux ont beaucoup investi dans les infrastructures de transport. La figure 0-1 nous montre les investissements en infrastructures de transport en pourcentage de PIB pour les grandes économies entre 1992-2008.



Note : CEECs : Les pays européens du centre et de l'est ; WECs : les pays européens de l'ouest.

Source : *International Transport Forum, Investment in Transport Infrastructure*

Figure 0-1 : L'investissement en infrastructures de transport terrestre en pourcentage du PIB en prix courants

Le lien entre les infrastructures de transport et la croissance peut aussi être abordé en étudiant l'accessibilité. Une amélioration des infrastructures de transport va améliorer le niveau d'accessibilité parmi les différents lieux et donc faciliter les interactions spatiales et les activités économiques. Comme la ville, le département, la région sont plus proches grâce à l'amélioration du transport, les endroits reliés entre eux vont d'avoir des relations économiques, politiques et culturelles plus proches, et favoriser ainsi les affaires économiques. La relation peut être exprimée par la figure suivante :

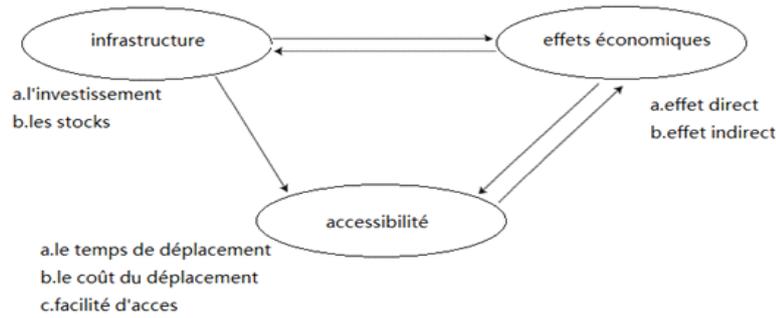


Figure 0-2 : La relation entre les infrastructures de transport et la croissance économique

La plupart des études ont tenté d'analyser quantitativement les impacts des infrastructures sur la croissance économique, sans différencier les types d'infrastructures. Ce faisant, il y a un risque de biaiser les résultats obtenus, car les différents types d'infrastructures influencent la croissance économique à des degrés divers. La littérature récente a commencé à distinguer différentes formes d'infrastructures. Dans la perspective d'une analyse empirique, on décompose le « noyau infrastructure » et, en particulier, on distingue la contribution des infrastructures de transport en termes des caractéristiques du réseau et la dépendance spatiale des diverses entités économiques, etc. En réaction à la baisse de la productivité dans les années 70 aux États-Unis, les chercheurs ont porté une attention particulière aux infrastructures de transport à partir des aspects suivants.

a. La condition nécessaire pour que les infrastructures de transport puissent influencer la croissance économique. Par exemple, la Banque Mondiale (1994) a indiqué que les infrastructures forment une condition nécessaire mais pas suffisante pour influencer la croissance économique. Par conséquent, s'appuyer uniquement sur un investissement en infrastructures ne garantit pas la croissance économique. Banister et Berechman (2001) ont proposé trois conditions nécessaires pour que les infrastructures de transport puisse promouvoir la

croissance économique : les externalités économiques ; les facteurs d'investissement tel que la disponibilité des fonds pour l'investissement ; les effets de réseau et les facteurs institutionnels et politiques. Sans ces trois conditions, les infrastructures de transport ne peuvent pas promouvoir la croissance économique. En l'absence d'une de ces conditions, l'impact des infrastructures devient quasiment nul. Sur cette base, l'impact des infrastructures de transport sur la croissance économique est aussi étudié par rapport aux niveaux locaux, régionaux et nationaux. Ils ont constaté que les investissements en infrastructures de transport peuvent avoir un effet multiplicateur et permettre à des agents économiques de gagner du temps, de réduire les coûts, d'augmenter l'accessibilité au réseau, d'étendre les activités économiques. Il peut aussi y avoir des externalités positives : les infrastructures peuvent générer un effet d'agglomération et un effet de réseau pour accélérer le développement économique régional

b. Analyser l'impact des infrastructures de transport par l'aspect de l'aménagement de l'industrie et le choix de la localisation géographique des entreprises. Par exemple, Buurman et Rietveld (1999) utilisent le système d'informations géographiques pour étudier l'industrie manufacturière de Thaïlande. Ils constatent que, quel que soit l'investissement en infrastructures de transport à court terme ou l'exploitation de ces infrastructures à long terme, un grand nombre d'emplois avait toujours été créés, ce qui offre une possibilité d'augmentation de revenu pour la population. En même temps, les autoroutes et les ports ont un impact positif sur l'aménagement des industries. Holl (2004) considèrent que les décisions de localisation des entreprises dépendent principalement des anticipations de la taille de marchés et du niveau des coûts de production. Les infrastructures de transport déterminent le niveau des coûts de transport des entreprises et, donc, influencent l'accès au marché des entreprises.

Ils utilisent des données portugaises de treize secteurs industriels et neuf secteurs de service pendant la période 1986-1997 et constatent que le réseau d'autoroutes portugais affecte d'une manière significative la localisation spatiale des entreprises. Par exemple, les entreprises préfèrent établir les nouvelles usines proches des nouvelles autoroutes. Fujita *et al.* (1999) ont montré qu'avec la réduction des coûts de transport, la distribution spatiale des entreprises va mettre en évidence un phénomène de diffusion, puis de concentration et finalement, de nouvelle diffusion. Par exemple, Kilkenny (1998) considère que lorsque le coût de transport est relativement important pour les entreprises, elles vont avoir une stratégie de concentration. Les entreprises choisissent alors une localisation plus proche du marché et des fournisseurs intermédiaires. Quand le coût de transport est relativement bas, les entreprises vont utiliser l'avantage lié au fait que le coût de la main d'œuvre est relativement bas. Elles s'étendront ainsi probablement aux zones périphériques. Holl (2004) ont souligné qu'une amélioration des infrastructures de transport contribue à la hausse des concentrations géographiques des industries avec un coût important de transport (comme le secteur de l'alimentation) et une dispersion des secteurs avec un faible coût de transport (comme le secteur mécanique, le secteur des plastiques, etc.). Ce qui indique que la construction de nouvelles routes influencera la répartition spatiale des entreprises manufacturières, mais l'ampleur de son impact varie selon les caractéristiques de l'industrie.

c. Les économistes ont également beaucoup étudié l'impact des différentes composantes des infrastructures de transport sur la croissance économique.

Garcia-Mila et McGuire (1992) ont estimé l'élasticité des autoroutes par rapport à la production, ils ont constaté que l'output économique d'un pays est déterminé par son niveau d'investissement dans les autoroutes. Boarnet (1998) a pris les routes et les autoroutes comme objet et a trouvé que les routes et les

autoroutes montrent un effet positif sur le comté d'origine et un effet négatif sur les comtés voisins. Chandra et Thompson (2000) ont considéré que les investissements en nouvelles autoroutes promeuvent le développement économique pour des régions avec un haut degré d'urbanisation. Par contre, les investissements dans les autoroutes rurales vont favoriser la croissance économique des villes voisines et détriment la croissance économique rurale. Cohen et Monaco (2008) ont trouvé que pour les États-Unis, la croissance économique de chaque état provient principalement d'une augmentation de la fourniture des infrastructures des ports dans son propre état, et non pas d'une augmentation des infrastructures de l'état voisin. Marazzo *et al.*, (2010) a fait une analyse économétrique en utilisant des données brésiliennes de 1996 à 2006, et a trouvé que le transport aérien joue un rôle positif pour la croissance économique. Il a confirmé que le transport aérien montre un effet multiplicateur sur le développement économique. Boopen (2006) a également constaté que le système routier est un élément important pour le développement économique, et son impact se présente par un effet de décalage dans la croissance économique.

En résumé, les chercheurs ont essayé d'estimer empiriquement l'influence des infrastructures de transport du point de vue d'économie du transport, de la nouvelle économie géographique, d'analyse économie régionale et du point de vue macroéconomique. Premièrement, les économistes ont trouvé que l'influence des infrastructures de transport sur la croissance économique a des effets multiples, et cette influence requiert certaines conditions nécessaires pour se produire. Deuxièmement, puisque le choix des échantillons, les sources des données, le choix des variables et même les pays à étudier sont très différents, l'ampleur de cette influence et en particulier, l'élasticité des infrastructures de transport par rapport à la production est très variables.

Une caractéristique commune aux études mentionnées jusqu'à présent est leur aspect « macroéconomique ». L'analyse macroéconomique suppose l'existence d'une fonction de production agrégée. Cette approche essaie de relier un agrégat de l'output de l'économie à des agrégats d'intrants privés (la main d'œuvre, les capitaux, l'énergie, etc.) et des éléments du capital public (les infrastructures). En regardant l'élasticité de l'output par rapport aux infrastructures, nous pouvons mesurer l'impact des infrastructures. Il y a deux effets potentiels. D'une part, si les infrastructures augmentent l'avantage compétitif des entreprises locales, cela attirera l'investissement extérieur ce qui favorise la croissance économique. D'autre part, si l'investissement en capital public est financé par une augmentation des taxes sur les entreprises, ceci diminue la compétitivité locale et risque de ralentir la croissance économique.

0-4 La problématique de la thèse

Les recherches réalisées sur la relation entre les infrastructures et l'activité économique ont été principalement conduites au niveau macroéconomique et de façon paramétrique. On a donc lié productivité et capital public au niveau macroéconomique sans mesurer l'incidence au niveau de la firme. En d'autres mots, malgré l'importance des infrastructures sur la performance individuelle des entreprises, très peu d'études ont été réalisées au niveau microéconomique et les rares qui s'y intéressent ne mettent pas la productivité au centre de l'analyse. Plus précisément, dans toutes les analyses sur ce sujet, les infrastructures sont présentées comme un bien public. La taille de ces infrastructures est déterminée par les instances publiques et les firmes (et le public en général) n'en assument les coûts que par le système fiscal et non en fonction de leur consommation. Lorsque qu'elles sont disponibles, toutes les entreprises en bénéficient. Et outre le fait que ces analyses sont, pour la plupart, au niveau macroéconomique, elles présupposent l'efficacité des firmes, des rendements d'échelle constants, la concurrence et un progrès technologique neutre. Ces conditions imposent alors une forme très particulière à la productivité. En fait dans un tel

contexte on peut montrer que la productivité multifactorielle est égale au changement technologique et ces analyses n'arrivent pas vraiment à capter les effets des infrastructures sur la performance des entreprises.

L'objet de la thèse est de combler le vide méthodologique et empirique des études précédentes. Plus précisément, nous nous attaquerons à la question de la mesure de l'efficacité en généralisant la méthode d'Aigner et Chu (1968) et nous mesurerons l'impact des dépenses publiques en infrastructures de transport au niveau de certaines industries françaises à l'aide de données individuelles de firmes. En particulier, la productivité de l'entreprise sera analysée en tenant compte de son fonctionnement réel afin d'obtenir une décomposition des gains de productivité totaux et d'identifier la partie imputable aux infrastructures. L'approche consiste, dans un premier temps, à décomposer la mesure de productivité multifactorielle à partir de l'approche semi-paramétrique, soit le modèle Aigner et Chu (1968), en introduisant les inputs quasi-fixes. Ce travail méthodologique se fera à la fois dans l'espace primal (la fonction de production) et dans l'espace dual (la fonction de coût). Nous aborderons également la question du choix de la forme fonctionnelle. Dans un deuxième temps, nous montrerons que les méthodes statistiques classiques ne nous permettent pas de tester la significativité des coefficients estimés par la méthode d'Aigner et Chu. Pour résoudre ce problème, dans un premier temps nous démontrerons la convergence des estimateurs obtenues par le modèle d'Aigner et Chu via le théorème de Wald (1949) et ensuite nous proposerons une méthode alternative pour tester les caractéristiques des paramètres obtenus grâce à la généralisation du modèle d'Aigner et Chu en utilisant une méthode de ré-échantillonnage. De la sorte, nous compléterons certains aspects théoriques manquants dans le domaine de mesure la productivité.

Nous testerons aussi nos résultats à l'aide d'une approche non-paramétrique plus traditionnelle, soit le *Data Envelopment Analysis* (DEA). La comparaison des

résultats (modèle semi-paramétrique et modèle DEA) nous permettra d'évaluer l'importance de notre contribution.

0-5 Le plan de la thèse

Nous présentons dans la première partie les aspects de la littérature théorique et empirique auxquels cette thèse se rapporte afin d'aborder les préalables nécessaires à la modélisation et de fournir les résultats de référence. Nous présentons d'abord le développement de la pensée sur la définition des infrastructures, son origine, les différentes définitions utilisées par les économistes et la définition adoptée dans notre thèse. Puis nous recensons les principales contributions empiriques dans la littérature visant à quantifier l'impact des infrastructures sur la productivité. Les études sont aujourd'hui nombreuses sur l'estimation des impacts des infrastructures au niveau macroéconomique mais sont rares au niveau microéconomique. Nous exposons les principales applications empiriques dans ce champ d'étude. Nous passerons ensuite en revue les méthodologies sur l'estimation de l'efficacité des entreprises. La théorie économique dispose de plusieurs méthodes permettant de mesurer et d'analyser la productivité. Dans notre étude, nous développerons un outil méthodologique permettant d'identifier les sources et les déterminants de la productivité des DMU. Pour répondre à ces exigences, nous présenterons aussi dans la première partie une revue de littérature synthétique ayant trait aux principales contributions traitant de la mesure de la productivité en commençant par Farrell (1957) et développée par Aigner, Lovell et Schmidt (1977), Charnes, Cooper et Rhodes (1978), etc. Ceci nous permettra non seulement de montrer l'évolution de la pensée économique en la matière, mais aussi de montrer la limite de ces approches. Le modèle théorique sous-jacent à chacune des approches (paramétrique, non paramétrique et semi-paramétrique) est présenté. Les avantages et inconvénients de chaque méthode sont discutés. Le modèle de base d'Aigner et Chu et la littérature autour du modèle est également introduite.

Dans une deuxième partie, nous adapterons le modèle d'Aigner et Chu (1968) en introduisant les inputs quasi-fixes. Nous transposerons ce modèle dans l'espace dual ce qui permettra d'étudier le comportement des firmes à l'aide de la fonction de coût. Les études empiriques utilisant la méthodologie d'Aigner et Chu ont eu recours à la forme fonctionnelle Cobb-Douglas. Comme il s'agit d'une forme contraignante, en ce sens que les caractéristiques de la technologie ne sont pas libres de prendre des valeurs arbitraires, nous utiliserons aussi une forme flexible de type Translog pour contourner les limites de la forme fonctionnelle Cobb-Douglas.

La méthode d'Aigner et Chu est basée sur des techniques de programmation mathématique, en particulier linéaire, et a toujours été utilisée en tant que méthode non statistique, c'est-à-dire, sans introduire de mesure de dispersion comme des écarts-types. Nous proposerons une méthode d'inférence statistique basée sur le ré-échantillonnage pour tester la significativité des paramètres estimés. Le développement du modèle Aigner et Chu dans le cas dual, l'utilisation des inputs quasi-fixes, l'introduction de l'indice d'accessibilité dans le modèle de frontière, avec, finalement, la construction d'intervalles de confiance pour les coefficients estimés constitueront nos principales contributions méthodologiques.

Dans une troisième partie, nous établissons une application empirique permettant de quantifier les déterminants de la productivité. Particulièrement, cette partie empirique analysera l'impact des infrastructures sur la productivité des entreprises françaises œuvrant dans le transport routier de voyageurs. Une innovation de notre thèse consistera à développer un indice de l'accessibilité au réseau routier afin de l'utiliser comme input dans le processus de production. Nous exposerons d'abord une description statistique de notre base de données et nous présenterons en détail le traitement des infrastructures dans notre modèle. Nous calculerons ensuite l'impact des infrastructures sur la productivité par la méthode de l'analyse par enveloppement

de données (DEA) et notre généralisation du modèle d'Aigner et Chu, tout en testant la significativité des coefficients estimés.

Première Partie.

Une revue de la littérature théorique et empiriques portant sur la mesure de l'effet des infrastructures sur la croissance économique

Chapitre 1. Les infrastructures et la croissance économique

L'objet de ce chapitre est de présenter les recherches réalisées pour mesurer l'impact des infrastructures sur la croissance économique. Dans un premier temps, nous nous attachons à donner une définition économique des infrastructures, puis, dans un second temps, nous présentons un cadre de référence qui nous servira à situer la littérature existante dans ce domaine et les principaux résultats.

1-1 Le développement sur la définition de l'infrastructure

Avant d'entreprendre une étude sur les impacts des infrastructures sur la performance des entreprises, la définition de ce qu'est une infrastructure est nécessaire. Dans la réalité, les infrastructures sont plus faciles à reconnaître qu'à définir. En fait, il n'y a pas de définition communément admise.

Gramlich (1994) insiste sur l'importance de la définition des infrastructures. Il dit qu'il existe de nombreuses définitions possibles des infrastructures en tant que stock de capital. Celle qui a le plus de sens du point de vue économique fait référence à des biens qui ont une grande intensité de capital, produits par un monopole public ou réglementé, comme les routes, le système de traitement d'eau et d'égout, et les systèmes de communication. Mais puisque certaines parties de ces systèmes sont de propriété privée et qu'il est presque impossible de les distinguer de celles qui ont une propriété nationale, il utilise une autre version qui se concentre sur le type de propriété et qui comprend seulement le stock de capital tangible dont la nation est propriétaire.

Au cours de la même année, dans un rapport publié en 1994 sur le développement dans le monde, la Banque Mondiale a publié une étude de la relation entre infrastructures et croissance économique et a discuté des moyens avec lesquels les

pays en voie de développement pouvaient améliorer l'offre et la qualité des infrastructures. Ce rapport définit les infrastructures aux sens économique et social. Elle correspond à la fois à un service et à l'ensemble de facteurs de production mis en œuvre pour les produire. Les infrastructures économiques regroupent les services publics, au sens de la Banque Mondiale (énergie, télécommunications, eau, gaz, collecte et stockage des déchets), les travaux publics (réalisation et entretien de route, barrage, canal, système pour l'adduction d'eau et d'assainissement) et les activités de transport (les voies ferrées entre les villes, le transport civils, les ports, les voies navigables et les aéroports). On retrouve aussi les infrastructures sociales comme les secteurs de la santé et de l'éducation. De la sorte, les infrastructures sont mieux définies, car les services sociaux sont inclus en rapport avec la définition de la géographie, mais il manque toujours la partie équipement qui occupe une grande place dans une économie.

En 2004, Prud'Homme nous donne une autre définition des infrastructures. Il définit les infrastructures selon les six caractéristiques qu'elles doivent satisfaire selon lui : « *capital goods that provide service instead of being directly consumed are lumpy and not incremental, have a long lifespan, are space specific and associated with market failures and consumed by both households and enterprises can be labeled as infrastructure* »¹. Dans son article, il sait que selon cette définition, certaines "infrastructures sociales" comme les écoles, les hôpitaux sont exclus. Donc, cette définition n'est pas aussi générale que la précédente car certains types de biens de capital publics sont exclus.

La même année, Chin et Chou (2004) développent les idées de Hall et Jones (1999), et définissent les infrastructures sociales comme les institutions, les politiques

¹ Prud'homme, R. (2004) "Infrastructure and Development ". In Bourguignon, F. & B. Pleskovic (Eds.) "*Lessons of Experience (Proceedings of the Annual World Bank Conference on Development Economics)* ", Washington : World Bank and Oxford University Press, page 155.

gouvernementales et la main d'œuvre associée qui construisent l'environnement économique dans lequel les individus et les entreprises investissent, créent et transfèrent des idées et produisent des biens et services. Mais cette définition va trop loin par rapport à celle qui nous intéresse.

Donc pour nous, en mélangeant les idées des économistes du développement, de la Banque Mondiale, de Prud'homme et d'autres économistes, nous pouvons identifier les caractéristiques que les infrastructures doivent avoir :

1. Elles possèdent un caractère d'immobilité. La non-mobilité se réfère principalement à l'emplacement des infrastructures qui est difficile à changer. Comparées au travail et au capital, les infrastructures, en particulier les biens immobiliers public tels que les barrages, les routes et les autres installations de transport sont généralement attachées à un emplacement particulier. Bien que le développement des technologies informatiques soit très rapide, les communications des utilisateurs du réseau peuvent utiliser les services de communication à un faible coût dans les différentes régions. En général, les infrastructures sont fixes dans un espace, et peuvent être considérées comme des variables exogènes pour cette région.
2. L'investissement dans les infrastructures a un coût d'installation relativement élevé et un coût d'exploitation relativement faible. Les infrastructures impliquent souvent des projets concernant l'économie nationale et le bien-être national, ou au minimum d'une région, comme les routes, les communications, l'énergie et de l'éducation, etc. La construction de ces infrastructures est souvent interdépendante et douées de liens mutuels. C'est-à-dire, une infrastructure est souvent indispensable pour une autre et elle peut en compléter une autre afin de jouer un rôle dans l'économie nationale ou régionale (par exemple, le fonctionnement d'un aéroport aura besoin des routes construits qui peuvent relier cet aéroport et les villes à côté) ; en plus, les investissements initiaux dans les infrastructures sont

souvent trop coûteux. Par conséquent, la construction initiale a besoin d'une intervention gouvernementale. Le gouvernement doit fournir un investissement minimum pour compenser les coûts initial et d'exploitation pour maintenir le fonctionnement des infrastructures à un prix relativement faible.

3. Les infrastructures ont une nature de monopole naturelle. Du point de vue historique, l'attribut de monopole naturel des infrastructures est déterminé par la nature de ses coûts de construction et d'exploitation. Basé sur un coût de construction assez élevé et un coût d'exploitation assez faible, le coût marginal pour fournir une unité de plus d'infrastructures est relativement faible et le coût fixe par unité est relativement élevé. Donc maintenir l'utilisation des infrastructures peut servir à réduire le coût fixe par unité. Plus les infrastructures ont un caractère de réseau, plus elles correspondent à un monopole naturel.
4. Les infrastructures génèrent également des externalités. Les infrastructures offrent des services pour les habitants et la plupart des entreprises d'une région. Par conséquent, les infrastructures auront certains effets externes économiques. Youngson (1967) croit que l'innovation est la clef pour le développement économique régionale et les infrastructures elles-mêmes sont la clef pour l'innovation de production et la réallocation des facteurs pour les nouvelles productions. Hulten *et al.* (2006) ont constaté que les infrastructures régionales montrent un effet externe considérable sur la productivité manufacturière. La construction des infrastructures de transport dans des zones reculées permettra le transfert de cette zone vers des zones à proximité des marchés et contribue à la croissance économique régionale.
5. Les infrastructures ont un caractère public. L'utilisation des infrastructures n'est pas limitée à quelques secteurs ou entreprises ou à quelques agents économiques. Tous les agents économiques dans le pays peuvent profiter des services fournis par les infrastructures installées. En même temps, la plupart des infrastructures

appartiennent à l'état, même s'il y a des infrastructures qui appartiennent à une entreprise privée, la supervision de l'exploitation de ces infrastructures sera très stricte et sujette à réglementation. Cela constitue une différence importante entre les entreprises qui possèdent des infrastructures et les entreprises privées normales.

La conception des infrastructures est assez large et plusieurs types de biens peuvent satisfaire les caractéristiques que nous avons mentionnées. En ce qui nous concerne ici, nous ne retiendrons que les routes et les autres types de moyens facilitant le transport.

1-2 Le développement de la pensée économique

L'importance des infrastructures pour les économistes peut remonter jusqu'à l'époque d'Adam Smith. Contrairement à la perception superficielle communément admise, quand on étudie la conception de l'État d'Adam Smith et que l'on se replonge dans la philosophie générale de son œuvre, nous sommes très loin de la vision selon laquelle l'intervention de l'État est inutile. En fait, la place que Smith attribue à l'État dans la vie économique et sociale est intimement liée à la théorie de la « main invisible ». Selon cette approche, la recherche de l'intérêt individuel est le meilleur moyen d'atteindre l'efficacité économique et l'allocation optimale des ressources, et par là même l'augmentation de la richesse nationale. Pourtant, il ne considère pas que le libre jeu de l'intérêt personnel permette à lui seul le fonctionnement de la société. Il considère, au contraire, que la satisfaction de certains besoins ne peut être laissée à l'initiative individuelle. Nous avons besoin de l'intervention de l'État. Cependant, cette intervention de l'État doit se limiter aux fonctions que le marché ne peut pas prendre en charge. On peut distinguer trois types de fonctions : 1. Assurer la sécurité nationale contre les agressions extérieures : c'est la défense nationale ; 2. Protéger les membres de la société de toute injustice de la part d'un autre membre de la société : un individu de la société ne peut pas opprimer un autre individu. Pour assurer cette fonction, l'État instaure l'administration de la justice, de la police, etc. L'État veille au

respect des libertés individuelles et du droit de propriété, et peut donc assurer la sécurité des biens et des personnes ; 3. Prendre en charge les dépenses nécessaires à la collectivité mais dont la rentabilité est insuffisante : construction d'infrastructures dont les ponts, les routes, les ports, l'éducation de la population. Dans cette troisième fonction, comme ce qui est écrit dans la « Richesse des Nations » (1776, page 375) : « *le troisième et dernier des devoirs du souverain ou de la République est celui d'élever et d'entretenir ces ouvrages et ces établissements publics dont une grande société retire d'immenses avantages* », on trouve l'affirmation d'une certaine conception des services publics à fournir dans l'intérêt de la société, soit les infrastructures.

Mais pendant une très longue période, les économistes ont porté très peu d'attention au rôle joué par les infrastructures dans une économie. Ce n'est qu'au début des années 40, dans le domaine de l'économie du développement, que les économistes commencent à faire attention aux infrastructures. Dans son article "Problems of Industrialization of Eastern and South- Eastern Europe", Rosenstein-Rodan (1943) propose la notion de "*Social Overhead Capital*" (SOC). Pour lui, l'accumulation de capital fixe pour une société avant son investissement dans les industries générales est une condition nécessaire pour le décollage économique. Puis, quand Prebisch (1950) évalue la théorie de croissance équilibrée de Rosenstein-Rodan, il met un accent particulier sur l'importance des différences régionales sur l'impact de la distribution de l'investissement dans les infrastructures : il y a beaucoup de facteurs incontrôlables qui posent d'énormes difficultés pour mesurer l'impact de l'investissement dans les infrastructures dans les différentes régions.

Rostow (1956) utilise lui-aussi le terme SOC au lieu du mot « infrastructures ». Pour lui, le SOC est une condition nécessaire mais pas suffisante pour un décollage économique. Par la suite, Hirschman (1958) réalise une analyse approfondie sur les infrastructures et il précise la notation de SOC. Pour lui, le SOC est généralement

défini comme les services essentiels qui peuvent assurer la bonne marche des secteurs primaires, secondaires et tertiaires. Le SOC en général inclut la loi, l'éducation, la santé, l'hygiène publique et le service public comme : le transport, les télécommunications, l'énergie, l'irrigation, etc. Toutefois, il croit également que le cœur du SOC peut se limiter au transport et à l'énergie. L'économiste P.H. Cootner (1964) croit que les infrastructures ne sont pas une condition nécessaire pour le développement économique, elle est simplement une façon ou un chemin intermédiaire qui nous dirige vers une croissance économique. Sans se prononcer sur le caractère nécessaire et/ou suffisant, la Banque mondiale en 1994 conclut que les infrastructures ont un impact significatif sur la croissance économique et le bien-être des humains.

Sachant qu'il existe une relation entre les infrastructures et le développement économique (même si la causalité n'est pas claire), les économistes ont essayé de mesurer l'impact des infrastructures sur la croissance économique. Avec le développement de bases de données, les chercheurs ont maintenant les moyens de tester empiriquement cet impact. À la fin des années 80 et au début des années 90, la théorie du développement durable se forme progressivement. Les recherches se concentrent sur les domaines suivants: 1. L'effet (positif) des infrastructures sur la productivité (implicitement, il y a l'influence des infrastructures sur le coût) ; 2. L'influence de la distribution des infrastructures sur la différence de la croissance économique régionale ; 3. L'augmentation de la compétitivité d'un pays pour attirer l'investissement direct étranger par une amélioration des infrastructures ; et 4. La possibilité que les infrastructures puissent favoriser le développement de marchés. Diverses recherches empiriques ont été réalisées pour démontrer les effets positifs des infrastructures. Toutes ces recherches ont commencé par la question posée par Aschauer en 1989 : les infrastructures exercent-elles un effet positif sur la croissance ?

En fait, les travaux empiriques montrent une grande homogénéité au niveau de la modélisation. L'outil de base est la fonction de production qui relie les quantités d'inputs et le niveau de production. Essentiellement, les distinctions entre les diverses études portent sur la mesure de l'output ou des inputs et sur le niveau d'agrégation des données. Nous présenterons ici un cadre théorique extrêmement simple qui permet de catégoriser les diverses études portant sur les infrastructures.

Sans infrastructures, la liste des inputs est représentée par le vecteur \mathbf{X} alors que l'output est noté y (que nous supposons unique). La frontière de l'ensemble des possibilités de production est représentée par la fonction de production f telle que $y = f(\mathbf{X})$. La présence d'infrastructures implique l'introduction d'un input ou d'un vecteur d'inputs supplémentaires, notés Inf . On obtient $y = f(\mathbf{X}, Inf)$. À partir de la fonction de production, nous pouvons calculer le taux de rendement implicite des infrastructures en prenant la dérivée logarithmique de la fonction de production par rapport à la quantité d'infrastructures :

$$\xi_{Inf,y} = \frac{\partial \ln y}{\partial \ln Inf} = \frac{\partial f}{\partial Inf} \frac{Inf}{y}, \quad (1-1)$$

où $\partial f / \partial Inf$ est le prix implicite des infrastructures relatif au prix de l'output. Le numérateur est donc le revenu marginal associé à une variation de taille des infrastructures.

La question revient à mesurer ce taux de rendement et à vérifier s'il est significativement positif. Naturellement, une valeur positive est une condition nécessaire pour justifier les investissements publics en infrastructures. Encore faut-il démontrer que cette valeur est suffisante pour compenser leur coût. En fait, un taux de rendement positif correspond à un effet positif des infrastructures sur l'output des secteurs et des économies nationales qui bénéficient de ces investissements.

Il est possible d'utiliser une représentation duale de la technologie. A la place de la fonction de production, nous pouvons opter pour la fonction de coût. Nous obtenons :

$$C = \mathbf{W}' \mathbf{X} = C(\mathbf{W}, Inf, y) \quad (1-2)$$

où C est le coût de production et w est le vecteur de prix des inputs privés. En utilisant les propriétés de la fonction de coût, on peut montrer que le taux de rendement des infrastructures est donné par :

$$\xi_{Inf,y} = - \left(\frac{\partial \ln C}{\partial \ln y} \right)^{-1} \frac{\partial \ln C}{\partial \ln Inf}. \quad (1-3)$$

Un taux de rendement positif implique une baisse des coûts de production, une fois la présence des rendements d'échelle prise en considération. Rappelons que $(\partial \ln C / \partial \ln y)^{-1}$ est la mesure des rendements d'échelle calculés à partir de la fonction de coût.

Les études macroéconomiques retiendront comme mesure de l'output, le PIB, total ou privé, de même que des indices agrégés pour le travail et le capital. Si cet aspect est plutôt standard, la mesure de l'input « infrastructures » est plus diversifié. Bien entendu, toutes ces études sont limitées par la disponibilité des données, particulièrement pour les infrastructures dont le stock (la quantité résiduelles nettes des investissements passés) n'est pas toujours mesuré par les organismes statistiques nationaux. Du côté des études microéconomiques, la seule différence tient au fait que les données sont, ou sectorielles ou individuelles.

En résumé, bien que Smith fût le premier à mentionner l'importance des infrastructures, ces dernières n'ont pas retenu l'attention des économistes pendant une longue période. Les économistes se questionnant sur les conditions de développement ont abordé la question mais ont élargi la définition pour englober les éléments de réglementation, de santé et d'éducation. Avec le développement des données, les chercheurs commencèrent à tester empiriquement l'impact des infrastructures sur

l'économie. Les premiers économistes cherchant à mesurer l'impact des infrastructures sur la croissance économique ont porté leur attention essentiellement sur les domaines suivants : l'accumulation de la richesse, le commerce extérieur, le développement des affaires, le transport maritime, les ports et les autres installations qui leurs sont associées. Parmi les résultats, on note une réduction des coûts d'expédition et de la capacité de transport, la spécialisation dans le transport et une amélioration de l'efficacité productive; Dans ce contexte, l'État doit établir et maintenir des services publics; les infrastructures de transport grâce à la réduction de coût affecteront la distribution spatiale des agents économiques.

1-3 Revue de littérature sur des recherches empiriques

Dans la littérature, il y a principalement quatre méthodes utilisées pour mesurer les impacts des infrastructures sur la croissance économique : par la fonction de production ; par la fonction du coût ; par l'auto régression vectorielle et les variantes de l'approche par la fonction de production.

Comme ce qui est présenté ci-haut, toutes ces recherches peuvent se classer selon le niveau des données utilisées : les recherches macroéconomiques utilisent des données au niveau national et les recherches microéconomiques utilisent des données au niveau plus fin (au niveau régional, industriel, etc.). Il y a un consensus que les données au niveau industriel sont préférables aux données au niveau national, mais selon la disponibilité des données, il y a tout de même beaucoup de recherches réalisées au niveau macroéconomique. Dans cette section, nous présenterons une revue de la littérature portant sur les recherches empiriques pour mesurer les impacts des infrastructures sur la croissance économique en fonction de la méthode utilisée (fonction de production ou fonction de coût) et en fonction des données utilisées (national ou régional, industriel, sectoriel, etc.).

1-3-1 Utilisation d'une fonction de production au niveau macro-économique

Parmi les économistes qui ont travaillé avec la fonction de production au niveau macroéconomique, on retrouve Aschauer (1989), Barro (1990), Munnell (1990a), etc. Les études de ce type conduites à une échelle nationale concluent que le capital public exerce une influence déterminante sur la production, la productivité et le taux de croissance. L'impact des infrastructures de transport en particulier apparaît très important.

Aschauer (1989a) examine la relation entre la productivité agrégée et les dépenses publiques. Les résultats empiriques indiquent que (i) le stock de capital public non militaire est considérablement plus important dans la détermination de la productivité que le flux des dépenses militaires ou non militaires, (ii) le capital militaire a peu de rapport avec la productivité, et (iii) le «noyau» des infrastructures, comme les rues, routes, aéroports, transports en commun, les égouts, les systèmes d'eau, etc., a un impact plus important sur la productivité. En estimant la fonction de production avec des données américaines entre 1949 et 1985, il trouve qu'une augmentation de 10 pourcent en stock de capital public augmentera de presque 3,9 pourcent la productivité multifactorielle. Dans son article, Aschauer introduit une constante et une variable de tendance comme approximation de l'indice de productivité. Le taux d'utilisation des capacités est ajouté pour contrôler l'influence du cycle économique. De nombreux articles subséquents ont utilisé cette spécification ou une spécification similaire. Munnell (1990a) estime une fonction de production pour les États-Unis et sur des séries temporelles couvrant la période de 1949 à 1987. Ses résultats font apparaître une forte élasticité de la production marchande, de la productivité du capital privé et de celle du travail par rapport au stock de capital public. Cette élasticité est encore plus forte par rapport aux infrastructures à finalité économique. Il déduit de ces estimations un taux marginal de rentabilité des investissements dans les infrastructures très élevé, évalué à 60 %, qui seraient deux fois plus élevé que le taux

de rentabilité du capital, estimé à 30 %. Après avoir recalculé au moyen des coefficients estimés de la fonction de production les taux de croissance annuels de la productivité globale des facteurs, il met en évidence une forte corrélation entre le ralentissement des gains de productivité et le déclin des investissements en infrastructures.

Barro (1990) introduit les dépenses publiques du gouvernement comme un bien public dans le processus de production et considère les infrastructures comme une contrainte sur le progrès technique. Selon Barro la production d'une firme i est représentée par une fonction de production de type Cobb-Douglas de la forme suivante :

$$y = Al^{1-\beta}k^{\beta}Inf^{\gamma}, \quad (1-4)$$

où y correspond au PIB ; l représente la quantité des mains d'œuvre ; k représente le stock de capital privé ; β correspond à l'élasticité de la production par rapport au stock de capital privé ; γ correspond à l'élasticité de la production par rapport aux investissements publics et Inf désigne les dépenses gouvernementales agrégées. Barro conclut qu'une augmentation des dépenses publiques conduira à une augmentation de la productivité des facteurs et favorise ainsi l'accumulation du capital privé. La dépense publique a deux effets opposés. Le premier est que le capital public rend le capital privé plus productif et évite que sa productivité marginale s'annule progressivement quand le revenu augmente. Cependant, l'impôt a un effet dépressif sur cette productivité, puisqu'il réduit son rendement privé en ôtant aux entreprises une part du revenu tiré de leur activité. Il montre que pour une petite taille de gouvernement (des dépenses publiques), le premier effet l'emportera. Puis, il montre qu'on peut déterminer une dépense publique optimale. À ce point, un dollar de dépenses publiques supplémentaire coûte plus en productivité que ce qu'il rapporte. Barro fait remarquer que, si les dépenses qui représentent une prestation de services aux consommateurs ont un impact en termes d'utilité, elles découragent la croissance.

En effet, elles n'accroissent pas la productivité du capital privé, mais leur financement par l'impôt pèse sur la rentabilité du capital. Seul l'effet négatif joue au niveau des entreprises. L'imposition liée à ces dépenses décourage l'épargne et l'investissement.

Un inconvénient avec cette approche est que le travail et le capital sont traités comme étant exogènes, alors qu'il est implicitement supposé que les deux facteurs sont payés en fonction de leur productivité marginale. Certaines études ont utilisé une fonction Translog, qui est plus souple que la fonction de Cobb-Douglas (Charlot et Schmitt, 1999 ; Canning et Bennathan, 2000 ; Albal-Bertrand et Mamatzakis, 2004 ; Everaert et Heylen, 2004). Un problème majeur dans l'estimation d'une fonction de production est que la causalité n'est pas très claire. Si les investissements en capital, définis comme suit :

$$I_t - \delta K_t = \Delta K_t, \quad (1-5)$$

dépendent du revenu par une fonction d'épargne :

$$I_t \equiv S_t = sY_t. \quad (1-6)$$

nous pouvons écrire :

$$\Delta K_t = sY_t - \delta K_t, \quad (1-7)$$

où δ est le taux de dépréciation. À l'état stationnaire nous avons :

$$K_t = \frac{sY_t}{\delta}. \quad (1-8)$$

Cela implique un effet de retour du revenu sur le stock du capital public, ce qui rend difficile d'identifier les résultats des régressions. Faire face à ce problème a été au cœur de la controverse sur la relation entre les infrastructures et la croissance économique. Diverses approches ont été proposées dans la littérature pour traiter le problème de la causalité. La première consiste à trouver un test pour vérifier le sens de la causalité. D'autres approches ont été suivies comme des estimations en panel, l'estimation du modèle à l'aide d'équations simultanées et l'utilisation des variables instrumentales. Parmi les chercheurs qui essaient de résoudre ce problème de

causalité, nous pouvons mentionner Canning et Bennathan (2000) qui soutiennent que l'approche sur les données panel peut résoudre le problème de la causalité ; Demetriades et Mamuneas (2000) qui estiment un système d'équations qui est dérivé d'un cadre de la maximisation du profit intertemporel ; et Calderón et Servén (2002) qui ont utilisé la méthode des variables instrumentales pour résoudre le problème de causalité. En utilisant la technique de panel de cointégration, Bronzini et Piselli (2009) montrent que le capital humain et les infrastructures « causent » au sens de Granger, la productivité à long terme alors que l'inverse n'est pas vrai.

1-3-2 Utilisation d'une fonction de production au niveau micro-économique

Les recherches empiriques en utilisant la fonction de production ont été effectuées à la fois au niveau national et au niveau régional. La plupart des études sur le niveau régional ont été, encore une fois, réalisées aux États-Unis. Elles concluent à un impact plus réduit mais toujours significatif du capital public sur les niveaux de production et de productivité. Par rapport aux estimations réalisées sur une base nationale au moyen de séries temporelles, ces travaux sont conduits le plus souvent sur des données de panel : individuelles-temporelles. Les individus, ici des régions ou des agglomérations, sont observés sur plusieurs années. Ceci présente un avantage puisque les relations peuvent être économétriquement mieux établies dans la mesure où les observations sont plus nombreuses et plus dispersées ; mais on a besoin de méthodes économétriques plus avancées, et ce n'est pas toujours évident de trouver une méthode pertinente.

Le tableau 1-1 résume les principales caractéristiques et les résultats des études passées en revue ci-dessus et d'autres études basées sur l'approche de la fonction de production au niveau national, et le tableau 1-2 résumant ceux au niveau microéconomique.

Tableau 1-1 : Études empiriques utilisant la fonction de production au niveau national

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Aschauer (1989)	États-Unis	1949-1985	Fonction de production Cobb-Douglas, une constante et une variable de tendance. rendements d'échelle constante (REC)	Le stock de capital public (infrastructures non militaire)	L'élasticité est 0,39
Munnell (1990b)	États-Unis	1949-1987	Fonction de production Cobb-Douglas ; REC pour tous les inputs ; REC pour le capital privé (en excluant le capital public) ;	Le stock de capital public (infrastructures non militaire)	L'élasticité est 0,34-0,41 ; taux marginal de rentabilité des investissements dans les infrastructures est 0,6.
Barro (1990)	96 pays	1970-1985	Fonction de production Cobb-Douglas avec REC	L'investissement en infrastructures	La dépense publique a deux effets opposés
Attaran et Auclair (1990)	États-Unis	1965-1985	Fonction de production Cobb-Douglas ;	Le capital routier autoroutier ; les autres infrastructures	L'élasticité est 0,22-0,24 pour les infrastructures routières ;
Duggal <i>et al.</i> (1999)	États-Unis	1960-1989	Intègre les infrastructures comme une contrainte technologique dans la fonction de production	Le stock de capital public	L'élasticité est 0,27
Canning et Bennathan (2000)	62 pays	1960-1990	Fonction de production Cobb-Douglas et Translog, plusieurs genres d'infrastructures	Nombre de téléphones, la capacité générale d'électricité, les routes pavées	Les pays avec un revenu faible ou intermédiaire bénéficient des infrastructures

Tableau 1-1 (suite) : Études empiriques utilisant la fonction de production au niveau national

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Calderón et Servén (2002)	101 pays	1960-1997	Fonctions de production Cobb-Douglas et Translog, plusieurs genres d'infrastructures	Transport, communication	Une élasticité de 0,16
Kamps (2004)	22 pays de l'OCDE	1960-2001	Même modèle qu'Aschauer, avec des données de panel	Le stock de capital public	Une élasticité de 0,22 avec données de panel
Duggal <i>et al.</i> (2007)	États-Unis	1975-2001	Fonction de production non linéaire, avec les infrastructures publiques et privées	Le stock de capital public (autoroute, route, préfecture, etc.) ; infrastructures privées (les câbles, la wifi, les équipements d'«Internet, etc.)	Des rendements d'échelle croissants pour l'ensemble d'économie américaine
Del Bo et Florio (2008)	Union Européenne	1995-2005	Fonction de production Cobb-Douglas avec indicateur d'infrastructure, et un modèle spatiale	Infrastructures de télécommunication, le temps d'arriver à certaine marché, la surface d'une région, etc.	Une élasticité de 0,15.
Calderón et Servén (2011)	88 pays	1960-2000	Fonction de production agrégée à long terme, capital humain, le capital physique et l'infrastructure	Infrastructures en quantité pas en valeur, (un indice synthétique)	Une élasticité entre 0,07 et 0,1

Tableau 1-2 : Étude empiriques utilisant la fonction de production au niveau régional

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Charlot et Schmitt (1999)	22 régions françaises	1982-1993	Fonction de production Cobb-Douglas et Translog avec trois facteurs (capital privé, emploi et capital public)	Le stock de capital public	Une élasticité de 0,3 pour Cobb-Douglas et 0,4 pour Translog, mais très sensible aux différentes régions et différentes périodes
Munnell (1990b)	48 régions américaines	1970-1986	Fonction de production Cobb-Douglas	Le stock de capital public	Une élasticité de 0,15
Holtz-Eakin et Schwartz (1995)	États américains	1971-1986	Une modèle de croissance néoclassique	Transportation et communication	Effet négatif
Seitz (1993)	99 villes allemandes	1980-1989	Fonction de production Cobb-Douglas et Translog	Le stock de capital public	Une élasticité de 0,08 à 0,19
Picci (1995)	20 régions italiennes	1970-1991	Fonction de production Cobb-Douglas et effets aléatoires	Le stock de capital public	Une élasticité de 0,08 à 0,43 à court terme, non significative à long terme
Prud'homme . et Fritsch (1997)	Régions française (hors Ile-de-France)	1970-1990	Fonction de production Cobb-Douglas, la croissance de PTF	Le stock d'infrastructures routier régional	Une élasticité de 0,08
Everaert et Heylen (2004)	Régions belges	1965-1996	Fonction de production Translog, un modèle d'équilibre général,	L'investissement public	Une élasticité de 0,31

Tableau 1-2 (suite) : Études empiriques utilisant la fonction de production au niveau régional

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Cadot <i>et al.</i> (2006)	Régions françaises	1985-1992	Fonction de production Cobb-Douglas avec équation politique pour les infrastructures de transport	Le stock des infrastructures de transport	Une élasticité de 0,08
Bronzini et Piselli (2009)	Régions italiennes	1980-2001	Fonction de production Cobb-Douglas, le taux de croissance de PTF, avec R & D et infrastructures publiques	Le flux d'investissement en infrastructures publiques	Une élasticité de 0,109

1-3-3 Par la fonction de coût agrégée

Il y a de nombreux inconvénients dans les modèles de fonction de production. La fonction de production ignore le rôle des prix des facteurs et de l'output dans le processus de décision pour une entreprise. La fonction de production ne prend en considération que la relation technologique alors que les modèles comportementaux de minimisation des coûts et de maximisation des profits indiquent que ces variables sont simultanément endogènes et dépendent des prix. Dans ce contexte, définir le caractère endogène ou exogène d'une variable au sein d'une fonction de production devient problématique. Pour cette raison, l'approche duale est préférable pour estimer l'impact des infrastructures. Dans cette approche, nous supposons que le capital public est fourni gratuitement par le gouvernement comme un input exogène. Les économistes spécifient une fonction de coût pour le secteur privé, en supposant que les entreprises visent à produire un niveau d'output donné avec un coût de production minimum. La fonction de coût peut être exprimée comme :

$$C = C(w_l, w_p, B, y, Inf). \quad (1-9)$$

Ici C représente le coût total de production, w_l est le prix de la main d'œuvre, w_p est le prix du capital privés, B représente le niveau technique, y l'output final et Inf le service des infrastructures.

Le problème pour l'entreprise est :

$$C = C(w_l, w_p, B, y, Inf) = \min_{l,k} \sum w_l l + w_p k, \quad (1-10)$$

$s.c : y = f(l, k, B, Inf).$

où l est la quantité de main d'œuvre et k est le stock de capital.

Si nous supposons que la fonction de coût est continûment différentiable, nous pouvons déduire le prix implicite des infrastructures Inf , w_{Inf} . Ce prix implicite des infrastructures reflète la réduction des coûts de production attribuable à un accroissement dans les investissements en infrastructures. Nous pouvons calculer ce prix implicite comme suit :

$$w_{Inf} = - \frac{\partial C(w_l, w_p, B, y, Inf)}{\partial Inf}. \quad (1-11)$$

Le service des infrastructures (ou le stock de capital public) entre dans la fonction de production comme un input quasi-fixe, aussi appelé input non discrétionnaire. Bien que le stock d'infrastructures soit considéré comme un input extérieur donné, chaque entreprise doit encore décider du montant d'infrastructures qu'elle souhaite utiliser. Cela implique que l'utilisation des infrastructures fait partie de son problème d'optimisation, qui à son tour, conduit à la nécessité d'une fonction de demande pour des infrastructures qui doit satisfaire les conditions sur la productivité marginale (Duggal *et al.*, 1999). En dehors de l'accent mis sur les effets directs dans l'approche de la fonction de production, les infrastructures peuvent également d'avoir effets indirects. Les entreprises pourraient ajuster leur demande d'inputs privés si les infrastructures sont un bien substitut ou un bien complémentaire à ces autres facteurs de production.

Tableau 1-3 : Études empiriques utilisant la fonction de coût au niveau macroéconomique

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Mamatzakis (1999a)	Grèce	1959-1990	Fonction de coût Translog	Le stock d'infrastructures (transports, communications)	L'élasticité des infrastructures varie de 0,02 à 0,78 selon différents secteurs
Demetriades et Mamuneas (2000)	12 pays de l'OCDE	1972-1991	Fonction de coût quadratique	Le stock de capital public	L'élasticité de l'output varie de 2,06 (Norvège) à 0,36 (Angleterre)
Cohen et Morrison Paul (2004)	États américains	1982-1996	Fonction de coût Leontief généralisée	Le stock d'autoroutes	L'investissement en infrastructures réduit le cout du même État et augmente celui de l'État voisin.
Albala-Bertrand et Mamatzakis (2007)	Chili	1960-2000	Fonction de coût Translog	Le changement du taux de croissance, principales composantes des infrastructures (Transports, communications)	Une hausse des infrastructures induit une diminution du coût total et améliore modérément la productivité de l'économie chilienne

Tableau 1-4 : Études empiriques utilisant la fonction de coût au niveau microéconomique

Auteur et Année	Zone d'étude	Période	Forme fonctionnelle	Variable d'infrastructure	Conclusion
Canaleta <i>et al.</i> (1998)	Régions espagnoles	1964-1991	Fonction de coût flexible	Le stock de capital public et les infrastructures de transport	Le capital public réduit le coût de production privé, le capital public et le capital privé sont des biens complémentaires et il existe l'effet externe
Boscá <i>et al.</i> (2000)	Régions espagnoles	1980-1993	Fonction de coût Leontief généralisée	Le stock d'infrastructures (transports, communications, ports)	Une élasticité de 0,08
Moreno <i>et al.</i> (2003)	Espagne, régions et secteurs	1980-1991	Fonction de coût Translog	Le stock d'infrastructures (transports, communications, ports)	L'investissement public et privée augmente l'efficacité
Teruel et Kuroda (2005)	Philippines, secteur agricole	1974-2000	Fonction de coût Translog	Le stock de capital public	Les infrastructures contribuent à la croissance de la productivité. L'importance est moindre dans les années 80 mais remonte dans les années

1-3-4 Les autres approches

En dépit de certaines contraintes statistiques, économétriques ou encore théoriques, l'analyse de la productivité du capital public est resté un domaine de recherche actif. La plupart des études récentes utilisent une approche basée sur les vecteurs autorégressifs (VAR). Par rapport aux approches basées sur de la fonction de production et la fonction de coût, elle n'impose pas de liens de causalité entre les variables sous enquête. Dans un modèle VAR, toutes les variables sont déterminées conjointement sans hypothèse a priori sur la causalité. Donc le modèle VAR permet de tester si la relation causale supposée dans d'autres approches est valide ou s'il y a des effets de rétroaction de l'output aux infrastructures. Selon Sims *et al.* (1990), un modèle VAR peut être simplement estimé par la méthode MCO qui donne une estimation cohérente et asymptotiquement normale.

La procédure la plus courante et la plus efficace utilisée pour l'approche d'auto-régression vectorielle est le test de Johansen (1995) employés par Lau et Sin (1997). Lau et Sin arrivent à la conclusion que le capital public n'est pas aussi productif que démontré par Aschauer et beaucoup de ses partisans. Selon les résultats obtenus par Lau et Sin, l'élasticité de production estimée par rapport à des capitaux publics sur la période 1925-1989 aux USA n'est que de 0,11, ce qui est plus faible que les résultats des études de régression simple. Pereira et Andraz (2012) utilisent l'approche VAR pour analyser les effets régionaux de l'investissement public dans le secteur du transport au Portugal. Ils ont trouvé que l'investissement public est un outil puissant pour encourager la croissance économique à long terme, par contre les effets régionaux sont déséquilibrés. C'est la plus grande région, Lisbonne, qui en bénéficie le plus.

Le principal inconvénient de cette approche est qu'elle nécessite généralement de grands échantillons. Pour cette raison, la plupart des chercheurs utilisant l'approche VAR ont utilisé les données sur l'investissement public au lieu de données sur le stock de capital public. Ce choix a été dicté non seulement par le manque de données sur le

stock de capital pour un grand nombre de pays, mais aussi par le fait que les données sur les investissements publics sont généralement disponibles chaque trimestre, alors que les données de stock de capital public ne sont disponibles que chaque année. Un inconvénient de ce choix est l'hypothèse implicite que les effets de l'investissement public sont indépendants du niveau du stock de capital correspondant. La théorie économique suggère que cette hypothèse est douteuse. Conformément à la loi des rendements d'échelle décroissants, une augmentation du stock de capital public aurait un effet négatif sur l'effet marginal des investissements publics. Il y a des raisons de croire au déclin rapide de la productivité marginale du capital public (Demetriades et Mamuneas, 2000).

Conclusion

Dans la littérature, bien qu'il n'y ait pas unanimité, les auteurs trouvent qu'il existe un consensus dans les années récentes à l'effet que le capital public favorise la croissance économique. Toutefois, l'impact indiqué par de récentes études n'est pas aussi grand que certaines études antérieures et les résultats varient selon le niveau des données qu'on utilise. L'approche macroéconomique démontre un impact positif des infrastructures sur la croissance économique, par contre, les résultats sont moins claires au niveau microéconomique. Ce qui est tout à fait normale, car l'influence des infrastructures comme les autoroutes sur les entreprises de transport est bien évidemment plus forte que sur d'autres entreprises. En plus, dans les études réalisées, les chercheurs n'ont pas regardé le fonctionnement des firmes, et supposent implicitement l'efficacité de ces dernières. Mais les études sur la productivité des firmes ont démontré que l'inefficacité était fréquente, ce qui invalide les études ne permettant pas le calcul des inefficiences. Donc dans notre thèse nous essayerons de mieux mesurer l'effet des infrastructures sur la performance des entreprises en prenant en considération la présence éventuelle d'inefficacités.

Chapitre 2. Mesure de l'efficience des entreprises : une revue et une synthèse méthodologique

Le défaut des approches présentées au premier chapitre (l'hypothèse d'efficacité des entreprises et son influence sur la performance des entreprises n'est pas claire) nous incite à utiliser une approche permettant de mesurer l'efficacité des entreprises. Cette approche relâche la contrainte d'efficacité des firmes, les entreprises peuvent bien fonctionner au-dessous de la frontière de production. La performance de chacune des entreprises sera au centre de l'analyse. Ce chapitre porte sur les différentes méthodes et approches afin de mesurer l'efficacité des unités de décision (« Decision Making Units » en anglais et nous utiliserons l'acronyme DMU par la suite). Dans un premier temps, nous définirons la productivité et l'efficacité et présenterons quelques contributions plus importantes dans ce domaine. Dans un deuxième temps, les modèles de frontière de production seront classifiés selon deux critères principaux : la spécification de la frontière (paramétrique ou non paramétrique) et la présence de bruits dans la procédure d'estimation. Finalement, nous présenterons deux approches principales utilisées dans la littérature : l'approche statistique (la frontière stochastique, la méthode des moindres carrés corrigées) ; l'approche basée sur la recherche opérationnelle (le modèle DEA et la méthode d'Aigner et Chu).

2-1 La productivité et l'efficacité

Dans la plupart des cas, les problèmes de définition sont à l'origine des difficultés quand vient le temps d'évaluer un concept économique. Avec une définition confuse, il est très difficile d'obtenir un résultat correct. Il est donc nécessaire de se concentrer sur les définitions de la productivité et de l'efficacité avant de commencer à présenter les méthodes de mesure de ces concepts. Après avoir défini les concepts fondamentaux, les méthodes seront exposées.

2-1-1 Les notations de base

Selon la définition classique, la productivité est le rapport entre un agrégat d'outputs et un agrégat des facteurs de production qui rendent possible cette production :

$$\text{Productivité} = \frac{Y(y_1, y_2, \dots, y_q)}{X(x_1, x_2, \dots, x_p)}, \quad (2-1)$$

où $Y(y_1, y_2, \dots, y_q)$ est une fonction agrégative des outputs et $X(x_1, x_2, \dots, x_p)$ est une fonction agrégative des inputs.

L'utilisation de ce ratio pour mesurer la productivité s'avère facile lorsqu'elles ne s'appliquent qu'à une seule DMU produisant un bien unique à partir d'un seul input de qualité uniforme. Dans un cas plus réaliste où les DMU vont utiliser des facteurs hétérogènes pour produire plusieurs outputs, la façon de définir $Y(y_1, y_2, \dots, y_q)$ et $X(x_1, x_2, \dots, x_p)$ est très importante. Le numérateur et le dénominateur doivent être agrégés en conformité avec la théorie économique. Nous pouvons donc distinguer une productivité partielle quand il s'agit d'un seul facteur de production et une productivité totale des facteurs qui s'exprime comme le ratio entre la production et l'ensemble des facteurs de production utilisés lors du processus de production. Par contre, en utilisant ce ratio, nous supposons $Y(y_1, y_2, \dots, y_q)$ et $X(x_1, x_2, \dots, x_p)$ sont identiques pour toutes les DMU, ce qui implique que toutes les DMU ont la même technologie et affrontent le même marché.

Un autre concept qui est relié étroitement avec le concept de productivité est l'efficacité. En fait, dans la littérature, les économistes ne font pas toujours la distinction entre la productivité et l'efficacité. Par exemple, Sengupta (1995) et Cooper, Seiford et Tone (2006) définissent la productivité et l'efficacité comme un ratio de l'output par rapport à l'input. Au lieu de définir l'efficacité comme le ratio entre les inputs et les outputs, nous pouvons le décrire comme une distance entre la quantité observée des inputs et des outputs, et la quantité des inputs et des outputs

situés sur la frontière de production. Une DMU est dite efficace si elle a produit le plus possible d'output avec les inputs utilisés, ou encore, si elle a produit cette quantité d'output avec un coût minimal.

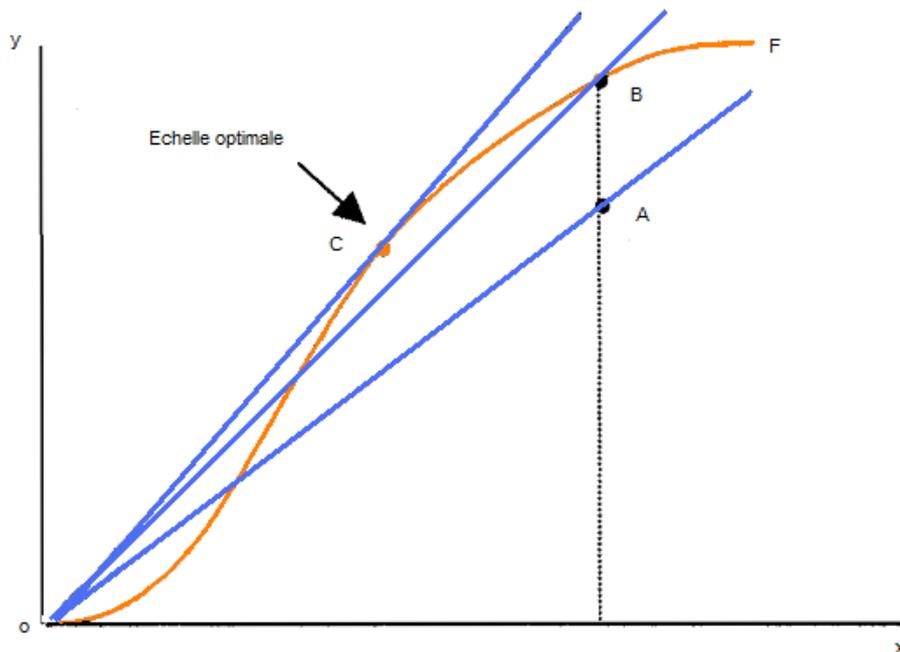
En fait, l'efficacité économique peut être décomposée en efficacité technique et efficacité allocative. L'efficacité technique se réfère à la capacité à éviter le gaspillage, soit en produisant le maximum d'output étant donnée la quantité des inputs utilisés, soit en utilisant le minimum d'inputs pour produire la quantité des outputs désirés. Donc il y a deux orientations d'efficacité : l'orientation outputs qui consiste à maximiser l'output étant donné les inputs et l'orientation inputs qui consiste à minimiser les inputs étant donné un output. L'efficacité allocative se réfère à la capacité de combiner des inputs dans une proportion optimale en prenant en considération les prix. La proportion optimale satisfait la condition du premier ordre pour le problème de minimisation du coût pour une DMU.

La différence entre l'efficacité et la productivité n'est pas très évidente. Nous allons donc ici prendre un exemple pour la montrer. Supposons qu'il y ait un simple processus de production dans lequel un seul input X est utilisé pour produire un seul output Y . La courbe OF dans la figure 2-1 représente une frontière de la production. Cette frontière de production représente la quantité d'output maximale réalisable compte tenu de chaque niveau d'entrée. Par conséquent, elle reflète l'état actuel de la technologie dans l'industrie. Les entreprises dans ce secteur fonctionnent soit sur cette frontière, si elles sont techniquement efficaces, soit sous la frontière si elles ne le sont pas. Donc, une entreprise opérant au point A est inefficace, parce que techniquement, elle pourrait accroître la production au niveau associé à un point B sans avoir besoin de plus d'inputs.

Pour la productivité, la pente de la droite qui passe par le point d'origine et le point observé pour une firme (combinaison input-output) fournit une mesure de la productivité moyenne des entreprises. Si la droite passe du point techniquement inefficace A au point techniquement efficace B , la pente de la droite serait plus

grande, ce qui implique une plus grande productivité au point *B* pour les entreprises en question. Toutefois, en passant au point *C*, qui est le point tangent à la frontière de production d'une droite passant par l'origine, l'entreprise passera à la production avec une échelle optimale, et nous retrouvons les entreprises qui ont une la productivité la plus élevée.

Pour les études économiques, quand nous parlons de la productivité, il s'agit souvent de la productivité totale des facteurs (PTF) qui est la *production réelle* de l'entreprise ou de l'industrie sur une certaine période, divisée par un agrégat des *inputs réels* utilisés par le même ensemble d'unités de production sur la même période.



Source : Coelli et al. (1998, page 5)

Figure 2-1 : Frontière de production et efficacité technique

Koopmans (1951) a proposé une définition de l'efficacité technique : une DMU est techniquement efficace si une augmentation de l'output nécessite une réduction d'au moins un autre output ou une augmentation d'au moins un input, ou si une réduction d'un input nécessite une augmentation d'au moins un autre input ou une réduction

d'au moins un output. Donc une DMU techniquement inefficace peut augmenter ses outputs sans augmenter les inputs ou diminuer ses inputs sans diminuer ses outputs.

Debreu (1951) et Farrell (1957) ont introduit une mesure de l'efficacité technique. Avec une orientation inputs, leur mesure est définie comme le plus grand pourcentage de réduction équiproportionnelle de tous les inputs sans diminuer la quantité d'output avec une technologie donnée. Avec une orientation outputs, leur mesure est définie comme la maximisation de l'expansion radiale de toutes les productions avec la quantité des inputs et le niveau technologique donnés. Dans les deux orientations, si le rapport est égal à un, les DMU sont dites techniquement efficaces. Une valeur différente de un indique que les DMU sont techniquement inefficaces.

Afin de relier les mesures de Debreu-Farrell à la définition de Koopmans, il est utile d'introduire quelques notations sur la technologie. Nous définissons la production comme un processus de transformation d'un ensemble d'inputs, $\mathbf{x} \in \mathbb{R}_+^p$, en un ensemble d'outputs, $\mathbf{y} \in \mathbb{R}_+^q$. Ce processus de transformation est représenté par une fonction de transformation :

$$T(\mathbf{x}, \mathbf{y}) \leq 0. \quad (2-2)$$

Alternativement, on peut le représenter par l'ensemble des possibilités de production, nommée Ψ :

$$\Psi = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in \mathbb{R}_+^p, \mathbf{y} \in \mathbb{R}_+^q, (\mathbf{x}, \mathbf{y}) \text{ est réalisable}\} = \{(\mathbf{x}, \mathbf{y}) \mid T(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}\}. \quad (2-3)$$

Donc l'ensemble des possibilités de production est défini dans une espace Euclidien de dimension $p + q$. Il montre simplement que pour cette DMU, il est physiquement possible de produire un vecteur d'output $\mathbf{y}(y_1, y_2, \dots, y_q)$ en utilisant un vecteur d'inputs $\mathbf{x}(x_1, x_2, \dots, x_p)$.

Il est aussi possible de décrire le processus de production par l'ensemble de besoin en facteurs, $L(\mathbf{y})$ défini comme suit :

$$L(\mathbf{y}) = \{\mathbf{x} \mid \mathbf{x} \in \Psi\} = \{\mathbf{x} \mid (\mathbf{x}, \mathbf{y}) \text{ peut être produit}\} = \{\mathbf{x} \mid T(\mathbf{x}, \mathbf{y}) \leq 0\}. \quad (2-4)$$

Il stipule que tous les éléments de $L(\mathbf{y})$ sont suffisants pour produire le vecteur d'outputs. Jusqu'à maintenant, nous n'avons pas encore défini l'efficacité. Avec $L(\mathbf{y})$, nous pouvons définir un isoquant pour une DMU :

$$I(\mathbf{y}) = \{\mathbf{x} \mid \mathbf{x} \in L(\mathbf{y}), \lambda \mathbf{x} \notin L(\mathbf{y}), \lambda < 1\}. \quad (2-5)$$

Cet isoquant définit en fait la frontière de l'ensemble de besoin en facteurs. De même manière, nous pouvons écrire l'ensemble des productions possibles :

$$P(\mathbf{x}) = \{\mathbf{y} \mid \mathbf{y} \in \Psi\} = \{\mathbf{y} \mid (\mathbf{x}, \mathbf{y}) \text{ peut être produit}\} = \{\mathbf{y} \mid T(\mathbf{x}, \mathbf{y}) \leq 0\}, \quad (2-6)$$

qui signifie que le vecteur d'output \mathbf{y} peut être produit par le vecteur d'inputs \mathbf{x} .

En théorie de la production, nous supposons souvent que la technologie représentée par le modèle de production satisfait certaines propriétés ou axiomes. Selon Shephard et Färe (1974), ces propriétés peuvent être présentées de la manière suivante :

- (i) $L(0) = \mathbb{R}_+^q$, l'ensemble de $L(\mathbf{y})$ est un ensemble fermé non vide qui ne contient pas l'origine pour \mathbf{y} positif.
- (ii) $0 \notin L(\mathbf{y}), \mathbf{y} > 0, \mathbf{y} \in P(\mathbf{x})$.
- (iii) $\bigcap_{\mathbf{y} > 0} L(\mathbf{y}) = \emptyset$.
- (iv) $\mathbf{x}_1 \geq \mathbf{x}_2, \mathbf{x}_2 \in L(\mathbf{y})$ signifie que $\mathbf{x}_1 \in L(\mathbf{y})$, soit la libre disposition en inputs.
- (v) Si $\mathbf{y}_1 > \mathbf{y}_2$, alors $L(\mathbf{y}_1) \in L(\mathbf{y}_2)$, soit la libre disposition en outputs.
- (vi) $L(\mathbf{y})$ est convexe pour $\mathbf{y} \in P(\mathbf{x})$.

Les axiomes 1 et 2 signifient que nous pouvons ne rien produire avec tout vecteur d'input non nul ; par contre, nous ne pouvons pas produire des outputs sans inputs. L'axiome 3 signifie qu'il n'y a pas de combinaison d'inputs qui permettent de produire un output infini. L'axiome 4 signifie que si en utilisant un vecteur d'inputs \mathbf{x}_2 , nous pouvons produire un vecteur d'output \mathbf{y} , alors pour tout vecteur \mathbf{x}_1 qui est

au moins aussi grand que x_2 , nous sommes capable de produire y . L'axiome 5 signifie qu'il est toujours possible de produire moins d'output avec une même quantité d'inputs. Finalement, l'axiome 6 signifie que si deux vecteurs d'input x_1 et x_2 sont dans $L(y)$, alors toutes les combinaisons convexes x_1 et x_2 , $x^\theta = \theta x_1 + (1-\theta)x_2$ avec $0 < \theta < 1$, sont aussi dans $L(y)$.

Une caractérisation de la forme de la frontière concerne la nature des rendements d'échelle. Selon la définition habituelle en économie, les rendements d'échelle mesurent de quelle façon varie les outputs quand tous les inputs varient simultanément. Si une augmentation proportionnelle des inputs implique une augmentation toute aussi proportionnelle des outputs, on dira que la technologie est à rendements d'échelle constants (REC). Si le pourcentage d'augmentation des outputs est moindre que le pourcentage d'augmentation des inputs, la technologie est à rendements d'échelle décroissants (RED). Finalement, si le pourcentage d'augmentation des outputs est plus important que celui des inputs, nous aurons une technologie à rendements d'échelle croissants (REI).

Pour définir une mesure de l'efficacité il est utile d'introduire la notion de fonction de distance. La fonction de distance en orientation inputs de Shephard (1953) est définie comme de suit :

$$D_I(y, x) = \max \left\{ \lambda \left| \left(\frac{1}{\lambda} \right) x \in L(y) \right. \right\}. \quad (2-7)$$

Il est évident que $D_I(y, x) \geq 1$ et l'isoquant est l'ensemble des x tel que $D_I(y, x) = 1$.

La mesure d'efficacité orientée en inputs de Farrell (1957) est :

$$TE(y, x) = \min \left\{ \theta \mid \theta x \in L(y) \right\}. \quad (2-8)$$

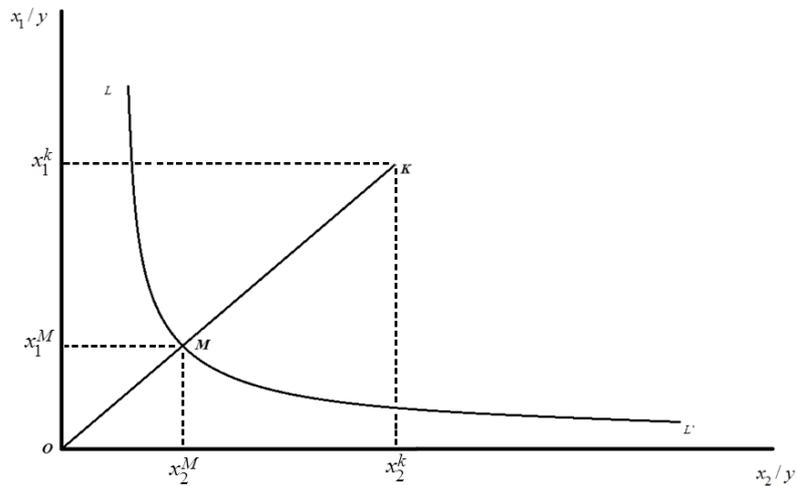
Selon la définition, il est évident que $0 < TE(y, x) \leq 1$ et qu'elle est la réciproque de la fonction de distance, $TE(y, x) = 1/D_I(y, x)$.

2-1-2 L'efficacité au sens de Farrell

La revue sur la mesure d'efficacité commence par les travaux de Farrell (1957) qui se sont appuyés sur les travaux de Debreu (1951) et Koopmans (1951). La contribution de Farrell pour mesurer l'efficacité technique des entreprises est décisive. Il construit un modèle de programmation linéaire (PL) en utilisant des données d'un échantillon d'entreprises, dont la solution donne une mesure numérique de l'efficacité technique d'une entreprise dans cet échantillon. En fait, comme nous l'avons montré ci-dessus, l'efficacité technique de Farrell est la réciproque que la fonction de distance de Shephard. En plus de fournir une mesure de l'efficacité technique, Farrell a également identifié l'efficacité allocative comme autre composante de l'efficacité économique globale. En d'autres termes, Farrell a proposé de décomposer en deux éléments l'efficacité d'une entreprise : l'efficacité technique, ce qui reflète la capacité d'une entreprise à produire un niveau de production maximale avec un vecteur d'inputs donnés et l'efficacité allocative, qui reflète la capacité d'une entreprise à utiliser des inputs dans des proportions optimales, compte tenu de leurs prix respectifs et de la technologie de production. Ces deux mesures sont ensuite combinées pour fournir une mesure de l'efficacité économique totale.

Pour montrer les idées de Farrell, nous pouvons utiliser un exemple simple. Supposons qu'une entreprise qui utilise deux inputs (x_1 et x_2) pour produire un output. La connaissance de l'isoquant des entreprises pleinement efficaces représenté par LL' dans la Figure 2-2, permet de mesurer l'efficacité technique des entreprises en question. Si une entreprise utilise des quantités d'inputs, définies par le point K , pour produire un certain niveau de production, l'inefficacité technique de cette entreprise peut être représentée par la distance MK . Elle représente la réduction de la quantité de tous les inputs sans une réduction de l'output. En fait, cette réduction est généralement exprimée par un pourcentage, donné par le rapport OM / OK . Ce rapport représente le pourcentage par lequel tous les inputs doivent être réduits par l'entreprise en question pour atteindre un point techniquement efficace. Donc

l'efficacité technique (*ET*) d'une entreprise est le plus souvent mesurée par le rapport $ET = OM / OK$, dont les valeurs sont comprises entre 0 et 1, Si elle est égale à 1, l'entreprise se situe sur la frontière de production, et elle est techniquement efficace (le point *M* par exemple). L'ensemble des entreprises techniquement efficaces forme la frontière de production.

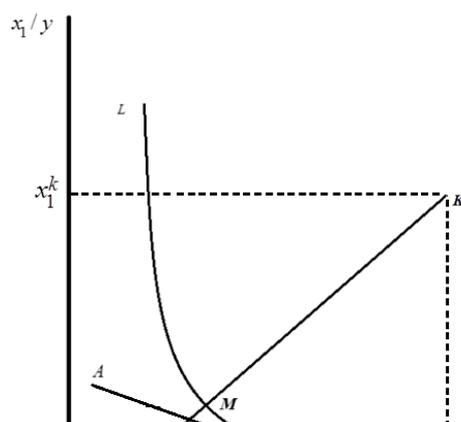


Source originale : Farrell (1957).

Figure 2-2 : Efficacité technique de Farrell

Avec la connaissance des prix des inputs, il est possible de mesurer l'efficacité allocative. La Figure 2-3 illustre cette mesure. La pente de la droite AA' est égale au rapport de prix des deux facteurs utilisés. Dans ce cas, l'entreprise au point *R'* est plus efficace que l'entreprise au point *M* au sens de la minimisation des coûts. La dépense de production en *R'* est seulement une fraction OR / OM de la dépense en *M*. Donc, l'efficacité allocative (*EA*) est mesurée par le rapport $EA = OR / OM$. Et l'efficacité économique (*EE*) est définie par :

$$EE = OR / OK = OM / OK * OR / OM = ET * EA .$$



Source originale : Farrell (1957)

Figure 2-3 : Efficacité allocative de Farrell

En pratique, la frontière de production des entreprises pleinement efficace n'est pas connue, par conséquent, nous devons l'estimer à partir des observations provenant d'un échantillon d'entreprises dans une industrie donnée. Donc la question qui se pose est : comment estimer la frontière de production ? La littérature distingue deux grandes approches :

- a. **L'approche paramétrique.** Dans ces modèles, l'ensemble des possibilités de production est défini par une frontière de production qui prend une forme analytique (ou polynomiale) et dépend de paramètres inconnus à estimer. L'avantage de cette approche est que nous pouvons facilement interpréter le sens économique des paramètres estimés et nous avons également les propriétés statistiques des estimateurs obtenus. Mais comme le choix de la forme fonctionnelle pour la modélisation de la relation entre les outputs et les inputs est arbitraire, il y a de fortes chances d'avoir une mauvaise spécification. Le principal problème est que dans le cas de mauvaise spécification, les calculs de la productivité marginale, des élasticités de substitution, de l'élasticité d'échelle, ainsi que divers tests statistiques peuvent être biaisés.
- b. **L'approche non-paramétrique.** Par définition les modèles non paramétriques n'imposent pas de forme fonctionnelle à la frontière de production, contrairement

à l'approche paramétrique. Par contre, ce n'est pas sans problème puisqu'il est généralement très difficile de produire une inférence statistique sur la localisation de la frontière ou même de produire des écart-types pour les mesures d'efficacité. Cependant, pour les modèles DEA (*Data Envelopment Analysis*) et FDH (*Free Disposal Hull*) cette lacune est en partie résolue.

Plus loin, nous exposerons deux approches qui introduisent des modèles statistiques de la mesure d'efficacité.

2-2 La revue méthodologique

Dans cette section, nous présenterons une synthèse méthodologique pour mesurer l'efficacité des entreprises. Nous commencerons par exposer l'approche économétrique, ensuite nous présenterons l'approche basée sur la recherche opérationnelle. Les méthodes issues de la recherche opérationnelle sont essentiellement des techniques d'enveloppement de données. Elles peuvent être non paramétriques comme les méthodes FDH, DEA et les indices de Malmquist, ou paramétrique comme la méthode de Aigner et Chu (1968). Quelle que soit la méthode retenue, on peut procéder à l'analyse dans l'espace primal (fonction de production ou de transformation) ou dans l'espace dual (fonction de coût).

2-2-1 L'approche paramétrique

Les recherches empiriques sur la fonction de production remontent à Cobb et Douglas en 1927. En fait, les travaux empiriques sur la fonction de production se sont concentrés sur la répartition des revenus entre le capital et la main d'œuvre au niveau macroéconomique jusque dans les années 50. L'analyse sur la structure de la production au niveau microéconomique a débuté avec les travaux de Dean (1951), Johnston (1959) et Nerlove (1963). Par contre, ces trois économistes se sont plutôt intéressés à la fonction de coût. L'attention se porte vers la fonction de production au niveau individuel au début des années 60 (Hildebrand et Liu (1965), Zellner et Revankar (1969)).

Les producteurs sont considérés comme des « preneurs de prix » sur le marché des inputs, donc le prix des inputs est considéré exogène. Dans la recherche empirique, nous utilisons souvent une fonction de production ce qui suppose un seul output (pour $y \in \mathbb{R}$) :

$$y \leq f(\mathbf{x}), \quad (2-9)$$

avec f la fonction de production d'un seul output utilisant un vecteur d'inputs \mathbf{x} . La mesure d'efficacité technique de Debreu-Farrell à orientation output est donnée par :

$$TE(y, \mathbf{x}) = \frac{y}{f(\mathbf{x})} \leq 1. \quad (2-10)$$

Notre cadre économétrique inclut l'interprétation de Debreu-Farrell ainsi que la définition classique d'une fonction de production. Donc nous commençons avec un modèle où l'inefficience est sous forme multiplicative :

$$y_i = f(\mathbf{x}_i, \boldsymbol{\beta}) TE_i, \quad (2-11)$$

avec $0 \leq TE_i \leq 1$, $\boldsymbol{\beta}$ le vecteur des paramètres de la fonction de production, et i est l'indice de la firme dont on mesure l'inefficacité. En plus de l'efficacité, nous pouvons aussi estimer $\boldsymbol{\beta}$, ce qui permettrait par exemple de mesurer les élasticités de substitution, les rendements d'échelle, etc. En prenant une transformation logarithmique, nous obtenons :

$$\ln y_i = \ln f(\mathbf{x}_i, \boldsymbol{\beta}) + \ln TE_i = \ln f(\mathbf{x}_i, \boldsymbol{\beta}) - u_i, \quad (2-12)$$

où $u_i \geq 0$ est une mesure d'efficacité technique, avec $u_i = -\ln TE_i \approx 1 - TE_i$. Notons que la dernière relation n'est acceptable que si l'inefficacité est faible. Pour obtenir l'efficacité des entreprises, nous devons donc estimer le vecteur de paramètre $\boldsymbol{\beta}$ économétriquement. Dans ce cas-là, nous devons faire des hypothèses sur le caractère statistique du terme u_i .

2-2-1-1 La méthode de moindres carrés corrigés

Prenons une fonction de production $f(x_i, \beta)$ de forme de log linéaire :

$$\ln y_i = \alpha + \beta^T \ln x_i + \varepsilon_i, \quad (2-13)$$

où $\varepsilon_i = -u_i$. Nous supposons que ε_i est un terme d'erreur classique pour une régression des moindres carrés ordinaires (MCO), sauf que $E(\varepsilon_i | x_i) \leq 0$. Afin de s'assurer que la frontière estimée est toujours au-dessus de toutes les observations, le modèle peut être écrit :

$$\begin{aligned} \ln y_i &= \alpha + \beta^T \ln x_i + \varepsilon_i \\ &= (\alpha + E(\varepsilon_i | x_i)) + \beta^T \ln x_i + (\varepsilon_i - E(\varepsilon_i | x_i)) \\ &= \alpha_i^* + \beta^T \ln x_i + \varepsilon_i^*. \end{aligned} \quad (2-14)$$

Nous déplaçons la fonction à estimer vers le haut jusqu'à ce que tous les résidus sauf un deviennent négatifs. Nous pouvons obtenir l'efficacité des DMU estimés par la méthode des moindres carrés corrigés (MCC) :

$$\alpha_i^{MCC} = \alpha_i^* + \max_j \{e_j\}. \quad (2-15)$$

Les résidus des DMU sont :

$$e_{i,MCC} = e_i - \max_i e_i. \quad (2-16)$$

La DMU qui a le plus petit e_i se verra accorder une valeur nulle et les autres auront toutes une valeur positive. Les paramètres e_i représentent les écarts d'efficience entre les DMU. Donc la frontière s'appuie sur les données d'au moins une DMU.

2-2-1-2 Les frontières stochastiques

Afriat (1972) a spécifié un modèle semblable à celui d'Aigner et Chu (1968). Mais il suppose que le terme ε_i suit une loi de probabilité gamma et les paramètres du modèle ont été estimés en utilisant la méthode du maximum de vraisemblance (MV). En fait,

Afriat propose le modèle $y = A \prod_{i=1}^K x_i^{\beta_i} u$ et suppose que $u = \exp(-z)$, où z suit une loi

gamma avec le paramètre θ :

$$G(z, \theta) = \frac{1}{\Gamma(\theta)} z^{\theta-1} \exp(-z), 0 < z < \infty, \theta > 0,$$

avec $\Gamma(\theta) = \int_0^{\infty} z^{\theta-1} \exp(-z) dz$. Le terme d'erreur u peut prendre une valeur entre 0

et 1, donc $z = \log 1/u$ est de loi gamma et prend des valeurs entre zéro et l'infini. La

fonction de densité de u est :

$$H(u, \theta) = \frac{1}{\Gamma(\theta)} \left(\log \left(\frac{1}{u} \right) \right)^{\theta-1}, 0 < u < 1, \theta > 0,$$

u est le paramètre d'efficacité, et $\theta < 1$ correspond au cas où la plupart des firmes

sont assez efficaces ; $\theta = 1$ correspond à une distribution uniforme et $\theta > 1$ nous

donne le cas où la plupart des firmes sont relativement inefficaces. Si nous avons un

échantillon de N firmes dans un secteur, pour la $n^{ième}$ firme, $y_n = f(\mathbf{x}_n) u_n$, Afriat

suggère que la mesure de l'efficacité pour la firme n soit égale à $u_n = y_n / f(\mathbf{x}_n)$.

Puisque $u = \exp(-z)$, où z suit une loi gamma, l'efficacité est naturellement mesurée

par :

$$\varepsilon = E(u) = \int_0^{\infty} \exp(-z) G(z, \theta) dz = 2^{-\theta}.$$

Par la suite, Richmond (1974) a noté que les paramètres du modèle d'Afriat pouvaient

être également estimés en utilisant une méthode qui est devenu connu sous le nom

moindre carré ordinaires corrigé (MCOC), où la méthode MCO fournit des

estimations non biaisées des paramètres de la pente, et l'estimateur MCO (biaisée à la

baisse) de l'ordonnée à l'origine est ajustée à la hausse.

Cette approche est utilisée par Richmond pour une fonction Cobb-Douglas :

$$y_i = A \prod_{j=1}^k x_{ij}^{\beta_j} u_i \text{ avec } i = 1, 2, \dots, N ,$$

ou encore ,sous forme log linéaire :

$$\ln y_i = \alpha + \sum_{j=1}^k \beta_j \ln x_{ij} - z_j ,$$

avec $E(z_i) = \theta$, $Var(z_{ij}) = \theta$ et $E(z_i z_j) = 0$ pour $i \neq j$. Ensuite, nous pouvons introduire $\beta_0 = \alpha - \theta$ et $v_{ij} = \theta - z_j$, nous aurons :

$$\ln y_i = \beta_0 + \sum_{j=1}^k \beta_j \ln x_{ij} + v_i ,$$

avec $E(v_i) = 0$, $Var(v_i) = \theta$ et $E(v_i v_j) = 0$ pour $i \neq j$. Finalement, nous pouvons estimer tous les paramètres par MCO. L'estimateur de θ est obtenu par $\hat{\theta} = \sum_i v_i^2 / N - k - 1$. Mais Meeusen et Broeck (1977) ont montré que l'efficacité obtenue par la frontière d' Afriat et Richmond sous-évalue l'efficacité des entreprises.

Finalement, Aigner, Lovell et Schmidt (1977), Battese et Corra (1977) et Meeusen et Broeck (1977) ont développé simultanément la méthode connue sous le nom d'estimation de frontière stochastique. Ils pensent que la position des entreprises par rapport à la frontière de production est déterminée à la fois par des facteurs qui peuvent être contrôlés par les firmes, comme l'inefficacité technique ou économique, la volonté et l'effort des producteurs et ses employeurs, et des facteurs qui sont hors du contrôle des firmes, comme les productions défectueuses ou endommagées. Mais la frontière elle-même est différente pour les différentes firmes, ou même différente pour une même firme pour des périodes différentes. Avec cette interprétation, la frontière est stochastique avec une perturbation v_i positive, négative ou nul. Et v_i est un résultat des éléments externes favorables ou défavorables, comme la chance, le climat, la topographie et la performance de la machine (Aigner et *al.* 1977). Une

forme possible pour la frontière stochastique est $y_i = f(\mathbf{x}_i)TE_i e^{v_i}$. Nous pouvons prendre un modèle log-linéaire :

$$\ln y_i = \alpha + \boldsymbol{\beta}^T \ln \mathbf{x}_i + v_i - u_i = \alpha + \boldsymbol{\beta}^T \ln \mathbf{x}_i + \varepsilon_i, \quad (2-17)$$

où $-u_i = \ln TE_i$ et $\varepsilon_i = v_i - u_i$. Comme ci-dessus, $u_i > 0$ représente l'inefficacité et v_i représente tous les autres éléments comme l'erreur de mesure, les autres bruits statistiques, la variation aléatoire de la frontière entre les firmes. Donc nous obtenons une frontière stochastique :

$$\alpha + \boldsymbol{\beta}^T \ln \mathbf{x}_i + v_i, \quad (2-18)$$

Notre objet est d'obtenir un estimateur de u_i et à cette fin, il nous faut tout d'abord correctement estimer les paramètres de la technologie, α et $\boldsymbol{\beta}$.

Nous pouvons aussi passer par la frontière de coût :

$$\ln C_i = \alpha + \boldsymbol{\beta}^T \mathbf{w}_i + v_i + u_i. \quad (2-19)$$

Le choix de la forme de frontière est important pour l'estimation et les hypothèses que nous faisons pour la distribution de v_i et u_i seront décisives. Nous retrouvons les combinaisons normale/semi-normale du modèle d'Aigner *et al.* (1977) ; le modèle normal/exponentiel et le modèle normal/gamma de Meeusen et van den Broeck (1977), Aigner *et al.* (1977), Stevenson (1980) et Greene (1980), et le modèle tronqué de Kumbhakar et Lovell (2000), etc.

Maintenant, si nous nous intéressons à la situation où l'output est un vecteur, nous devons utiliser la fonction de transformation. Nous pouvons l'écrire de la façon suivante $T(\mathbf{y}, \mathbf{x}) = 0$ où \mathbf{y} représente un vecteur d'outputs et \mathbf{x} représente le vecteur d'input utilisé pour produire \mathbf{y} . Ici nous avons besoin de quelques hypothèses pour établir un modèle empirique. Par exemple, nous pouvons supposer la séparabilité homothétique. Dans ce cas, la fonction de transformation peut être exprimée comme:

$$g(\mathbf{y}) = h(\mathbf{x}). \quad (2-20)$$

La fonction $g(\mathbf{y})$ est un agrégateur de l'output qui relie les outputs agrégés à une fonction de production connue. En supposant une élasticité constante de la fonction de transformation, Fernandez et *al.* (2000) proposent le modèle suivant :

$$\left(\sum_{m=1}^M \alpha_m^q y_{i,t,m}^q \right)^{\frac{1}{q}} = \boldsymbol{\beta}^T \mathbf{x}_{it} + v_{it} - u_{it}. \quad (2-21)$$

Ces méthodes sont caractérisées par une sensibilité des résultats au choix de la forme fonctionnelle, mais surtout au choix de la structure stochastique des résidus et des erreurs de mesures.

2-2-2 L'approche non paramétrique

Les limites des approches paramétriques pour mesurer l'efficacité découlent des hypothèses sur la forme fonctionnelle, sur le comportement des résidus, sur la courbure de certaines fonctions, etc. Elles ont toutes un impact sur le résultat final (Gagné et Ouellette, 1998 et 2002). Farrell propose plutôt une approche d'enveloppement de données. C'est la première apparition de la technique DEA bien que n'en portant pas encore le nom. Dans les années soixante, il y a beaucoup d'économistes et de mathématiciens qui ont amélioré cette approche jusqu'à l'apparition du premier modèle appelé DEA en 1978.

La méthode DEA est une approche non paramétrique pour mesurer l'efficacité d'une DMU, comme une entreprise ou une agence publique, etc. Le premier modèle de DEA est à rendements d'échelle constants et est développé par Charnes, Cooper et Rhodes (le modèle CCR) et publié dans une revue de recherche opérationnelle (*European Journal of Operational Research*, 1978). Le principe consiste à former le plus petit cône convexe qui enveloppe toutes les données observées. Dans ce modèle, la principale limitation est qu'il faut supposer que les DMU ont une technologie à rendements d'échelles constants. Banker, Charnes, et Cooper (1984) étendent le

modèle CCR en enlevant l'hypothèse sur les rendements d'échelles (connu sous l'acronyme BCC). L'enveloppe convexe n'est plus nécessairement un cône, parce qu'elle est plus serrée que le cône convexe et enveloppe ainsi les données de plus près. De plus, nous pouvons également chercher à interpréter cette différence par l'efficacité d'échelle des entreprises. Dans les années qui ont suivi beaucoup de chercheurs ont contribué au développement et à la mise en œuvre de ces deux modèles.

2-2-2-1 Le modèle DEA classique

Dans les méthodes paramétriques, on maximise la fonction de vraisemblance ou on minimise la somme des erreurs au carré une seule fois en utilisant toutes les observations et la fonction estimée s'applique à toutes les DMU. Avec le DEA, chaque résultat est optimisé pour chaque DMU individuellement. Pour illustrer le principe, supposons qu'il y ait p inputs et q pour chacune des N DMU.

Soient $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$ et $\mathbf{y}_i = (y_{1i}, y_{2i}, \dots, y_{qi})^T$ représentent les vecteurs des inputs et outputs de la $i^{\text{ème}}$ DMU. La matrice \mathbf{X} d'inputs ($p \times N$) et la matrice \mathbf{Y} d'outputs ($q \times N$) représentent les données des N entreprises. L'objectif est de mesurer la performance de chaque DMU relativement à la meilleure pratique observée dans l'échantillon. Pour simplifier la notation, le vecteur des inputs et des outputs pour la DMU₀ est noté (x_o, y_o) . Alors le modèle CCR peut s'écrire de manière suivante :

$$\theta_0 = \min_{\theta_0} \left\{ \theta_0 \mid \sum_{j=1}^N \lambda_j y_{rj} \geq y_{r0}, \sum_{j=1}^N \lambda_j x_{sj} \leq \theta_0 x_{s0}, \lambda_j \geq 0 \right\}, \quad (2-22)$$

où $r = 1, 2, \dots, q$ et $s = 1, 2, \dots, p$.

Comme le problème ne contraint pas la relation entre les multiplicateurs λ_j , l'enveloppe des données sera un cône. Plus généralement, il satisfait la condition suivante :

$$\Psi^{REC} = \{(\mathbf{x}, \mathbf{y}) \in \Psi \mid ((\alpha \mathbf{x}, \alpha \mathbf{y}) \in \Psi \text{ pour tous les } \alpha \geq 0)\}. \quad (2-23)$$

Cette enveloppe correspond à une technologie de rendements d'échelle constants.

Divers auteurs, tels que Färe, Grosskopf et Logan (1983) et Banker, Charnes et Cooper (1984) ont suggéré d'ajuster le modèle CCR afin de tenir compte des rendements d'échelle non constants.

Pour avoir un modèle à rendements d'échelles variables, il suffit d'ajouter une contrainte de convexité, de la forme $\sum_{i=1}^N \lambda_i = 1$ au modèle initial. Ce modèle en orientation input peut s'écrire la manière suivante :

$$(D_{BCC}^e) \left\{ \begin{array}{ll} \max & \theta, \\ \text{s.c} & \sum_{j=1}^N \lambda_j x_j - \theta \cdot x_o = -s_j^- \quad (a) \\ & \sum_{j=1}^N \lambda_j y_j - y_o = s_j^+ \quad (b) \\ & \sum_{j=1}^N \lambda_j = 1 \quad (c) \\ & \lambda_j \geq 0, j = 1, 2, \dots, N, \quad s^- \geq 0, s^+ \geq 0, 0 < \theta \leq 1 \quad (d) \end{array} \right. \quad (2-24)$$

où θ représente le coefficient d'efficacité technique; x_j représente le vecteur d'input pour l'entreprise j ; y_j représente le vecteur d'output pour l'entreprise j ; λ_j représente le vecteur de multiplicateurs associé à chaque entreprise; s_j^- représente les variables d'écart de l'input j et s_j^+ représente les variables d'écart de l'output j . La contrainte (a) impose que les inputs corrigés utilisés par la DMU évaluée (θx_o) soient au moins égaux aux quantités utilisées par la combinaison convexe des inputs utilisés par les unités de référence. La contrainte (b) implique que la DMU évaluée ne doit pas produire plus d'outputs que la combinaison convexe des inputs utilisés par les unités de référence. La contrainte (c) force à envelopper les données au plus près, ce qui permet d'avoir des rendements d'échelle variables en pratique. C'est là que la formulation BCC du problème d'optimisation est plus souple que celle la CCR. Pour

des rendements d'échelle non croissants, l'équation (c) devient une inégalité faible :

$\sum_{j=1}^N \lambda_j \leq 1$. Le programme BCC peut être écrit de façon plus compacte sous la forme

suivante :

$$\theta_0 = \min_{\theta_0} \left\{ \theta_0 \mid \sum_{j=1}^N \lambda_j y_{rj} \geq y_{r0}, \sum_{j=1}^N \lambda_j x_{sj} \leq \theta_0 x_{s0}, \sum_{j=1}^N \lambda_j = 1 \right\}, \quad (2-25)$$

où $r = 1, 2, \dots, q$ et $s = 1, 2, \dots, p$.

Comme nous avons vu, le modèle CCR suppose que l'ensemble des possibilités de production est à rendements constants. On postule que la croissance et la réduction radiales de toutes les DMU observées ou leurs combinaisons non négatives sont possibles. Par conséquent, le coefficient d'efficacité obtenu par le modèle CCR est appelé efficacité technique globale. D'autre part, le modèle BCC suppose que la combinaison convexe des DMU observées forme l'ensemble des possibilités de production, le coefficient d'efficacité obtenu par le modèle BCC est appelé efficacité technique locale. En générale, puisque l'ensemble des possibilités de production sous l'hypothèse de rendements d'échelle variables est un sous-ensemble de l'ensemble des possibilités de production sous l'hypothèse de rendements d'échelles constants (REC), soit $\Psi^{REV} \subset \Psi^{REC}$, nous avons alors la relation $\theta^{REC} < \theta^{REV} < 1$.

2-2-3 L'approche semi-paramétrique : le modèle d'Aigner et Chu

Dans le cas des frontières indiquées de la sous-section précédente, l'écart d'une observation par rapport à son output maximum est attribué uniquement à l'inefficacité de l'entreprise. Ce genre de frontière est essentiellement déterministe. Ceci est en contraste avec la spécification de la frontière dans lequel l'output maximal pour un producteur est déterminé à la fois par la fonction de production et par le hasard des facteurs externes, comme la chance ou des perturbations imprévues sur un marché. Dans cette seconde interprétation, le modèle est caractérisé comme une frontière stochastique comme nous l'avons vu plus haut. Dans cette section nous présenterons

un modèle qui retient le caractère déterministe de l'approche d'enveloppe et la spécification explicite d'une forme fonctionnelle pour la frontière comme dans le cas des frontières stochastiques. C'est le modèle d'Aigner et Chu (1968).

2-2-3-1 Modèle de base

Le modèle de Aigner et Chu (1968) est un modèle semi-paramétrique car il combine les méthodes paramétriques, notamment le choix d'une forme fonctionnelle, avec des techniques non paramétriques comme l'utilisation la programmation linéaire ou quadratique.

Aigner et Chu (1968) utilisent une fonction de production de type Cobb-Douglas avec un output, deux inputs :

$$y_i = Ax_{1i}^{\beta_1} x_{2i}^{\beta_2} U_i, \quad (2-26)$$

où y_i est l'output, (x_{1i}, x_{2i}) le vecteur des inputs, U_i représente l'inefficience, et (A, β_1, β_2) sont des paramètres de la fonction de production qui sont à estimer. U_i est une variable de perturbation qui prend une valeur entre 0 et 1. Sous forme logarithmique, nous obtenons :

$$\ln y_i = \ln A + \sum_{k=1}^K \beta_k \ln x_{ki} + \varepsilon_i = \alpha + \sum_{k=1}^K \beta_k \ln x_{ki} - \mu_i, \quad (2-27)$$

où $\varepsilon_i = \ln U_i = -\mu_i$. La partie déterministe du côté droit de l'équation correspond à la frontière. Tous les éléments stochastiques sont regroupés dans le terme ε_i . Aigner et Chu (1968) proposent deux méthodes pour calculer les paramètres qui limitent les résidus μ_i aux valeurs non négatives:

Un programme linéaire :

$$\begin{aligned} \min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i \\ \text{s.c. : } \ln y_i - \alpha - \beta^T \ln x_i \leq 0 \quad \forall i \end{aligned} \quad (2-28)$$

et un programme quadratique :

$$\min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i^2 \quad (2-29)$$

s.c : $\ln y_i - \alpha - \beta^T \ln x_i \leq 0 \quad \forall i.$

Dans ces deux méthodes, les variables d'écart dans les contraintes nous donne une mesure de $-\mu_i$. Avec cette méthode, nous pouvons comparer les résidus de chaque DMU avec le $-\mu_i$ estimé :

$$\hat{\mu}_i = \hat{\alpha} + \hat{\beta}^T \ln x_i - \ln y_i. \quad (2-30)$$

L'inefficacité technique de l'ensemble d'échantillon est mesurée par :

$$\frac{1}{N} \sum_{i=1}^N \hat{\mu}_i. \quad (2-31)$$

Les termes d'inefficience sont faciles à récupérer avec ces méthodes. Comme le terme d'erreur est toujours de même signe, il est inutile de prendre les erreurs au carré et la simple minimisation de la somme des erreurs suffit, simplifiant d'autant la résolution du problème. Dans le programme quadratique, l'influence des valeurs extrêmes dans la fonction de critère sera plus importante, mais dans certaine situation, ce critère peut mieux convenir (un modèle avec système d'équations simultanées). Dans leur article, Aigner et Chu ont utilisé la forme fonctionnelle de Cobb-Douglas. Le choix de la forme fonctionnelle, du fait qu'elle impose des restrictions sur la nature des résultats, n'est pas sans conséquence ; par exemple, la forme Cobb-Douglas implique que l'élasticité de substitution entre facteurs de production soit égale à un. Pour relâcher ce genre de contrainte imposée par la forme fonctionnelle, nous prendrons une forme plus souple dans notre thèse. L'intérêt de mettre ce modèle au centre de notre analyse est que par rapport au modèle non paramétrique, nous avons une forme explicite de la fonction de production et la fonction de coût, donc il est possible de mesurer les caractéristiques technologiques selon les approches classiques (en supposant une frontière différentiable) : économies d'échelles, élasticités de substitution, etc., comme dans les modèles statistiques. En même temps, ce modèle garde l'avantage du

modèle non paramétrique : facile à calculer et évite la nécessité de faire des hypothèses sur les termes d'erreurs et leurs distributions.

2-2-3-2 La littérature autour de modèle d'Aigner et Chu

A partir du travail pionnier d'Aigner et Chu en 1968, quelques chercheurs ont essayé d'appliquer ou améliorer cette approche. Dans cette section, nous exposerons cette littérature.

Cette approche est utilisée par Førsund et Hjalmarsson (1979), Albritsen et Forsund (1990), Førsund(1992). Dans ces articles, les auteurs utilisent une fonction de production généralisée de Zellner et Revankar (1969) :

$$\gamma_0 \ln y_i + \gamma_1 y_i = \alpha + \sum_{k=1}^K \beta_k \ln x_{k_i} .$$

Ils minimisent la somme des résidus sous contrainte d'homogénéité et tous les paramètres sont positifs.

Timmer (1971)

Aigner et Chu (1968) suggèrent que les restrictions d'inégalité du programme pourraient être altérées pour introduire une frontière probabiliste de telle sorte que certaines observations d'output pourraient se situer au-dessus de la frontière estimée (page 838). Timmer (1971) a retenu cette suggestion et proposé une approche probabiliste pour la modélisation de la frontière de production. Dans sa modélisation, une proportion des observations peut être au-dessus de la frontière estimée. Dans son article, il relâche l'hypothèse d'homogénéité linéaire de la fonction de production (le cas de Farrell (1957)), mais le coût de cette généralisation est de supposer une forme fonctionnelle Cobb-Douglas. Timmer utilise une frontière probabiliste :

$$\Pr \left(\prod_{i=0}^m \hat{\alpha}_i x_{ij} \geq y_{it} \right) > P .$$

où P est une probabilité spécifiée de façon exogène (par exemple, 98 % ou 95 %). La méthode proposée par Timmer est d'éliminer $(1 - p)$ pourcent des firmes efficaces estimé par la méthode d'Aigner et Chu pour atteindre le niveau de satisfaction.

Timmer applique son modèle sur une base de données américaine dans le secteur agricole pour la période 1960 à 1967. Une exploitation agricole moyenne pour chaque état et chaque année est considérée comme une observation. La production agricole brute en dollars de l'état j à l'année t divisée par le nombre d'exploitation est utilisée comme output; le nombre d'employeurs, le flux de construction déflaté et la valeur des terres sont utilisés comme inputs. Il a trouvé une efficacité technique moyenne de 92,4 %. Mais le problème de Timmer est qu'il élimine les observations d'une façon « arbitraire ». Pour cette raison, cette méthode n'est pas largement utilisée.

Førsund et Jansen (1977)

Førsund et Jansen (1977) ont proposé une nouvelle approche de la programmation linéaire et un modèle paramétrique pour étendre la recherche sur la fonction de coût. Ils partent d'une fonction de production homothétique, $y_i = F[f(\mathbf{x}_i)]$, où F est une fonction croissante, f correspond une fonction homogène de degré un. Le modèle empirique est obtenu en spécifiant la fonction :

$$y_i = F[f(\mathbf{x}_i)v_i].$$

Où v_i représente l'efficacité technique de la firme i . On suppose que la densité de l'efficacité technique est donnée par :

$$h(v_i) = (1 + \theta)v_i^\theta, 0 < v_i < 1, \theta > 0.$$

Par la dualité, nous pouvons trouver la fonction de coût correspondante :

$$\ln C_i = \ln F^{-1}(y_i) + \ln C(\mathbf{w}_i) - \ln(v_i),$$

où \mathbf{w}_i est le vecteur de prix et $C(\mathbf{w}_i)$ est la fonction de coût unitaire. Ensuite, ils dérivent la fonction de vraisemblance correspondante. Les paramètres de la fonction

de coût sont obtenus par l'utilisation d'un programme linéaire pour minimiser la somme des coûts en respectant la contrainte que le coût observé supérieur ou égal à celui donné par la frontière de coût. Ce modèle est intéressant, car c'est l'une des premières applications pour estimer une frontière de coût au lieu d'une frontière de production. Cette méthode est controversée à cause de la possible contribution de l'inefficacité allocative à l'estimation de l'inefficacité observée des firmes.

Nishimizu et Page (1982)

Nishimizu et Page (1982) ont été les premiers à isoler le changement technique et le changement d'efficacité comme les composantes de la croissance de la productivité. Ils utilisent l'indice de productivité de Malmquist avec le modèle de frontière de production développé par Aigner et Chu (1968). Ils ont estimé la croissance de la productivité totale, le progrès technologique et le changement d'efficacité technique en Yougoslavie pour la période de 1965 à 1978. Ils ont trouvé que dans les cinq premières années, l'efficacité technique a augmenté, mais pour les cinq années suivantes, l'efficacité technique a diminué.

Pour commencer, ils ont défini la fonction de production qui décrit la meilleure pratique :

$$y_{it}^F = f[\mathbf{x}_i(t), t],$$

où y_{it}^F représente le niveau d'output potentielle pour la firme i au temps t , et $\mathbf{x}_i(t)$ est un vecteur d'input. La fonction dépend directement du temps, donc le déplacement de la frontière de production peut se produire sans changement d'input. Comme dans l'article de Solow (1957), la variable temps t est supposé capter le progrès technique neutre dans la production. Pour tous les outputs $y(t)$ observés en utilisant un vecteur d'input $x(t)$, nous pouvons estimer le niveau d'efficacité technique obtenu par la fonction de distance orientée en output :

$$D_o^t = [\mathbf{x}_i(t), y_i(t)] = \frac{y_i(t)}{y_{it}^F},$$

où $D_o^t = 1$ pour les entreprises techniquement efficaces et $0 < D_o^t < 1$ pour les entreprises techniquement inefficaces. En utilisant une approche paramétrique, Nishimizu et Page supposent que la fonction de distance est :

$$e^u = \frac{y_i(t)}{y_{it}^F},$$

où e^u est le rapport d'efficacité technique. Donc la fonction de production pour la firme i peut s'écrire :

$$y_i(t) = y_{it}^F e^u = f[\mathbf{x}_i(t), t] e^u.$$

La dérivation du logarithme de cette équation par rapport au temps nous donne :

$$\frac{\dot{y}_i(t)}{y_i(t)} = \varepsilon_x \frac{\dot{x}_i(t)}{x_i(t)} + \varepsilon_t + \dot{u},$$

avec ε_x et ε_t représentant l'élasticité de l'output par rapport aux inputs, x , et de la tendance, t , et où $\dot{y}_i(t) = \partial y_i(t) / \partial t$, $\dot{x}_i(t) = \partial x_i(t) / \partial t$. Donc le changement d'output est décomposé en trois parties : la première partie correspond au changement des inputs pondérés par les élasticités de l'output, la deuxième partie est le taux de progrès technique qui correspond au déplacement en pourcentage de la frontière de production et la dernière partie nous donne le changement d'efficacité technique pendant la période t . En prenant une fonction de production Translog (Christensen *et al.*, 1973) Nishimizu et Page utilisent l'approche de la frontière paramétrique d'Aigner et Chu pour estimer les paramètres de la frontière.

Conclusion

L'approche économétrique pour construire la frontière et estimer l'efficacité des DMU a des similitudes et des différences avec l'approche basée sur la recherche opérationnelle. Les deux méthodes mesurent la distance des DMU par rapport à la frontière estimée. Toutefois, les deux méthodes utilisent des techniques différentes pour envelopper les données. On fait des hypothèses différentes sur la distribution du bruit et on exige différentes souplesses pour la structure de production. Autrement dit, l'approche économétrique est stochastique et paramétrique. Cela lui permet de distinguer les effets du bruit statistique de l'inefficacité, et donc nous pouvons obtenir une inférence statistique pour les estimateurs. Par contraste, l'approche basée sur la recherche opérationnelle est non paramétrique et déterministe, cela permet d'éviter de confondre l'inefficacité avec des erreurs causées par une mauvaise spécification de la forme fonctionnelle. La méthode de Aigner et Chu se trouve entre les deux. Elle n'est pas une méthode économétrique, car nous n'estimons pas les paramètres à partir de techniques statistiques; elle n'est pas non plus une méthode non paramétrique, car nous devons supposer une forme fonctionnelle spécifique et essayer d'estimer tous les paramètres inconnus. C'est ce que nous appelons une méthode semi-paramétrique. L'avantage de cette méthode est qu'elle prend les avantages de la méthode DEA et ceux de la méthode économétrique. Ainsi, le recours à la programmation linéaire, comme dans la méthode DEA, mesure directement l'efficacité des DMU et évite les hypothèses sur les termes d'erreurs et leur distribution. Aussi, il est possible de mesurer les caractéristiques technologiques selon les approches classiques (en supposant une frontière différentiable): économies d'échelles, élasticités de substitution, etc., comme dans les modèles statistiques. Pour mesurer l'influence des inputs quasi-fixes, comme les infrastructures, sur la performance des entreprises, la méthode semi-paramétrique reste l'un des meilleurs choix.

Deuxième partie.

Modélisation : introduction des inputs quasi-fixes et de l'indicateur d'accessibilité dans le modèle de Aigner et Chu.

Introduction

Nous avons exposé au cours des chapitres précédents les différentes méthodes traitant de la mesure de l'impact des infrastructures et la mesure de l'efficacité des entreprises. Les études réalisées dans ce champ d'étude étaient surtout conduites au niveau macro-économique et de façon paramétrique. L'impact des infrastructures sur la performance des entreprises n'a pas attiré beaucoup d'attention. De plus, les hypothèses de rendements d'échelle constants et d'efficacité des entreprises ne correspondent pas vraiment au fonctionnement de ces dernières. Au chapitre 2, nous avons vu de quelle façon nous pouvions relâcher l'hypothèse d'efficacité des entreprises. Par contre, dans la littérature, il y a très peu de recherche utilisant l'approche de frontières, tant stochastique que non paramétrique, pour mesurer l'impact des infrastructures sur la croissance économique et la performance des entreprises.

L'objet de cette partie est d'introduire les modèles que nous utiliserons pour estimer l'impact des infrastructures sur la croissance économique au niveau micro-économique. Cette partie s'articule comme suit. Dans un premier temps, nous présenterons le modèle DEA avec les inputs quasi-fixes, et la façon de calculer l'efficacité d'échelle avec ce modèle. Dans un deuxième temps, nous exposerons le modèle semi-paramétrique avec inputs quasi-fixes dans l'espace primal (avec la fonction de production) et l'espace dual (avec la fonction de coût). Dans un troisième temps, nous présenterons les problèmes statistiques du modèle semi-paramétrique et nous proposerons une approche permettant de révéler le caractère statistique des coefficients d'efficacité estimés (estimation par intervalle de confiance). Cette approche s'articule autour de deux parties. La première propose une démonstration de la convergence des estimateurs obtenus par le modèle semi-paramétrique, la seconde, en se basant sur une méthode de ré-échantillonnage, propose une méthode d'estimation des paramètres intervalle de confiance pour les estimateurs obtenus. La méthode utilisée est une adaptation de la méthode de sous-échantillonnage.

Finalement, nous montrerons comment introduire les infrastructures dans notre modèle théorique.

Chapitre 3. Modélisation

Selon la théorie classique, les firmes se trouvent sur leur frontière de production. Elles produisent un niveau maximal d'output avec un niveau d'inputs choisi. Cette relation est généralement décrite par une fonction de production. Ainsi toutes les firmes, par définition, sont techniquement efficaces. La seule source d'inefficience économique réside dans l'inefficience allocative. Mais en pratique, la firme produit parfois au-dessous de la frontière des possibilités de production, en étant inefficace techniquement parce qu'elle peut avoir une connaissance incomplète des meilleures méthodes pour utiliser les inputs, ou être influencée par des facteurs qui l'empêchent d'être sur la frontière. On a vu dans chapitre 2 que plusieurs méthodes sont utilisées par les chercheurs. Globalement, il y a trois types de méthodes ou modèles : Premièrement, les méthodes paramétriques, qui sont des modèles statistiques dont le plus important est le modèle de la frontière stochastique. Deuxièmement, les méthodes non-paramétriques basées sur la recherche opérationnelle, dont la plus connue est la méthode DEA. Troisièmement, les méthodes semi-paramétriques basées sur l'approche d'Aigner et Chu (1968). Dans les sections suivantes, nous présenterons la modélisation du modèle non paramétrique et du modèle paramétrique en introduisant les inputs quasi-fixes.

3-1 Modélisation du modèle non paramétrique

3-1-1 Modèle DEA avec inputs quasi-fixes

Dans la section 2-2-2, nous avons présenté les fondements du modèle non paramétrique pour estimer l'efficacité des entreprises : soit les modèles CCR et BCC. Le modèle BCC améliore le modèle CCR en permettant des rendements d'échelle variables, mais il existe également un autre phénomène dans le processus de production qui doit être considéré : les inputs quasi-fixes aussi appelés inputs non discrétionnaires.

Les quasi-fixes inputs sont des inputs qui ne peuvent pas être contrôlés par les DMU au cours de la période observée. En 1984, Banker et Morey, introduisent un modèle incorporant des inputs quasi-fixes. On a la représentation suivante (Figure 3-1) :

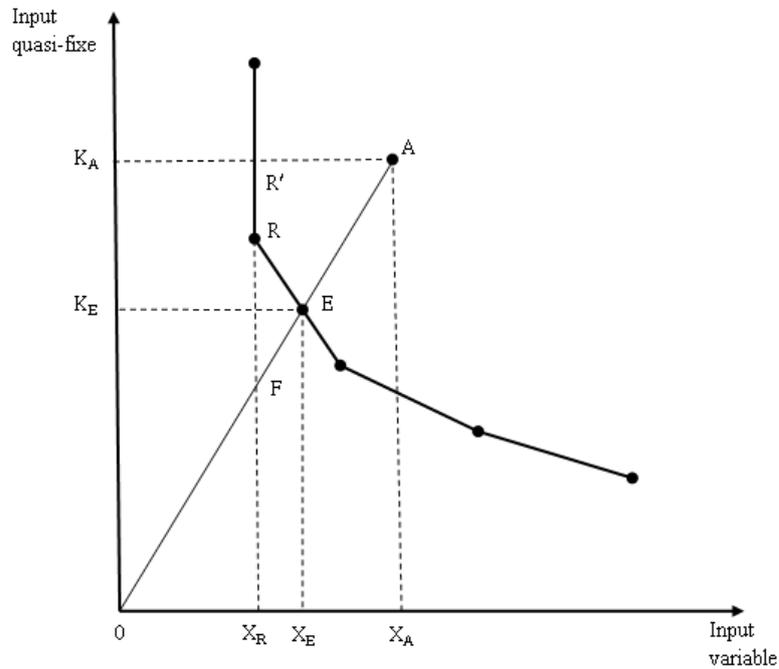


Figure 3-1 : Modèle BCC avec les inputs quasi-fixes

Supposons qu'un processus de production utilise un input variable x et un input quasi-fixe k pour produire un seul output y (Figure 3-1). L'efficacité de la DMU A sera mesurée par comparaison avec la DMU E (possiblement fictive), qui se trouve sur la frontière de production, et le coefficient d'efficacité est $\theta = \frac{OE}{OA}$. La firme A doit diminuer tous ses inputs par un facteur $(1-\theta)$ pour atteindre l'efficacité. Mais puisque k est un input qui ne peut pas être contrôlé par A, il est plus logique de la comparer à la DMU R', qui se trouve également dans la frontière de production, utilise la même quantité de k mais moins de x . Donc en introduisant la contrainte sur les inputs quasi-fixes, le modèle devient :

$$\theta_0 = \min_{\theta_0} \left\{ \theta_0 \mid \sum_{j=1}^N \lambda_j y_{rj} \geq y_{r0}, \sum_{j=1}^N \lambda_j x_{sj} \leq \theta_0 x_{s0}, \sum_{j=1}^N \lambda_j k_{hj} \leq k_{h0}, \sum_{j=1}^N \lambda_j = 1 \right\}, \quad (3-1)$$

et le coefficient d'efficacité θ_0 n'influencera que les inputs variables : x_p

3-1-2 L'efficacité d'échelle avec DEA

L'efficacité d'échelle caractérise l'écart existant entre les performances constatées et celles qui seraient obtenues dans une situation d'équilibre concurrentiel de long terme où le profit est nul, c'est-à-dire par rapport à une situation où les rendements d'échelle sont constants. Ainsi, une entreprise est inefficace à l'échelle si sa situation initiale est caractérisée par des rendements d'échelle croissants ou décroissants. Dans ce contexte de long terme, tous les facteurs de production peuvent être ajustés par le producteur pour réduire son inefficacité. La méthode DEA nous permet de mesurer l'efficacité d'échelle d'une DMU et nous donne plus précisément la nature de ses rendements d'échelle : décroissants, constants ou croissants. Avec l'hypothèse de REC, toutes les inefficacités des firmes sont supposées dérivées d'opérations inefficaces, on néglige la possibilité que les firmes ne fonctionnent pas à la taille optimale. Par contre, le modèle BCC prend en compte le problème d'échelle. La différence entre les modèles CCR et BCC nous donne un moyen pour mesurer l'efficacité d'échelle.

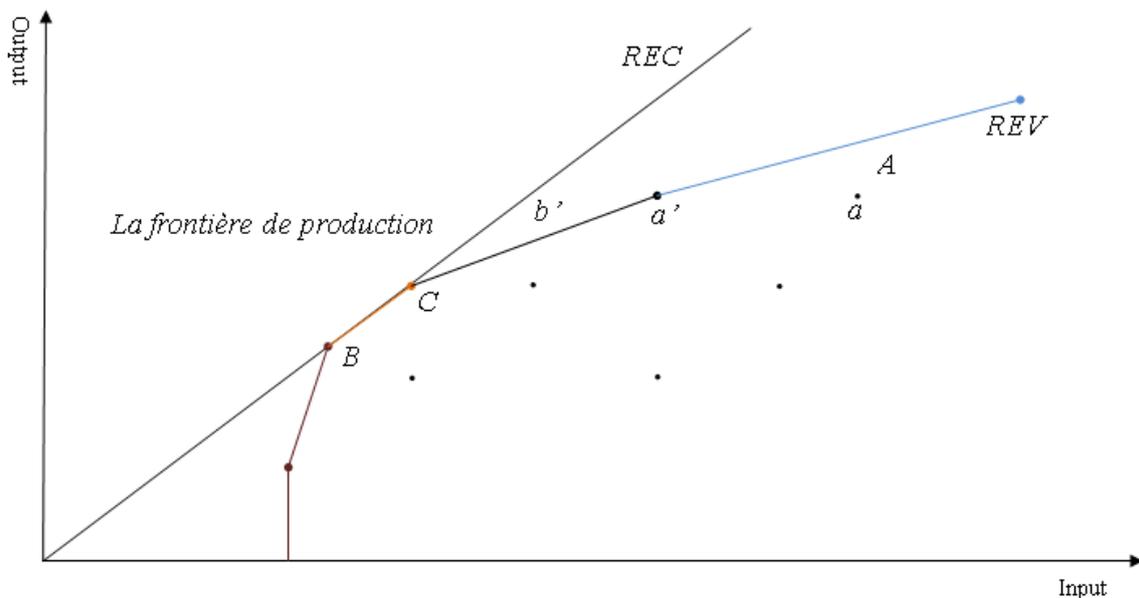


Figure 3-2 : Frontière d'efficacité estimée par CCR et BCC

À partir de la Figure 3-2, nous pouvons voir que les deux frontières d'efficacités sont estimées séparément, l'une par le modèle CCR sous l'hypothèse REC et l'autre par le modèle BCC sous l'hypothèse REV. Toutes les entreprises à gauche du point C produisent avec des rendements d'échelle croissants, c'est-à-dire que la taille est au-dessous de la taille optimale. Toutes les entreprises à droite du point C produisent avec des rendements d'échelle décroissants, leur taille est plus grande que la taille optimale. Maintenant regardons l'entreprise A, sous l'hypothèse REC, l'efficacité est donnée par le ratio bb'/ba . Par contre, sous l'hypothèse REV, l'inefficacité est peut-être causée par la taille sous-optimale des firmes et l'efficacité est mesurée par le ratio ba'/ba . Donc l'efficacité d'échelle est donnée simplement par bb'/ba' . L'efficacité technique d'une firme peut se composer sous la forme suivante :

$$\theta_{REC} = \theta_{REV} \times s(\mathbf{x}, \mathbf{k}, \mathbf{y}), \quad (3-2)$$

avec θ_{REC} le coefficient d'efficacité obtenu par le modèle CCR, θ_{REV} le coefficient d'efficacité obtenu par le modèle BCC et $s(\mathbf{x}, \mathbf{k}, \mathbf{y})$ le coefficient d'efficacité d'échelle.

Toute amélioration dans l'efficacité technique peut découler d'une amélioration de la gestion des ressources au sein de l'entreprise ou d'un changement de taille des firmes afin de tirer l'avantage des économies d'échelle.

La mesure d'efficacité d'échelle est définie comme :

$$s(\mathbf{x}, \mathbf{k}, \mathbf{y}) = \frac{\theta_{REC}(\mathbf{x}, \mathbf{k}, \mathbf{y})}{\theta_{REV}(\mathbf{x}, \mathbf{k}, \mathbf{y})} \leq 1. \quad (3-3)$$

D'une part, si $s(\mathbf{x}, \mathbf{k}, \mathbf{y}) = 1$, la DMU est efficace à l'échelle, ce qui signifie que la combinaison des inputs et outputs est optimale, l'entreprise fonctionne à sa taille optimale. D'autre part, si $s(\mathbf{x}, \mathbf{k}, \mathbf{y}) < 1$, la DMU n'est pas efficace à l'échelle, et pour déterminer la nature d'inefficacité, nous utilisons un autre rapport :

$$\eta(\mathbf{x}, \mathbf{k}, \mathbf{y}) = \frac{\theta_{RENC}(\mathbf{x}, \mathbf{k}, \mathbf{y})}{\theta_{REV}(\mathbf{x}, \mathbf{k}, \mathbf{y})} \leq 1, \quad (3-4)$$

où $\theta_{niiri}(\mathbf{x}, \mathbf{k}, \mathbf{y})$ est le coefficient d'efficacité obtenu sous l'hypothèse rendements d'échelle non croissants (RENC), il peut être dérivé du modèle suivant :

$$\theta_{RENC} = \min_{\theta_{RENC}} \left\{ \theta_{RENC} \left| \sum_{j=1}^N \lambda_j x_{sj} \leq \theta_{RENC} x_{sRENC}, k_{h0} \geq \sum_{j=1}^N k_{hj} \lambda_j, \sum_{j=1}^N \lambda_j y_{rj} \geq y_{r0}, \sum_{j=1}^N \lambda_j = 1 \right. \right. \\ \left. \left. \forall r = 1, 2, \dots, q; \forall s = 1, 2, \dots, p, \forall h = 1, 2, \dots, m \right\}. \quad (3-5)$$

Si $\eta(\mathbf{x}, \mathbf{k}, \mathbf{y}) = 1$, la technologie de la DMU est caractérisée par des rendements d'échelle décroissants ; par contre, si $\eta(\mathbf{x}, \mathbf{k}, \mathbf{y}) < 1$, les rendements d'échelle sont croissants.

La méthode présentée ci-dessus est une méthode qualitative. Pour mesurer quantitativement les rendements d'échelles des DMU avec des inputs quasi-fixes, nous pouvons utiliser l'équation suivante (Ouellette, Quesnel et Vigeant, 2012) :

$$RTS = \left[1 - \frac{v_0}{\theta + \sum_{s=1}^l v_s^k k_{sh}} \right]^{-1}. \quad (3-6)$$

Dans cette partie, nous avons présenté les deux modèles classiques en DEA : le modèle CCR et le modèle BCC. Ensuite, nous avons introduit les inputs quasi-fixes dans ces modèles et présenter finalement la façon de vérifier le nature des rendements d'échelle des entreprises. L'avantage de cette approche est très claire : elle s'intéresse aux observations individuelles plutôt qu'aux moyennes; elle ne nécessite pas la spécification de la fonction de transformation et permet de déterminer la meilleure pratique possible ainsi que les potentiels d'amélioration pour chaque DMU. En d'autres termes, elle indique les changements souhaitables au niveau des inputs et des outputs pour les unités de décision qui seraient sous de la frontière d'efficacité. Cependant plusieurs critiques ont été adressées à l'encontre de cette méthode. Entre

autres, nous ne pouvons pas mesurer les caractéristiques technologiques directement car nous n'avons pas la forme de la fonction de production spécifique et le résultat est très sensible à la qualité des données car les valeurs extrêmes peuvent influencer le résultat.

3-2 Modélisation du modèle semi-paramétrique

Dans cette section, nous allons présenter d'une manière plus précise le modèle semi-paramétrique. Pour ce faire, en nous basant sur l'article original d'Aigner et Chu (1968), nous exposerons d'abord le modèle primal en intégrant les inputs quasi-fixes. Nous passerons ensuite au modèle semi-paramétrique en coût. Nous terminons par une présentation d'une solution possible pour extraire une caractérisation statistique des coefficients estimés.

Nous pensons qu'il est malheureux que ce modèle n'ait pas suscité beaucoup d'attention malgré qu'il ait beaucoup d'avantages comme nous allons de le voir. Il reste aussi du travail de développement de ce modèle à faire, notamment dans l'espace dual. Nous tentons de remédier à ces constats dans notre thèse.

3-2-1 Le modèle semi-paramétrique primal

Dans le prolongement des analyses pionnières de Farrell (1957), Aigner et Chu (1968) propose un modèle semi-paramétrique, où il combine l'avantage des méthodes paramétriques et non paramétriques. Il permet de choisir une forme fonctionnelle, d'introduire facilement les inputs quasi-fixes et de procéder à des analyses sur la nature de la technologie (les rendements d'échelle, les élasticités de substitution, etc.) sans avoir à faire des hypothèses sur la distribution conjointe des termes d'erreur. Malheureusement, la méthode d'Aigner et Chu possède aussi certains défauts des méthodes paramétriques et non paramétriques. Au stade actuel de nos connaissances, il n'est pas possible de réaliser une inférence statistique dans ce modèle car les mesures n'ont pas de propriétés statistiques suffisantes démontrées jusqu'à maintenant; le choix de la forme fonctionnelle peut être limitant et les observations

extrêmes influencent les résultats obtenus. Le modèle proposé par Aigner et Chu est un modèle représentatif de frontières déterministes. Nous pouvons le considérer comme une modification de la frontière de production de Farrell. La frontière de Farrell est une frontière enveloppe pour chaque DMU. La fonction objective de Farrell est d'essayer d'obtenir le minimum de distance entre l'output réel et l'output efficace pour cette DMU, donc la frontière construite est en fait un polyèdre convexe composé par plusieurs surfaces pleines. La frontière de Farrell est en fait une combinaison des programmes linéaires. Par contre, le modèle d'Aigner et Chu essaye de construire une frontière pour toutes les DMU en même temps. Dans ce modèle, nous cherchons à minimiser la somme de la distance entre l'output observé et l'output efficace pour toutes les DMU. La frontière d'Aigner et Chu peut prendre plusieurs formes selon les différentes formes fonctionnelles choisies.

Dans notre thèse, nous proposons d'introduire les infrastructures dans la fonction de production. Puisque les DMU ne peuvent pas contrôler la quantité des infrastructures fournie à court terme, il est naturel que nous les introduisons dans la fonction de production comme des inputs quasi-fixes. Généralisons le cas de plusieurs inputs variables et plusieurs inputs quasi-fixes dans le processus de production : pour une DMU $i (i = 1, 2, \dots, N)$ qui utilisent P inputs variables $\mathbf{x} = (x_1, x_2, \dots, x_p)$ et S inputs quasi-fixes $\mathbf{k} = (k_1, k_2, \dots, k_s)$ pour produire un output y , la fonction de production sous forme Cobb-Douglas peut s'écrire :

$$y_i = e^{\alpha_0} \prod_{p=1}^P x_p^{\alpha_p} \prod_{s=1}^S k_s^{\beta_s} U_i. \quad (3-7)$$

Prenons la forme logarithme :

$$\ln y_i = \alpha_0 + \sum_{p=1}^P \alpha_p \ln x_{pi} + \sum_{s=1}^S \beta_s \ln k_{si} + \varepsilon_i = \alpha_0 + \boldsymbol{\alpha}^T \ln x_i + \boldsymbol{\beta}^T \ln k_i - \mu_i. \quad (3-8)$$

La partie déterministe à la droite de l'équation est considérée comme étant la frontière. C'est parce que la partie stochastique est entièrement incluse dans le terme

$-\mu_i$ et que nous ne modélisons pas ce terme dans le modèle que nous avons ici un modèle de frontière déterministe. Aigner et Chu (1968) proposent deux méthodes pour calculer les paramètres de façon à limiter les résidus μ_i aux valeurs non négatives. Prenons $-\mu_i = \varepsilon_i$, dans un premier temps, nous avons un programme linéaire :

$$\begin{aligned} \min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i \\ \text{s.c : } \ln y_i - \alpha_0 - \boldsymbol{\alpha}^T \ln \mathbf{x}_i - \boldsymbol{\beta}^T \ln \mathbf{k}_i \leq 0 \quad \forall i \\ \alpha_0 \geq 0, \boldsymbol{\alpha}^T \geq 0, \boldsymbol{\beta}^T \geq 0 \end{aligned} \quad (3-9)$$

ou un programme quadratique :

$$\begin{aligned} \min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i^2 \\ \text{s.c : } \ln y_i - \alpha_0 - \boldsymbol{\alpha}^T \ln \mathbf{x}_i - \boldsymbol{\beta}^T \ln \mathbf{k}_i \leq 0 \quad \forall i \\ \alpha_0 \geq 0, \boldsymbol{\alpha}^T \geq 0, \boldsymbol{\beta}^T \geq 0. \end{aligned} \quad (3-10)$$

Dans ces deux modèles, les variables d'écart dans les contraintes nous donne une mesure de ε_i . Avec cette méthode, l'inefficacité de chaque DMU est donnée par le ε_i estimé :

$$\hat{\varepsilon}_i = y_i - \left(\hat{\alpha}_0 + \hat{\boldsymbol{\alpha}}^T \mathbf{x}_i + \hat{\boldsymbol{\beta}}^T \mathbf{k}_i \right). \quad (3-11)$$

La moyenne de l'inefficacité technique, notée MIT, de l'ensemble de l'échantillon est mesurée par :

$$MIT = \frac{1}{N} \sum_{i=1}^N \hat{\varepsilon}_i. \quad (3-12)$$

3-2-2 Le modèle semi-paramétrique dual

Dans cette section nous suivrons Ouellette et Petit (2010) pour introduire une frontière de coût dans le modèle d'Aigner et Chu.

Pour une entreprise ou une DMU, la technologie peut être présentée dans l'espace primal par la fonction de production ou dans l'espace dual par la fonction de coût. Le problème de minimisation de coût peut être écrit comme suit :

$$C(\mathbf{w}, \mathbf{k}, y) = \min_x \{ \mathbf{w}^T \mathbf{x} \mid f(y, \mathbf{x}, \mathbf{k}) \leq 0 \}. \quad (3-13)$$

La solution pour ce problème est donnée par le vecteur des demandes conditionnelles de facteurs :

$$\mathbf{x} = \mathbf{x}(\mathbf{w}, \mathbf{k}, y). \quad (3-14)$$

Pour un échantillon, le coût observé de l'entreprise i est défini par :

$$C_i^{obs} = \sum_{p=1}^P w_p^{obs} x_{ip}^{obs}, \quad (3-15)$$

où C_i^{obs} est le coût observé pour la DMU i dans notre échantillon, w_p^{obs} représente le prix observé pour l'input p , et x_p^{obs} représente la quantité de l'input p observé pour la DMU i . En pratique, il peut toujours y avoir un écart entre le coût minimum théorique et le coût observé. Cet écart peut avoir deux sources : les erreurs d'observation et les erreurs d'optimisation. Les erreurs d'optimisation peuvent provenir de la mauvaise gestion et sont de nature aléatoire ou de mauvaise gestion répétée suite à un système d'incitations défaillant.

En supposant une structure d'erreurs additive, nous pouvons décrire le prix observé et les inputs observés par les équations suivantes² :

$$w_p^{obs} = w_p + \mathcal{E}_{w_p} \quad (3-16)$$

$$x_{ip}^{obs} = x_{ip} + \mathcal{E}_{x_p} + V_{x_p} \quad (3-17)$$

$$y_{ij}^{obs} = y_{ij} + \mathcal{E}_{y_j} \quad (3-18)$$

² Pour le modèle multiplicatif, voir Ouellette et Petit (2010).

$$k_{is}^{obs} = k_{is} + \varepsilon_{k_s}, \quad (3-19)$$

où w_p est le prix réel de l'input p , ε_{w_p} est l'erreur d'observation du prix pour l'input p , x_{ip} est la vraie quantité utilisée de l'input p par la DMU i ,³ ε_{x_p} est l'erreur d'observation pour l'input p , v_{x_p} est l'erreur d'optimisation pour l'input r , ε_{y_j} est l'erreur d'observation pour l'output j et ε_{k_s} est l'erreur d'observation pour l'input quasi-fixe s . Les quantités des inputs utilisés par une DMU découlent des comportements d'optimisation de la DMU. Les DMU sont considérées comme des preneurs de prix. Pour cette raison, nous supposons qu'il n'y a que l'erreur d'observation pour le prix des inputs, alors qu'il y a à la fois erreur d'observation et erreur d'optimisation pour la quantité des inputs variables.

Nous pouvons déduire l'écart entre le coût variable observé et le coût variable minimum par la formule suivante :

$$\begin{aligned} C(\mathbf{w}, \mathbf{k}, y) &= \mathbf{w}^T \mathbf{x}(\mathbf{w}, \mathbf{k}, y) \\ &= \left[\mathbf{w} + (\mathbf{w}^{obs} - \mathbf{w}^{obs}) \right]^T \left[\mathbf{x}(\mathbf{w}, \mathbf{k}, y) + (\mathbf{x}^{obs} - \mathbf{x}^{obs}) \right] \\ &= C^{obs} - (\mathbf{w}^{obs})^T (\varepsilon_x + v_x) - (\mathbf{x}^{obs})^T \varepsilon_w + \varepsilon_w^T (\varepsilon_x + v_x) \end{aligned} \quad (3-20)$$

ou encore :

$$\begin{aligned} C^{obs} &= C(\mathbf{w}, \mathbf{k}, y) + (\mathbf{w}^{obs})^T (\varepsilon_x + v_x) + (\mathbf{x}^{obs})^T \varepsilon_w - \varepsilon_w^T (\varepsilon_x + v_x) \\ &= C(\mathbf{w}, \mathbf{k}, y) + \left[(\mathbf{w}^{obs})^T \varepsilon_x + (\mathbf{x}^{obs})^T \varepsilon_w - \varepsilon_w^T \varepsilon_x \right] + (\mathbf{w}^{obs} - \varepsilon_w)^T v_x. \end{aligned} \quad (3-21)$$

Ici les variables sont bien sûr des vecteurs. Le coût observé est égale la somme du coût variable minimum ($C(\mathbf{w}, \mathbf{k}, y)$ le premier terme), des erreurs d'observations ($\varepsilon_c = (\mathbf{w}^{obs})^T \varepsilon_x + (\mathbf{x}^{obs})^T \varepsilon_w - \varepsilon_w^T \varepsilon_x$, le deuxième terme) et des erreurs d'optimisations (

³ Tous simplement, nous supposons que pour le même type d'input, l'erreur d'observation et l'erreur d'optimisation restent le même pour les différentes DMU.

$v_c = (\mathbf{w}^{obs} - \varepsilon_w)^T v_x$, le troisième terme). Donc l'écart entre le coût variable observé et le coût minimum, notée μ , peut s'exprimer par :⁴

$$\mu = \varepsilon_c + v_c. \quad (3-22)$$

Donc nous pouvons adapter le modèle d'Aigner et Chu (1968) pour une fonction de coût variable. Nous pouvons tout d'abord écrire la relation entre le coût observé et la fonction de coût minimale :

$$C^{obs} - C(\mathbf{w}^{obs} - \varepsilon_w, y^{obs} - \varepsilon_y, \mathbf{k}^{obs} - \varepsilon_k) - \varepsilon_c - v_c \geq 0. \quad (3-23)$$

Cette formulation soulève maintenant le problème du choix de la forme fonctionnelle (par exemple, la forme Cobb-Douglas comme dans le l'article original d'Aigner et Chu). Si nous supposons que la technologie est linéaire dans les paramètres, nous avons la forme suivante :

$$C^{obs} - (\mathbf{w}^{obs} - \varepsilon_w)^T \boldsymbol{\beta}_w - (y^{obs} - \varepsilon_y)^T \boldsymbol{\beta}_y - (\mathbf{k}^{obs} - \varepsilon_k)^T \boldsymbol{\beta}_k - \varepsilon_c - v_c \geq 0. \quad (3-24)$$

Maintenant, nous pouvons suivre l'idée d'Aigner et Chu en supposant que les termes d'erreur d'observation pour le prix, les inputs variables, les inputs quasi-fixes et l'output sont nuls, c'est-à-dire $\varepsilon_w = 0$, $\varepsilon_x = 0$, $\varepsilon_k = 0$ et $\varepsilon_y = 0$. Nous aurons $\varepsilon_c = (\mathbf{w}^{obs})^T \varepsilon_x + (\mathbf{x}^{obs})^T \varepsilon_w - \varepsilon_w^T \varepsilon_x = 0$ et $v_c = (\mathbf{w}^{obs} - \varepsilon_w)^T v_x = (\mathbf{w}^{obs})^T v_x$. Si nous linéarisons la fonction de coût après transformation logarithmique, l'équation (3-24) devient :

$$\ln C^{obs} - (\alpha + \boldsymbol{\beta}_w^T \ln \mathbf{w}^{obs} + \boldsymbol{\beta}_y^T \ln y^{obs} + \boldsymbol{\beta}_k^T \ln \mathbf{k}^{obs}) - (\mathbf{w}^{obs})^T v_x \geq 0. \quad (3-25)$$

Maintenant, nous pouvons suivre Aigner et Chu en imposant la non-négativité pour le terme d'erreur, pour chaque DMU i , nous avons notre fonction duale :

$$\ln C_i^{obs} - (\alpha + \boldsymbol{\beta}_w^T \ln \mathbf{w}_i^{obs} + \boldsymbol{\beta}_y^T \ln y_i^{obs} + \boldsymbol{\beta}_k^T \ln \mathbf{k}^{obs}) = (\mathbf{w}_i^{obs})^T v_{x_i} = \varepsilon_i \geq 0. \quad (3-26)$$

⁴ La démonstration est basée sur Ouellette et Petit (2010), page 5.

Pour illustrer la procédure, prenons une fonction de coût Cobb-Douglas avec Q output, P inputs variables et S inputs quasi-fixes :

$$C_i = e^{\alpha_o} \prod_{p=1}^P w_{pi}^{\alpha_p} \prod_{s=1}^S k_{si}^{-\beta_s} \prod_{q=1}^Q y_{qi}^{\beta_q} U_i, \quad (3-27)$$

où $U_i \geq 1$ représente l'efficacité de la DMU i . En prenant le log des deux côtés, le modèle devient log-linéaire :

$$\ln C_i = \alpha_o + \sum_{p=1}^P \alpha_p \ln w_{pi} - \sum_{s=1}^S \beta_s \ln k_{si} + \sum_{q=1}^Q \beta_q \ln Y_{qi} + \varepsilon_i, \quad (3-28)$$

où $\varepsilon_i = \ln U_i$ représente le terme d'erreur qui doit être non-négatif pour tout $i = 1, 2, \dots, n$.

Pour trouver les paramètres, il faut spécifier le critère d'optimisation. Pour la programmation linéaire, il s'agit en fait de minimiser la somme des erreurs sujette à n contraintes avec en plus la non-négativité des coefficients et du terme d'erreur, et l'homogénéité dans les prix des inputs variables. Pour la programmation quadratique, il s'agit de minimiser le terme d'erreur au carré comme pour les MCO et sous les mêmes contraintes. Pour le programme linéaire, nous avons :

$$\begin{aligned} \min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i &= \min_{\alpha, \beta} \sum_{i=1}^N \left(\ln C_i - \left(\alpha_o + \sum_{p=1}^P \alpha_p \ln w_{pi} - \sum_{s=1}^S \beta_s \ln k_{si} + \sum_{q=1}^Q \beta_q \ln Y_{qi} \right) \right) \\ \text{s.c : } \ln C_i - \left(\alpha_o + \sum_{p=1}^P \alpha_p \ln w_{pi} - \sum_{s=1}^S \beta_s \ln k_{si} + \sum_{q=1}^Q \beta_q \ln Y_{qi} \right) &\geq 0 \quad \forall i \end{aligned} \quad (3-29)$$

et pour le programme quadratique :

$$\begin{aligned} \min_{\alpha, \beta} \sum_{i=1}^N \varepsilon_i^2 &= \min_{\alpha, \beta} \sum_{i=1}^N \left(\ln C_i - \left(\alpha_o + \sum_{p=1}^P \alpha_p \ln w_{pi} - \sum_{s=1}^S \beta_s \ln k_{si} + \sum_{q=1}^Q \beta_q \ln Y_{qi} \right) \right)^2 \\ \text{s.c : } \ln C_i - \left(\alpha_o + \sum_{p=1}^P \alpha_p \ln w_{pi} - \sum_{s=1}^S \beta_s \ln k_{si} + \sum_{q=1}^Q \beta_q \ln Y_{qi} \right) &\geq 0 \quad \forall i. \end{aligned} \quad (3-30)$$

À partir de la fonction de coût, il est possible de tester certains indicateurs de performance : les rendements d'échelle, la relation entre les biens d'inputs, la diversification ainsi que la flexibilité dans le choix des inputs et des outputs.

L'efficacité technique sur l'ensemble de l'échantillon est mesurée par la moyenne de l'inefficacité de coût, MIC :

$$MIC = \frac{1}{N} \sum_{i=1}^N \varepsilon_i. \quad (3-31)$$

Conclusion

Dans le chapitre 3, nous avons présenté deux modèles non paramétriques et un modèle semi-paramétrique pour mesurer l'efficacité des entreprises. Pour le modèle non paramétrique, nous avons également présenté la façon d'introduire les inputs quasi-fixes. Dans le cas du modèle semi-paramétrique, en plus d'introduire les inputs quasi-fixes, nous avons également développé le modèle semi-paramétrique dans l'espace duale, via la fonction de coût.

Chapitre 4. La convergence des estimateurs et l'inférence statistique du modèle semi-paramétrique

L'approche proposée par Aigner et Chu est une approche intermédiaire entre l'approche purement économétrique et les approches fondées sur la recherche opérationnelle. Bien que cette méthode possède des avantages évidents, comme la simplicité pour caractériser la technologie, cette procédure n'est pas fondée sur un modèle statistique. Les propriétés statistiques des estimateurs sont donc ambiguës et les caractéristiques statistiques des estimateurs dépendent de la façon dont les données sont générées. L'inconvénient notable de ces estimateurs est qu'ils ne produisent pas d'écart-types pour les coefficients. Par contre, nous pouvons utiliser des méthodes de simulations (le "*bootstrap*" par exemple) pour déduire les propriétés statistiques des estimateurs et tenter de construire des intervalles de confiance valides. Dans les sous-sections qui suivent nous présenterons une caractérisation statistique des deux modèles de base d'Aigner et Chu (le modèle en programmation linéaire et en programmation quadratique). Les deux résultats les plus importants sont d'une part que les techniques habituelles pour déduire les propriétés statistiques sont inopérantes et, d'autre part, qu'en faisant appel à l'approche de Wald (1949) il est tout de même possible de démontrer la convergence de ces estimateurs. Armé de ce résultat, il devient alors possible de se livrer à un exercice d'inférence statistique par des méthodes de ré-échantillonnage. Donc le but de ce chapitre est d'abord de montrer la convergence des estimateurs du modèle semi-paramétrique, et ensuite proposer une méthode d'inférence statistique pour les paramètres estimés.

4-1 Sur la convergence du modèle de frontière de Aigner et Chu

L'inférence statistique n'a de sens que si l'estimateur utilisé est convergent. Par conséquent, il faut s'assurer que les estimateurs d'Aigner et Chu sont convergents et

c'est seulement une fois ce résultat démontré que nous pourrions nous interroger sur la nature de la loi de répartition de l'estimateur ou sur une façon d'en obtenir une pour nous permettre d'obtenir une forme d'inférence statistique sur les paramètres estimés. Comme nous l'avons annoncé plus haut, compte tenu des difficultés quant à l'obtention d'une loi de probabilité pour l'estimateur, nous aurons recours aux méthodes de simulations pour estimer la loi de répartition des différents paramètres du modèle.

Supposons que la frontière à estimer est linéaire dans les paramètres et que les observations de production (output), y_i , ne peuvent se trouver que "sous" la frontière.

La frontière de l'unité ou firme i est donnée par $X_i\beta$ et l'inefficacité est donnée par ε_i . La relation entre la production y_i et ces deux groupes de variables est :

$$y_i = \tilde{X}_i\tilde{\beta} + \varepsilon_i = \alpha + X_i\beta + \varepsilon_i, \quad (4-1)$$

$X_i\beta$ est la forme fonctionnelle de la frontière pour l'observation i . Elle est linéaire dans les paramètres mais pas nécessairement dans les variables. y_i représente l'output ou le coût pour l'entreprise i , $X_i = [X_{i1} \ \dots \ X_{ik}]$ est un vecteur ligne de variables explicatives de dimension $(1 \times k)$ pour l'observation i , où k représente le nombre de variables explicatives, $\alpha \in \mathbb{R}$ et $\beta = [\beta_1 \ \dots \ \beta_k]^T$ sont des paramètres de dimension (1×1) et $(k \times 1)$, ε_i représente le terme d'inefficacité, qui est de dimension (1×1) est par définition non positif. Par conséquent, nous avons $\tilde{X}_i = [1 \ X_i]$, et $\tilde{\beta}^T = [\alpha \ \beta^T]$.

Pour un échantillon de taille n , nous pouvons écrire les frontières de chaque unité de la façon suivante :

$$y = \alpha I_n + X\beta + \varepsilon, \quad (4-2)$$

où premièrement, $\mathbf{y}^T = [y_1 \ \dots \ y_n] = [y_i]$ est un vecteur ligne d'ordre $(1 \times n)$, qui représente les n observations de la variable dépendante y , α et $\boldsymbol{\beta}$ sont des paramètres de dimension (1×1) et $(k \times 1)$ respectivement, et \mathbf{I}_n est un vecteur colonne de « 1 » de dimension $(n \times 1)$. En plus, nous avons :

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{11} & \dots & \mathbf{X}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{n1} & \dots & \mathbf{X}_{nk} \end{bmatrix} = [\mathbf{X}_{is}] \quad (4-3)$$

qui est une matrice de dimension $(n \times k)$, composé de n observations pour chacune des k variables explicatives et finalement $\boldsymbol{\varepsilon}^T = [\varepsilon_1 \ \dots \ \varepsilon_n] = [\varepsilon_i]$ est un vecteur ligne composé des n observations de la variable d'inefficience ε . On notera une ligne de la matrice \mathbf{X} , $\mathbf{X}_i = [X_{i1} \ \dots \ X_{ik}]$ (un vecteur ligne de dimension $(1 \times k)$) avec X_{is} la variable explicative s , pour $s = 1, \dots, k$, de l'observation i , pour $i = 1, \dots, n$.

Revenons à notre fonction de frontière pour chaque observation, comme nous supposons que les observations ne peuvent se trouver que sur ou en-dessous de la vraie frontière, nous avons $\varepsilon_i \leq 0$. Par conséquent, nous en déduisons que :

$$y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}. \quad (4-4)$$

Nous pouvons associer le terme “d'erreur” ε_i à l'inefficacité. La modélisation statistique implique de donner un comportement probabiliste à ε_i .

4-1-1 Le modèle linéaire/exponentiel

Supposons dans un premier temps que les erreurs de ce modèle suivent une loi exponentielle, c'est-à-dire $f(\eta) = \lambda e^{-\lambda \eta}$ pour $\eta > 0$. Nous avons alors :

$$f(\varepsilon, \sigma) = \frac{1}{\sigma} e^{\frac{\varepsilon}{\sigma}}, \quad (4-5)$$

avec $-\varepsilon = \eta > 0$, $\sigma = 1/\lambda$ et $\lambda \neq 0$ (en fait, $\lambda > 0$). Montrons maintenant que nous avons un estimateur du maximum de vraisemblance.

4-1-1-1 Estimation par maximum vraisemblance

Sachant maintenant que $\varepsilon_i = \alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0 - y_i$ (l'indice inférieur 0 désigne le vrai paramètre), posons $\partial \varepsilon_i / \partial y_j = \delta_{ij}$, avec $\delta_{ii} = 1$ et $\delta_{ij} = 0$ pour $i \neq j$. Nous avons alors :

$$J(\varepsilon_i \rightarrow y_i) = \det \left[\frac{\partial \varepsilon_i}{\partial y_j} \right], \quad (4-6)$$

qui est le déterminant de la matrice $[\delta_{ij}]$ de dimension $(n \times n)$. Cela implique que nous avons $|J| = 1$. La vraisemblance de notre échantillon peut être écrite comme suit :

$$\begin{aligned} L(\sigma, \alpha, \boldsymbol{\beta} | y) &= \prod_{i=1}^n \frac{1}{\sigma} \exp \left\{ \frac{1}{\sigma} (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} |J| \\ &= \prod_{i=1}^n \frac{1}{\sigma} \exp \left\{ \frac{1}{\sigma} (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} \\ &= \left(\frac{1}{\sigma} \right)^n \exp \left\{ \frac{1}{\sigma} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} \end{aligned} \quad (4-7)$$

Le logarithme naturel de la fonction de vraisemblance est :

$$\ell(\sigma, \alpha, \boldsymbol{\beta} | y) = \ln L(\sigma, \alpha, \boldsymbol{\beta} | y) = -n \ln \sigma + \frac{1}{\sigma} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}), \quad (4-8)$$

avec $y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta} \leq 0, \forall i = 1, \dots, n$. Adoptons la notation $\ell_w = \partial \ell / \partial w$, alors les dérivées premières de la vraisemblance sont :

$$\begin{aligned} \ell_\sigma &= -\frac{n}{\sigma} - \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \\ \ell_\alpha &= -\frac{n}{\sigma} \\ \ell_{\boldsymbol{\beta}} &= -\frac{1}{\sigma} \sum_{i=1}^n \mathbf{X}_i. \end{aligned} \quad (4-9)$$

ℓ_σ et ℓ_α sont deux scalaires et ℓ_β est un vecteur de dimension $(1 \times k)^5$.

À partir de $\ell_\sigma = 0$, nous pouvons solutionner pour σ , ce qui donne :

$$\hat{\sigma}_n = -\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}), \quad (4-10)$$

où n indique la taille de l'échantillon⁶.

Maintenant, réinsérerons cette expression dans la vraisemblance pour obtenir la vraisemblance concentrée :

$$\begin{aligned} \ell^c(\alpha, \boldsymbol{\beta} | y) &= -n \ln \hat{\sigma}_n - \frac{n}{\hat{\sigma}_n} \hat{\sigma}_n \\ &= -n \ln \hat{\sigma}_n - n \\ &= -n \ln \left\{ \frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} - n. \end{aligned} \quad (4-11)$$

La vraisemblance concentrée, $\ell^c(\alpha, \boldsymbol{\beta} | y)$, est maximisée par la minimisation du premier terme, sous la contrainte de négativité de ε_i . Puisque le logarithme naturel est une transformation monotone, ceci est équivalent à la minimisation suivante :

$$\begin{aligned} \min_{\alpha, \boldsymbol{\beta}} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \\ s.c : y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}, \forall i = 1, \dots, n. \end{aligned} \quad (4-12)$$

Ceci est exactement le problème d'Aigner et Chu pour le programme linéaire. En somme, nous avons le résultat suivant :

⁵ Puisque $\boldsymbol{\beta}$ est un vecteur colonne de dimension $(k \times 1)$, la dérivée d'un scalaire par un vecteur colonne donne un vecteur ligne.

⁶ Nous utilisons la façon classique l'accent circonflexe pour désigner un estimateur et l'indice inférieur n pour donner la taille de l'échantillon pour lequel il a été calculé. Ceci a pour but de garder à l'esprit que la convergence est étudiée lorsque n tend à l'infini, $n \rightarrow \infty$.

Proposition 4.1. *La solution du programme linéaire de minimisation de la somme des erreurs unilatères d'Aigner et Chu est un estimateur du maximum de vraisemblance lorsque l'inefficacité est de loi exponentielle.*

Nous pouvons trouver le même résultat dans Schmidt (1975). Le fait d'avoir un estimateur du maximum de vraisemblance est intéressant en soit, mais ce ne sera pas très utile tant que nous n'aurons pas explicité ses propriétés statistiques. Il est par conséquent désirable d'établir la convergence de l'estimateur et de déduire une loi de répartition. La technique habituelle pour un estimateur du maximum de vraisemblance est d'utiliser la théorie asymptotique pour déduire la convergence et la normalité à partir du comportement du score et de la matrice d'information. Cependant, les outils habituels ne sont pas applicables dans le cas qui nous occupe. Procédons par étape pour le démontrer.

L'approche traditionnelle utilise une condition d'identification pour les paramètres de la vraisemblance. La condition limite est généralement de la forme :

$$E[S_n(\boldsymbol{\theta})] = 0 \text{ ou } \lim_{n \rightarrow \infty} S_n(\boldsymbol{\theta}) = 0, \quad (4-13)$$

avec $\boldsymbol{\theta} = [\sigma \quad \alpha \quad \boldsymbol{\beta}^T]$ et $S_n(\boldsymbol{\theta}) = n^{-1} \partial \ln L / \partial \boldsymbol{\theta}$ est le score de la vraisemblance.

Dans le cas qui nous occupe, nous avons :

$$\begin{aligned} E[n^{-1} \ell_\sigma] &= E \left[\frac{1}{n} \left(-\frac{n}{\sigma} - \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right) \right] \\ &= -\frac{1}{\sigma} - \frac{1}{\sigma^2} E \left[\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right] \\ &= -\frac{1}{\sigma} - \frac{1}{\sigma^2} (-\sigma) \\ &= 0 \end{aligned} \quad (4-14)$$

$$E[n^{-1} \ell_\alpha] = E \left[-\frac{1}{n} \frac{n}{\sigma} \right] = -\frac{1}{\sigma}$$

$$E[n^{-1} \ell_\beta] = E \left[-\frac{1}{n} \frac{1}{\sigma} \sum_{i=1}^n \mathbf{X}_i \right] = -\frac{1}{\sigma} \bar{\mathbf{X}},$$

où $\bar{X} = (1/n)1_n^T X = \sum_i^n X_{is}/n, s = 1, \dots, k$, est un vecteur ligne de dimension $(1 \times k)$, qui représente la moyenne arithmétique de la variable explicative pour un échantillon de taille n , k est le nombre de variables explicatives autres que la constante (par exemple, pour la forme Cobb-Douglas, c'est le nombre d'inputs).

Nous obtenons :

$$E[S(\theta)] = \left[0 \quad -\left(\frac{1}{\sigma}\right) \quad -\left(\frac{1}{\sigma}\right)\bar{X} \right] \neq 0. \quad (4-15)$$

qui est un vecteur de dimension $(1 \times (1+1+k))$, car 0 est un scalaire, $-1/\sigma$ est un scalaire, et $(-1/\sigma)\bar{X}$ est un vecteur de dimension $(1 \times k)$.

Ceci démontre que l'espérance du score n'est pas nulle, donc la méthode de démonstration classique de la convergence utilisant le score n'est donc pas possible⁷. Cependant, cette technique de démonstration n'est pas unique et il est possible de procéder avec une autre approche. Plus loin, nous nous proposons de vérifier les conditions d'applicabilité du théorème de convergence de Wald (1949). Cette approche ne nécessite pas d'avoir recours au score ou à la matrice hessienne de la vraisemblance, les conditions étant restreintes plus ou moins exclusivement à la vérification de conditions de dominance stochastique. Cependant, avant de nous lancer sur cette voie, nous voudrions aussi voir de quelle façon les outils classiques peuvent ou non être utilisés pour étudier la loi asymptotique de cet estimateur.

Rien n'empêche de supposer que l'estimateur est convergent, seule la technique habituelle ne s'applique pas ici. Supposons-le et tentons de voir si nous pouvons établir une forme de normalité asymptotique. Ceci pourrait permettre d'établir directement une règle pour l'inférence statistique - la loi asymptotique se substitue alors à la loi de petit échantillon inconnue.

⁷ Voir Stuart, Ord et Arnold (1999, page 57) pour un exemple d'une telle démonstration utilisant la méthode de Cramér.

La méthode traditionnelle pour établir la convergence en loi pour un estimateur du maximum de vraisemblance nécessite l'utilisation de l'espérance de l'inverse de la matrice hessienne. Pour la construction de la matrice hessienne, rassemblons les dérivées secondes de la vraisemblance et prenons l'espérance de ces termes. Nous avons, dans un premier temps :

$$\begin{aligned}
\ell_{\sigma\sigma} &= \frac{n}{\sigma^2} + \frac{2}{\sigma^3} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \quad (\text{un scalaire}) \\
\ell_{\alpha\alpha} &= \frac{n}{\sigma^2} \quad (\text{un scalaire}) \\
\ell_{\sigma\boldsymbol{\beta}^T} &= \frac{1}{\sigma^2} \sum_i \mathbf{X}_i^T \quad (\text{un vecteur de dimension } (k \times 1)) \\
\ell_{\alpha\alpha} &= 0 \quad (\text{un scalaire}) \\
\ell_{\alpha\boldsymbol{\beta}^T} &= \mathbf{0}_k \quad (\text{un vecteur de dimension } (k \times 1)) \\
\ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T} &= \mathbf{0}_{(k \times k)} \quad (\text{une matrice de dimension } (k \times k))
\end{aligned} \tag{4-16}$$

Il s'ensuit que :

$$\begin{aligned}
\mathbb{E} \left[n^{-1} \ell_{\sigma\sigma} \right] &= \frac{1}{\sigma^2} + \frac{2}{\sigma^3} (-\sigma) = \frac{1}{\sigma^2} \\
\mathbb{E} \left[n^{-1} \ell_{\alpha\alpha} \right] &= \frac{1}{\sigma^2} \\
\mathbb{E} \left[n^{-1} \ell_{\sigma\boldsymbol{\beta}^T} \right] &= \frac{1}{\sigma^2} \bar{\mathbf{X}}^T \quad (\text{un vecteur de dimension } (k \times 1)) \\
\mathbb{E} \left[n^{-1} \ell_{\alpha\alpha} \right] &= 0 \\
\mathbb{E} \left[n^{-1} \ell_{\alpha\boldsymbol{\beta}^T} \right] &= \mathbf{0}_k \quad (\text{un vecteur de dimension } (k \times 1)) \\
\mathbb{E} \left[n^{-1} \ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T} \right] &= \mathbf{0}_{(k \times k)} \quad (\text{une matrice de dimension } (k \times k))
\end{aligned} \tag{4-17}$$

En rassemblant ces termes, nous obtenons la matrice hessienne de la vraisemblance :

$$-\mathbb{E} \left[\frac{1}{n} \frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right] = -\mathbb{E} \begin{bmatrix} \ell_{\sigma\sigma} & \ell_{\alpha\sigma} & \ell_{\boldsymbol{\beta}\sigma} \\ \ell_{\sigma\alpha} & \ell_{\alpha\alpha} & \ell_{\boldsymbol{\beta}\alpha} \\ \ell_{\sigma\boldsymbol{\beta}^T} & \ell_{\alpha\boldsymbol{\beta}^T} & \ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & -\frac{1}{\sigma^2} & -\frac{1}{\sigma^2} \bar{\mathbf{X}} \\ -\frac{1}{\sigma^2} & 0 & \mathbf{0}_k^T \\ -\frac{1}{\sigma^2} \bar{\mathbf{X}}^T & \mathbf{0}_k & \mathbf{0}_{(k \times k)} \end{bmatrix}. \tag{4-18}$$

Puisque le déterminant de cette matrice est nul, elle est singulière. Par conséquent, la méthode de Cramér pour établir la normalité asymptotique ne fonctionne pas. Il devient alors hasardeux de déduire une loi asymptotique traditionnelle pour cet estimateur. C'est ici que s'arrête en général l'analyse classique (voir Green, 1980).

Par contre, si la convergence est assurée, la loi de répartition des estimateurs peut être approchée par la méthode de ré-échantillonnage (*bootstrap*) si le modèle statistique est adéquatement construit. En résumé, nous ne pouvons pas utiliser les méthodes habituelles pour déduire la convergence de l'estimateur et il n'est pas non plus possible d'obtenir facilement une caractérisation asymptotique des lois de répartition de l'estimateur. Cependant, si nous pouvons établir la convergence de l'estimateur, il devient alors possible d'utiliser des méthodes de simulation pour inférer une caractérisation statistique de l'estimateur.

La suite de cette section présente une démonstration de la convergence de l'estimateur de Aigner et Chu sous l'hypothèse que l'inefficacité est de loi exponentielle en utilisant la démonstration de Wald (1949) de la convergence d'un estimateur du maximum de vraisemblance.

4-1-1-2 Convergence de l'estimateur.

La méthode de Wald fait appel à un ensemble de conditions différentes de la méthode habituelle et elle a entre autres la caractéristique de ne pas faire usage des dérivées premières et secondes de la vraisemblance. L'esprit de la démonstration que nous présentons ici repose sur la vérification des hypothèses de Wald (1949). En fait, ces hypothèses fournissent un environnement idéal pour vérifier la convergence de la solution du modèle d'Aigner et Chu puisqu'elle évite les problèmes aux bornes et n'utilise pas tant le score que la matrice hessienne de la vraisemblance.

Pour rappel, nous avons la vraisemblance suivante :

$$L(\sigma, \alpha, \beta | y) = \left(\frac{1}{\sigma}\right)^n \exp\left\{\frac{1}{\sigma} \sum_{i=1}^n (y_i - \alpha - X_i \beta)\right\},$$

avec $y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta} \leq 0 \forall y_i, \mathbf{X}_i$. La log-vraisemblance est donnée par :

$$\ell(\sigma, \alpha, \boldsymbol{\beta} | y) = -n \ln \sigma + \frac{1}{\sigma} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}).$$

Pour pouvoir démontrer la convergence nous devons contrôler le comportement de la vraisemblance dans les valeurs extrêmes. Les définitions qui suivent seront donc des éléments centraux pour la convergence. En fait, les seules difficultés de la démonstration sont liées au comportement de ces fonctions.

Définition 4.1. Soit, pour la fonction de densité exponentielle et un vecteur de variables explicatives \mathbf{X}_i

$$f(y, \boldsymbol{\theta}) = f(y, \sigma, \alpha, \boldsymbol{\beta}) = \frac{1}{\sigma} \exp \left\{ \frac{1}{\sigma} (y - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} \text{ avec } \boldsymbol{\theta} = [\sigma \quad \alpha \quad \boldsymbol{\beta}^T]. \quad (4-19)$$

i. Pour un $\boldsymbol{\theta}$ et un ρ donnés :

$$f(y, \boldsymbol{\theta} | \rho) := \sup_{\boldsymbol{\theta}'} f(y, \boldsymbol{\theta}') = \sup_{\sigma', \alpha', \boldsymbol{\beta}'} \frac{1}{\sigma'} \exp \left\{ \frac{1}{\sigma'} (y - \alpha' - \mathbf{X}_i \boldsymbol{\beta}') \right\} \quad (4-20)$$

$$\text{avec } \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \rho \text{ et } \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| = \left[(\sigma - \sigma')^2 + (\alpha - \alpha')^2 + (\boldsymbol{\beta} - \boldsymbol{\beta}')^T (\boldsymbol{\beta} - \boldsymbol{\beta}') \right]^{1/2}.$$

ii. Pour un $\boldsymbol{\theta}$ et un r donnés :

$$\varphi(y, r) := \sup_{\boldsymbol{\theta}} f(y, \boldsymbol{\theta}) = \sup_{\sigma, \alpha, \boldsymbol{\beta}} \frac{1}{\sigma} \exp \left\{ \frac{1}{\sigma} (y - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\}. \quad (4-21)$$

$$\text{pour tout } \|\boldsymbol{\theta}\| \geq r, \text{ avec } \|\boldsymbol{\theta}\| = \left[\sigma^2 + \alpha^2 + \boldsymbol{\beta}^T \boldsymbol{\beta} \right]^{1/2}$$

iii.

$$f^*(y, \boldsymbol{\theta} | \rho) = \begin{cases} f(y, \boldsymbol{\theta} | \rho) & \text{si } f(y, \boldsymbol{\theta} | \rho) > 1 \\ 1 & \text{si } f(y, \boldsymbol{\theta} | \rho) \leq 1 \end{cases} \quad (4-22)$$

iv.

$$\varphi^*(y, r) = \begin{cases} \varphi(y, r) & \text{si } \varphi(y, r) > 1 \\ 1 & \text{si } \varphi(y, r) \leq 1 \end{cases} \quad (4-23)$$

Notons ici que X_i est simplement une ligne arbitraire de X . On aurait pu utiliser un vecteur de variables arbitraires, noté x . La seule chose importante ici est que ces X_i sont exogènes et conditionnent la position de la frontière.

Maintenant, réglons d'emblée une question technique. *A priori*, il n'y a pas de raison de borner l'ensemble des paramètres dans la démonstration de Wald, c'est d'ailleurs une des particularités de cette démonstration. En d'autres termes, chez Wald nous avons $\theta \in \Theta$ et Θ est uniquement fermé. Il faut au minimum supposer que l'espace des paramètres est fermé (il contient toutes les suites convergentes et le point de convergence - la suite converge dans l'ensemble). Donc, nous supposons que Θ est un ensemble fermé. Ceci étant, cette restriction est insuffisante pour nos besoins. Par exemple, dans le cas linéaire Cobb-Douglas, nous devons supposer que $\theta \geq 0$. Par conséquent, la généralité de l'hypothèse de fermeture est insuffisante et il nous faudra la renforcer.

En d'autres termes, l'hypothèse de Wald,

Hypothèse 4.0. *L'ensemble des paramètres θ , noté Θ , est un sous-ensemble de dimension \mathbb{R}^{k+2} , $\Theta \subset \mathbb{R}^{k+2}$ et Θ est fermé.*

peut s'avérer insuffisante dans notre environnement.

Dans nos applications, puisque nous essayons d'estimer les coefficients pour une fonction de production avec une forme précise, que les paramètres appartiennent à un ensemble borné va de soi. Cependant, il est possible aussi de vouloir les borner "loin" de zéro. Entre autres, le paramètre σ de la loi exponentielle ne peut pas prendre la valeur zéro. Si c'était le cas, la loi serait dégénérée en un point, ce qui enlèverait toute substance à notre modèle.

Nous supposons alors que l'espace des paramètres admissibles ne contient pas cette valeur pathologique. En d'autres termes, l'espace des paramètres sera contraint de telle sorte que $\sigma \in [\underline{\sigma}, \infty]$ avec $\underline{\sigma} > 0$. De plus, pour que le problème ait un sens économique, la valeur de la frontière doit être $\alpha + X_i \beta$ bornée inférieurement. Une façon d'obtenir cette borne est de supposer que, pour un M fini, possiblement négatif, $M \in \mathbb{R}$, $\alpha \geq M > -\infty$ et $\beta_i \geq M > -\infty$ pour tout $i = 1, \dots, k$. D'où l'hypothèse suivante :

Hypothèse 4.1. *L'espace des paramètres, Θ , est fermé avec $\Theta \subseteq [M, \infty)^{k+1} \times [\underline{\sigma}, \infty)$ et $\underline{\sigma} > 0$.*

Ceci nous assure que la septième condition de Wald est satisfaite. Les sept lemmes qui suivent nous permettront de démontrer que les sept conditions de Wald restantes tiennent bien.

Commençons par caractériser la fonction cumulative. En guise de notation, retenons que $F(y, \theta)$ est la fonction cumulative de la densité exponentielle $f(y, \theta)$, avec

$$F(y, \theta) = \int_y^{\alpha + X_i \beta} f(s, \theta) ds.$$

Lemme 4.1. *La fonction cumulative $F(y, \theta)$ est absolument continue.*

Démonstration. Ceci découle directement de la proposition 4.14 de Royden (1988) puisque f est continue pour toutes les valeurs de paramètres θ données.

CQFD

Donc la définition même du problème nous assure que la première condition du théorème de convergence de Wald est satisfaite. Les six lemmes suivants permettent de démontrer que les six autres conditions du théorème de convergence de Wald tiennent.

Lemme 4.2. Soit $\theta_0 = [\sigma_0 \quad \alpha_0 \quad \beta_0^T]$ le vecteur des vrais paramètres et soit ρ suffisamment petit et r suffisamment grand. Alors les espérances $\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \theta | \rho) dF(y, \theta_0)$ et $\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln \varphi^*(y, r) dF(y, \theta_0)$ sont finies pour la mesure de probabilité vraie, c'est-à-dire pour $F(y, \theta)$ évaluée en $\theta_0 = [\sigma_0 \quad \alpha_0 \quad \beta_0^T]$.

Démonstration. Commençons par démontrer que la valeur de $\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \theta | \rho) dF(y, \theta_0)$ est finie. Notons, dans un premier temps, que $\ln 1 = 0$ et que $\ln f^* \geq 0$. Pour un vecteur de paramètres admissibles donné, $[\sigma \quad \alpha \quad \beta^T] = \theta$ avec $\sigma > 0$, la fonction f est monotone décroissante avec l'inefficacité. Par conséquent, pour $y \leq \alpha + X_i \beta \forall \alpha, \beta$ admissibles, la fonction f est monotone croissante en y jusqu'en $y = \alpha + X_i \beta$. Par construction, nous avons aussi $f(y, \theta) = 0$ pour $\alpha + X_i \beta \leq y \leq \alpha_0 + X_i \beta_0$. Donc, supposons que f prenne des valeurs supérieures à un sur l'intervalle $[\underline{y}, \bar{y}]$ avec $\underline{y} < \alpha_0 + X_i \beta_0$, pour σ suffisamment petit.

Soit une boule fermée de θ , noté $N_\rho(\theta)$, tel que $\|\theta - \theta'\| \leq \rho$. Notons que l'on prend $N_\rho(\theta) \cap \Theta$ si $N_\rho(\theta)$ n'est pas entièrement dans Θ . $N_\rho(\theta)$ est compact. On remarque que f est continue à droite en θ , par conséquent elle atteint son maximum dans $N_\rho(\theta)$, $\max f = \sup f \leq e^M$ avec $M \leq \infty$ et ce pour tout y , puisque f est monotone croissante en y pour θ donné.

Soit σ tel que $f > 1$ et notons la fonction indicatrice $I(f > 1)$ pour $f > 1$, alors

$$\begin{aligned}
\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \boldsymbol{\theta} | \rho) dF(y, \boldsymbol{\theta}_0) &= \int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \boldsymbol{\theta} | \rho) f(y, \boldsymbol{\theta}_0) dy \\
&= \int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln[f(y, \boldsymbol{\theta} | \rho)] I(f > 1) f(y, \boldsymbol{\theta}_0) dy \\
&\leq \int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln e^M f(y, \boldsymbol{\theta}_0) dy \\
&= \ln e^M \int_{-\infty}^{\alpha_0 + X_i \beta_0} f(y, \boldsymbol{\theta}_0) dy \\
&= M \int_{-\infty}^{\alpha_0 + X_i \beta_0} f(y, \boldsymbol{\theta}_0) dy \\
&= M. ?
\end{aligned} \tag{4-24}$$

Ceci complète la démonstration de la première partie.

Maintenant, considérons l'intégrale $\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln \varphi^*(y, r) dF(y, \boldsymbol{\theta}_0)$. Comme dans le premier cas, $f(y, \boldsymbol{\theta}_0) dy$ peut remplacer de façon équivalente $dF(y, \boldsymbol{\theta}_0)$. Si, pour un X et y donnés, f prend des valeurs inférieures à un, alors $\ln 1 = 0$, l'intégrale est trivialement finie et le résultat tient.

Donc, nous devons considérer le cas des valeurs de f supérieures à un. Pour un X donné et pour toute valeur de y , la fonction $\sigma^{-1} \exp\{\sigma^{-1}(y - \alpha - X_i \beta)\}$ tend vers son supremum lorsque α et β sont choisis de telles sortes que $y - \alpha - X_i \beta \rightarrow 0$ et $\sigma \rightarrow \underline{\sigma}$. Donc, pour un X_i donné et une valeur de y , il faut choisir α et β pour maximiser (en fait, le supremum) la fonction sachant que σ sera choisi de telle sorte que $\sigma \rightarrow \underline{\sigma}$. Nous notons que puisque $\sigma \geq \underline{\sigma} > 0$, la fonction de densité est toujours bornée pour tout α et β tels que $\alpha + X_i \beta \geq y$ (i.e. y est admissible).

Donc nous devons choisir α et β tel que $\alpha + X_i\beta \rightarrow y$ (en fait $y = \alpha + X_i\beta$ tel que $\|(\alpha, \beta)\| > r$). Notons aussi que si $y - \bar{\alpha} - X_i\bar{\beta} < 0$ pour $\|(\bar{\alpha}, \bar{\beta})\|^2 - \underline{\sigma}^2 = r^2$ et $\sigma = \underline{\sigma}$ alors $f(y, \theta) < f(\bar{\alpha} + X_i\bar{\beta}, \bar{\theta})$ avec $y = \bar{\alpha} + X_i\bar{\beta}$ et $\sigma = \underline{\sigma}$. Alors le sup φ^* est atteint pour $\sigma = \underline{\sigma}$, puisque $f(y, \theta)$ est bornée entre zéro et M (f atteint son maximum en $\sigma = \underline{\sigma}$ et $y = \alpha + X_i\beta$ pour tout y).

Maintenant, puisque f est monotone en y , s'il n'existe pas de y tel que $f \geq 1$, pour

tout $\|\theta\| > r$, alors $\int_0^{\alpha_0 + X_i\beta_0} \log \varphi(y, r) dF = 0$. Supposons alors qu'il existe des valeurs de

y telles que $f \geq 1$. Nous avons alors :

$$\begin{aligned}
 \int_y^{\alpha_0 + X_i\beta_0} \ln \left\{ \sup_{\|\theta\| > r} f(y, \theta) \right\} dF(y, \theta_0) &\leq \int_y^{\alpha_0 + X_i\beta_0} \ln \{e^M\} dF(y, \theta_0) \\
 &\leq \int_{-\infty}^{\alpha_0 + X_i\beta_0} M dF(y, \theta_0) \\
 &= M \int_{-\infty}^{\alpha_0 + X_i\beta_0} dF(y, \theta_0) \\
 &= M \\
 &< \infty.
 \end{aligned} \tag{4-25}$$

CQFD

Ce premier lemme nous a permis d'établir, entre autres, que le comportement de la log-vraisemblance dans ses valeurs extrêmes n'est pas pathologiques, ou en d'autres mots, que le poids accordé aux valeurs extrêmes est relativement faible. Le lemme suivant étudie la continuité de la densité dans les paramètres.

Lemme 4.3. *Si la suite $\{\theta_n\}$ converge, $\lim_{n \rightarrow \infty} \theta_n = \theta$, alors $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ pour tout y sauf possiblement pour un ensemble de mesure zéro par rapport à $F(y, \theta_0)$ qui peut dépendre de θ mais pas de la suite.*

Démonstration. Puisque θ appartient à un ensemble fermé, toute suite converge dans l'ensemble et nous avons $\lim_{n \rightarrow \infty} \theta_n = \theta$. La fonction f est continue en dans ses paramètres pour tout $\theta \in \Theta$, sauf pour les α et β tels que $y = \alpha + X\beta$. Dans ce cas nous avons $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ pour $y \leq \alpha_n + X\beta_n$ et $\lim_{n \rightarrow \infty} f(y, \theta_n) = 0$ pour $y > \alpha_n + X\beta_n$. Puisque l'ensemble les α et β tels que $y = \alpha + X\beta$ forment un ensemble de mesure zéro dans Θ , alors toute suite convergente, $\lim_{n \rightarrow \infty} \theta_n = \theta$, implique que $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ sauf pour un ensemble de mesure zéro qui ne dépend pas de la suite.

CQFD

Nous avons une mesure continue sur un espace fermé, donc la suite ne peut pas conduire à des cas pathologiques, sauf à l'infini. Cependant nous avons montré au Lemme 4.2 que cela n'était pas possible. Ce lemme est presque une définition une fois la densité spécifiée. Le lemme suivant montre que la fonction cumulative exponentielle identifie les paramètres.

Lemme 4.4. *Si θ_1 est un vecteur de paramètres différent de θ_0 , c'est-à-dire $\theta_1 \neq \theta_0$, alors $F(y, \theta_1) \neq F(y, \theta_0)$ pour au moins une valeur de y .*

Démonstration. Par construction, pour deux valeurs différentes du vecteur de paramètres la densité exponentielle aura un nombre fini de point pour lesquels $f(y, \theta_1) = f(y, \theta_0)$. Par conséquent, sauf en un nombre fini de y , nous avons

$$f(y, \theta_1) = \frac{1}{\sigma_1} \exp\left\{\frac{1}{\sigma_1}(y - \alpha_1 - X_i\beta_1)\right\} \neq \frac{1}{\sigma_0} \exp\left\{\frac{1}{\sigma_0}(y - \alpha_0 - X_i\beta_0)\right\} = f(y, \theta_0). \quad (4-26)$$

La démonstration procède par construction en montrant que si $F(\tilde{y}, \theta_1) = F(\tilde{y}, \theta_0)$ en un point \tilde{y} , alors il existe au moins un autre point tel que $F(y, \theta_1) \neq F(y, \theta_0)$.

Considérons maintenant la valeur de F pour $\theta_1, \theta_0 \in \Theta$ tels que $\theta_1 \neq \theta_0$ au point \tilde{y} et supposons que :

$$F(\tilde{y}, \theta_1) = \int_{\tilde{y}}^{\alpha_1 + X_i \beta_1} \ln f(s, \theta_1) ds = \int_{\tilde{y}}^{\alpha_0 + X_i \beta_0} \ln f(s, \theta_0) ds = F(\tilde{y}, \theta_0). \quad (4-27)$$

Considérons alors $y = \tilde{y} - \Delta y$ pour Δy arbitrairement petit, nous avons alors :

$$\begin{aligned} F(y, \theta_1) &= \int_y^{\alpha_1 + X_i \beta_1} f(s, \theta_1) ds \\ &= \int_{\tilde{y}}^{\alpha_1 + X_i \beta_1} f(s, \theta_1) ds + \int_y^{\tilde{y}} f(s, \theta_1) ds \\ &\approx F(\tilde{y}, \theta_1) + f(y, \theta_1) \Delta y \end{aligned}$$

et

$$\begin{aligned} F(y, \theta_0) &= \int_y^{\alpha_0 + X_i \beta_0} f(s, \theta_0) ds \\ &= \int_{\tilde{y}}^{\alpha_0 + X_i \beta_0} f(s, \theta_0) ds + \int_y^{\tilde{y}} f(s, \theta_0) ds \\ &\approx F(\tilde{y}, \theta_0) + f(y, \theta_0) \Delta y. \end{aligned}$$

Alors

$$F(y, \theta_1) - F(y, \theta_0) \cong (f(y, \theta_1) - f(y, \theta_0)) \Delta y. \quad (4-28)$$

Et puisque $f(y, \theta_1) - f(y, \theta_0) \neq 0$ (sauf en un nombre fini de points, auquel cas nous prendrons alors $\bar{y} \neq y$) nous avons : $F(y, \theta_1) \neq F(y, \theta_0)$.

CQFD

Le prochain résultat permet de relâcher l'hypothèse forçant l'ensemble de paramètres à être compact.

Lemme 4.5. Si $\lim_{n \rightarrow \infty} \|\boldsymbol{\theta}_n\| = \infty$, alors $\lim_{n \rightarrow \infty} f(y, \boldsymbol{\theta}_n) = 0$ pour tout y , (à l'exception d'un ensemble de mesure nulle par rapport à $F(y, \boldsymbol{\theta}_0)$ qui peut dépendre de $\boldsymbol{\theta}_0$, mais pas de la suite $\{\boldsymbol{\theta}_n\}$).

Démonstration. De la loi de répartition exponentielle, $f(y, \boldsymbol{\theta}) = \sigma^{-1} \exp\{\sigma^{-1}(y - \alpha - \mathbf{X}_i \boldsymbol{\beta})\}$, on note que $\sigma_n^{-1}(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)$ est $O(1)$ lorsque $\|\boldsymbol{\theta}\| \rightarrow \infty$ avec $\sigma_i \rightarrow \infty$ et possiblement $\alpha, \boldsymbol{\beta} \rightarrow \infty$, donc nous avons :

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma_n} \exp\left\{\frac{y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n}{\sigma_n}\right\} \rightarrow 0 \quad \forall y \quad (4-29)$$

Dans le cas où $\alpha, \boldsymbol{\beta} \rightarrow \infty$, nous avons $\exp\{\sigma_n^{-1}(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)\} \rightarrow 0$ et le résultat suit de façon triviale. Finalement, la borne $\alpha + \mathbf{X}_i \boldsymbol{\beta} \geq y$ et le fait que $\|\boldsymbol{\theta}\| \rightarrow \infty$ sont suffisants pour assurer le comportement approprié de la densité pour les valeurs possiblement négatives de α et $\boldsymbol{\beta}$, puisque $\alpha \geq M > \infty$ et $\boldsymbol{\beta}_j \geq M > \infty, j = 1, \dots, k$ par l'hypothèse 4.1. En d'autres termes, on doit avoir que $\sigma \rightarrow \infty$ plus rapidement que $\|\tilde{\boldsymbol{\beta}}\| = \|\begin{bmatrix} \alpha & \boldsymbol{\beta}^T \end{bmatrix}\| \rightarrow \infty$.

CQFD

Le lemme suivant affirme que la log-vraisemblance est absolument bornée en mesure.

Lemme 4.6. Soit $\boldsymbol{\theta}_0$ le vecteur des vrais paramètres, alors

$$\int_{-\infty}^{\infty} |\ln f(y, \boldsymbol{\theta}_0)| dF(y, \boldsymbol{\theta}_0) \leq M < \infty.$$

Démonstration. Par la continuité de la loi exponentielle, nous avons $dF(y, \boldsymbol{\theta}_0) = f(y, \boldsymbol{\theta}_0) dy = \sigma_0^{-1} \exp\{\sigma_0^{-1}(y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0)\} dy$. Maintenant, par définition $y \leq \alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0$,

d'où :

$$\begin{aligned}
\int_{-\infty}^{\infty} \left| \ln f(y, \boldsymbol{\theta}_0) \right| dF(y, \boldsymbol{\theta}_0) &= \int_0^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| -\ln \sigma_0 + \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| \frac{1}{\sigma_0} \exp \left\{ \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right\} dy \\
&\leq \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| -\ln(\sigma_0) \right| \frac{1}{\sigma_0} \exp \left\{ \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right\} dy \\
&\quad + \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| \frac{1}{\sigma_0} \exp \left\{ \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right\} dy \\
&\leq \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| -\ln(\sigma_0) \right| dF_0 + \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| dF_0.
\end{aligned} \tag{4-30}$$

Le premier terme est clairement borné. Maintenant, notons que :

$$\lim_{y \rightarrow -\infty} \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \cdot \exp \left\{ \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right\} \rightarrow 0. \tag{4-31}$$

Nous pouvons, réécrire le dernier terme comme suit :

$$\begin{aligned}
&\int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| dF(y, \boldsymbol{\theta}_0) \\
&= \int_0^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \left| \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| dF(y, \boldsymbol{\theta}_0) + \int_{-\infty}^0 \left| \frac{1}{\sigma_0} (y - \mathbf{X}_i \boldsymbol{\beta}_0 - \alpha_0) \right| dF(y, \boldsymbol{\theta}_0).
\end{aligned} \tag{4-32}$$

Alors que le premier terme de (4-29) est clairement fini, c'est un simple exercice d'intégration de montrer que le second terme est fini. Ceci complète la démonstration.

CQFD

Lemme 4.7. $f(y, \boldsymbol{\theta} | \rho)$ est une fonction mesurable de y pour tout $\boldsymbol{\theta}$ et ρ .

Démonstration. On note que puisque Θ est fermé alors $N_\rho(\boldsymbol{\theta}) = \{\boldsymbol{\theta}' \mid \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \rho\}$ est tel que $N_\rho(\boldsymbol{\theta}) \subseteq \Theta$ ou $N_\rho(\boldsymbol{\theta}) \cap \Theta$ est un ensemble compact. Par la définition même de $f(y, \boldsymbol{\theta})$, $f(y, \boldsymbol{\theta} | \rho)$ est nécessairement définie et atteint son maximum pour la valeur minimale de σ dans $N_\rho(\boldsymbol{\theta})$ puisque $N_\rho(\boldsymbol{\theta})$ est compact et que bien entendu

cette valeur de σ est dans $N_\rho(\boldsymbol{\theta})$. En d'autres mots, c'est un point limite de ce voisinage. Maintenant, notons que pour un tel choix de σ , disons $\underline{\sigma} = \min \sigma \in N_\rho(\boldsymbol{\theta})$, nous avons que le supremum de

$$f(y, \boldsymbol{\theta}') = \sup_{\alpha, \boldsymbol{\beta}, \underline{\sigma}} \frac{1}{\underline{\sigma}} \exp \left\{ \frac{1}{\underline{\sigma}} (y - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \right\} \quad (4-33)$$

est atteint pour le plus petit $y - \alpha - \mathbf{X}_i \boldsymbol{\beta}$ possible. Donc, lorsque cela est possible, nous choisirons $\alpha, \boldsymbol{\beta}$, de telle sorte que $y - \alpha - \mathbf{X}_i \boldsymbol{\beta} = 0$, avec $\alpha, \boldsymbol{\beta} \in N_\rho(\boldsymbol{\theta})$. Pour un tel choix de α et $\boldsymbol{\beta}$, $f(y, \boldsymbol{\theta} | \rho)$ est constante et égale à son maximum. Pour tout autre y tel que $y - \alpha - \mathbf{X}_i \boldsymbol{\beta} < 0$ avec $\underline{\sigma}, \alpha, \boldsymbol{\beta} \in N_\rho(\boldsymbol{\theta})$, nous posons $\alpha = \underline{\alpha}$ et $\boldsymbol{\beta} = \underline{\boldsymbol{\beta}}$, ces valeurs de paramètres correspondant au maximum de $y - \alpha - \mathbf{X}_i \boldsymbol{\beta}$ dans $N_\rho(\boldsymbol{\theta})$, d'où $y - \underline{\alpha} - \mathbf{X}_i \underline{\boldsymbol{\beta}} \neq 0$. Donc, $f(y, \boldsymbol{\theta} | \rho)$ est soit constante ou décroissante et dans les deux cas, continue, sauf possiblement en un point, d'où $f(y, \boldsymbol{\theta} | \rho)$ est mesurable.

CQFD

Ce résultat découle aussi du Lemme 4.2 qui donne la continuité de $f(y, \boldsymbol{\theta} | \rho)$.

Théorème 4.1. Posons $\boldsymbol{\theta} = [\sigma \quad \alpha \quad \boldsymbol{\beta}^T]$ et définissons l'estimateur du maximum de

vraisemblance, $\hat{\boldsymbol{\theta}}_n$, comme l'argument qui maximise $-n \ln \sigma + (1/\sigma) \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})$

sous la contrainte $y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}$. C'est-à-dire :

$$\begin{aligned} \hat{\boldsymbol{\theta}}_n &= \arg \max_{\boldsymbol{\theta}} \{ \ell(\sigma, \alpha, \boldsymbol{\beta} | y) | y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta} \} \\ &= \arg \max_{\alpha, \boldsymbol{\beta}, \sigma} \left\{ -n \ln \sigma + \frac{1}{\sigma} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) | y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta} \right\}. \end{aligned}$$

Alors, sous l'Hypothèse 1, l'estimateur du maximum de vraisemblance, $\hat{\boldsymbol{\theta}}_n$, est convergent, $\text{plim}_{n \rightarrow \infty} \hat{\boldsymbol{\theta}}_n = \boldsymbol{\theta}_0$.

Démonstration. Puisque l'espace des paramètres est fermé par hypothèse, l'Hypothèse 7 de Wald (1949) est aussi satisfaite. Les Lemmes 4.1 à 4.7 nous assurent que les sept autres hypothèses de Wald (1949) sont vérifiées. Par conséquent, le résultat découle du Théorème 2 de Wald (1949).

CQFD

Le théorème et les hypothèses menant au théorème de Wald (1949) sont présentés à l'Annexe A.

4-1-2 Le modèle quadratique /semi-normal

Nous pouvons maintenant nous tourner vers le modèle avec une répartition des inefficacités semi-normale. Cette structure stochastique peut être mise en relation avec le modèle de programmation quadratique d'Aigner et Chu (1968), comme Schmidt (1975) l'a montré et, comme Green (1980), nous démontrerons que les propriétés statistiques de l'estimateur ne peuvent pas être déduites avec la technique classique habituelle.

4-1-2-1 Estimation par maximum vraisemblance

Comme dans le cas précédant, supposons que la frontière est linéaire dans les paramètres et que les observations de production, y_i , peuvent se trouver exclusivement sur ou sous la frontière. Nous avons alors, comme précédemment :

$$y_i = \tilde{X}_i \tilde{\beta} + \varepsilon_i = \alpha + X_i \beta + \varepsilon_i,$$

avec $\varepsilon_i \leq 0$. Dans ce cas nous avons :

$$y_i \leq \alpha + X_i \beta.$$

Pour se dégager du modèle de la section précédente, il faut se donner une nouvelle structure stochastique pour ε_i . Comme nous l'avons annoncé plus haut, nous supposons que les erreurs suivent de loi semi-normale :

$$f(\eta) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{\eta^2}{2\sigma^2}\right\}, \quad (4-34)$$

avec $\eta > 0$. Les moments de cette densité sont $\mu_\eta = (2\sigma^2/\pi)^{1/2} = 0,7978\sigma$ et $\sigma_\eta^2 = \sigma^2[(\pi-2)/\pi] = 0,3653\sigma^2$.

Par conséquent, lorsque l'inefficacité, ε , suit de loi semi-normale nous avons :

$$f(\varepsilon, \sigma) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{\varepsilon^2}{2\sigma^2}\right\}. \quad (4-35)$$

Comme dans le cas précédent, sachant que $J(\varepsilon_i \rightarrow y_i) = \det[\partial\varepsilon_i/\partial y_j]$. Nous avons alors $|J|=1$. Notons que les moments de ε_i sont identiques à ceux de η . La vraisemblance de notre échantillon peut être écrite comme suit :

$$\begin{aligned} L(\sigma^2, \alpha, \boldsymbol{\beta} | y) &= \prod_{i=1}^n \left(\sqrt{\frac{2}{\pi\sigma^2}} \right) \exp\left\{ \left(-\frac{1}{2}\right) \left[\frac{1}{\sigma} (y_i - \tilde{\mathbf{X}}_i \tilde{\boldsymbol{\beta}}) \right]^2 \right\} |J| \\ &= \prod_{i=1}^n \left(\sqrt{\frac{2}{\pi\sigma^2}} \right) \exp\left\{ -\left(\frac{1}{2\sigma^2}\right) (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right\} \\ &= \left(\frac{2}{\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp\left\{ -\left(\frac{1}{2\sigma^2}\right) \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right\}. \end{aligned} \quad (4-36)$$

La log-vraisemblance est :

$$\ell(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \ln L(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \quad (4-37)$$

avec $y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta} \leq 0$ ou $y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}$.

Adoptons la notation $\ell_w = \partial\ell / \partial w$. Pour w , un vecteur colonne, ℓ_w est un vecteur ligne du même ordre. Alors les dérivées premières de la vraisemblance sont :

$$\begin{aligned}
\ell_{\sigma^2} &= -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \\
\ell_{\alpha} &= \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \\
\ell_{\boldsymbol{\beta}} &= \frac{1}{\sigma^2} \left[\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta}) \mathbf{X}_i \right],
\end{aligned} \tag{4-38}$$

où ℓ_{σ^2} est un scalaire, ℓ_{α} est un scalaire, $\ell_{\boldsymbol{\beta}}$ est un vecteur de dimension $(1 \times k)$.

Avec $\ell_{\sigma^2} = 0$, nous trouvons une solution pour σ^2 :

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2, \tag{4-39}$$

avec n qui rappelle la taille de l'échantillon.

Pour obtenir la vraisemblance concentrée, nous réinsérerons cette expression dans l'équation (4-37):

$$\begin{aligned}
\ell^c(\sigma^2, \alpha, \boldsymbol{\beta} | y) &= \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - \frac{n}{2} \ln \hat{\sigma}_n^2 - \frac{1}{2\hat{\sigma}_n^2} \frac{n}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \\
&= \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - \frac{n}{2} \ln \hat{\sigma}_n^2 - \frac{n}{2} \frac{\hat{\sigma}_n^2}{\hat{\sigma}_n^2} \\
&= \frac{n}{2} \left(\ln\left(\frac{2}{\pi}\right) - 1 \right) - \frac{n}{2} \ln \hat{\sigma}_n^2 \\
&= \frac{n}{2} \left(\ln\left(\frac{2}{\pi}\right) - 1 \right) - \frac{n}{2} \ln \left(\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right).
\end{aligned} \tag{4-40}$$

Nous déduisons immédiatement que $\ell^c(\sigma^2, \alpha, \boldsymbol{\beta} | y)$ est maximisée par la minimisation du dernier terme de cette expression sous la contrainte de négativité. Mais puisque le logarithme naturel est une transformation monotone, ceci est équivalent à la minimisation suivante :

$$\begin{aligned}
\min_{\alpha, \boldsymbol{\beta}} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \\
s.c : y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}.
\end{aligned} \tag{4-41}$$

Ceci est exactement le problème d'Aigner et Chu pour le programme quadratique. Par conséquent, nous avons démontré la proposition suivante :

Proposition 4.2. *La solution du programme quadratique de minimisation de la somme des erreurs au carré d'Aigner et Chu est un estimateur du maximum de vraisemblance lorsque l'inefficacité suit une loi semi-normale.*

Nous pouvons trouver le même résultat dans Schmidt (1975). Cependant, nous avons les mêmes problèmes ici qu'avec le modèle linéaire et la loi exponentielle pour déduire les propriétés statistiques. Il est impossible de démontrer la convergence par les canaux habituels (méthode de Cramér) comme nous le mentionnions plus haut. Cependant, il n'est pas exclu de pouvoir démontrer cette convergence comme nous le verrons plus loin.

Un autre problème se pose aussi. La méthode classique pour démontrer la loi asymptotique de l'estimateur est inopérante ici aussi. Ceci semble être un problème insurmontable puisque la hessienne de la vraisemblance est singulière. Encore une fois, ceci impliquerait qu'il ne serait pas possible de conduire la moindre inférence statistique. Cependant, en dernière ressource, nous pouvons nous tourner vers les méthodes de simulation, entre autres les méthodes de ré-échantillonnage, ce qui permettrait d'obtenir une caractérisation statistique possible de l'estimateur.

Les deux problèmes que nous avons soulevés sont que l'espérance du gradient de la vraisemblance (le *score*) n'est pas nulle et que la hessienne de la vraisemblance est singulière.

Commençons par considérer le cas du score. Nous avons :

$$\begin{aligned}
 E\left[n^{-1}\ell_{\sigma^2}\right] &= E\left[\left(\frac{1}{n}\right)\left(\left(-\frac{n}{2}\right)\left(\frac{1}{\sigma^2}\right) + \left(\frac{1}{2\sigma^4}\right)\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right)\right] \\
 &= E\left[\left(-\frac{1}{2}\right)\left(\frac{1}{\sigma^2}\right)\right] + E\left[\left(\frac{1}{2\sigma^4}\right)\left(\frac{1}{n}\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right)\right] \\
 &= \left(-\frac{1}{2\sigma^2}\right) + \left(\frac{1}{2\sigma^4}\right)E\left[\left(\frac{1}{n}\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right)\right].
 \end{aligned} \tag{4-42}$$

Comme nous l'avons vu plus haut pour une loi semi-normale, nous avons

$\mu_{\varepsilon_i} = (2\sigma^2/\pi)^{1/2}$ et $\sigma_{\varepsilon_i}^2 = \sigma^2(\pi - 2/\pi)$. Il s'ensuit que :

$$E\left[\left(\frac{1}{n}\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right)\right] = \sigma^2\left(1 - \frac{2}{\pi}\right) + \sigma^2\frac{2}{\pi} = \sigma^2, \quad (4-43)$$

d'où,

$$E\left[n^{-1}\ell_{\sigma^2}\right] = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}\sigma^2 = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^2} = 0. \quad (4-44)$$

Maintenant, en utilisant nos résultats sur les moments de la loi semi-normale, nous obtenons :

$$E\left[n^{-1}\ell_{\alpha}\right] = E\left[\frac{1}{\sigma^2}\left(\frac{1}{n}\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})\right)\right] = \frac{1}{\sigma^2}\left(-\sigma\sqrt{\frac{2}{\pi}}\right) = -\sqrt{\frac{2}{\pi\sigma^2}} \quad (4-45)$$

et

$$E\left[n^{-1}\ell_{\boldsymbol{\beta}}\right] = E\left[\frac{1}{\sigma^2}\left(\frac{1}{n}\sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})\mathbf{X}_i\right)\right] = \frac{1}{\sigma^2}\bar{\mathbf{X}}(-\sigma)\sqrt{\frac{2}{\pi}} = -\sqrt{\frac{2}{\pi\sigma^2}}\bar{\mathbf{X}}, \quad (4-46)$$

où $E_0[n^{-1}\ell_{\sigma^2}]$ est un scalaire, $E_0[n^{-1}\ell_{\alpha}]$ est un scalaire, $E_0[n^{-1}\ell_{\boldsymbol{\beta}}]$ est un vecteur ligne de dimension $(1 \times k)$.

En définitive, nous avons :

$$E\left[n^{-1}\ell_{\boldsymbol{\theta}}\right] = \begin{bmatrix} 0 & -\sqrt{\frac{2}{\pi\sigma^2}} & -\sqrt{\frac{2}{\pi\sigma^2}}\bar{\mathbf{X}} \end{bmatrix} \neq 0, \quad (4-47)$$

avec $\boldsymbol{\theta} = [\sigma^2 \quad \alpha \quad \boldsymbol{\beta}^T]$. Ceci démontre que l'espérance du score n'est pas nulle.

Maintenant, tournons-nous vers le problème de la matrice hessienne de la vraisemblance, nous démontrons ici qu'elle est singulière. Nous avons :

$$\begin{aligned}
\ell_{\sigma^2\sigma^2} &= \frac{n}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2 && \text{(un scalaire)} \\
\ell_{\sigma^2\alpha} &= -\frac{1}{\sigma^4} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta}) && \text{(un scalaire)} \\
\ell_{\sigma^2\boldsymbol{\beta}^T} &= -\frac{1}{\sigma^4} \left(\sum_{i=1}^n (\alpha + \mathbf{X}_i\boldsymbol{\beta} - y_i) \mathbf{X}_i^T \right) && \text{(un vecteur de dimension } (k \times 1)) \\
\ell_{\alpha\alpha} &= -\frac{n}{\sigma^2} && \text{(un scalaire)} \\
\ell_{\alpha\boldsymbol{\beta}^T} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \mathbf{X}_i^T && \text{(un vecteur de dimension } (k \times 1)) \\
\ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T} &= -\frac{1}{\sigma^2} \mathbf{X}_i^T \mathbf{X}_i. && \text{(une matrice de dimension } (k \times k))
\end{aligned}
\tag{4-48}$$

En prenant l'espérance de ces termes, nous obtenons :

$$\begin{aligned}
\mathbb{E}\left[n^{-1}\ell_{\sigma^2\sigma^2}\right] &= \frac{1}{2\sigma^4} - \frac{1}{\sigma^6} \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right] = \frac{1}{2\sigma^4} - \frac{1}{\sigma^4} = -\frac{1}{2\sigma^4} \text{ (un scalaire)} \\
\mathbb{E}\left[n^{-1}\ell_{\sigma^2\alpha}\right] &= \mathbb{E}\left[-\frac{1}{\sigma^4} \left(\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta})\right)\right] = -\frac{1}{\sigma^4} (-\sigma) \left(\frac{2}{\pi}\right)^{1/2} = \frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} \text{ (un scalaire)} \\
\mathbb{E}\left[n^{-1}\ell_{\sigma^2\boldsymbol{\beta}^T}\right] &= \mathbb{E}\left[-\frac{1}{\sigma^4} \left(\frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i\boldsymbol{\beta}) \mathbf{X}_i\right)\right] = -\frac{1}{\sigma^4} (-\sigma) \left(\frac{2}{\pi}\right)^{1/2} \bar{\mathbf{X}}^T = \frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} \bar{\mathbf{X}}^T \\
&&& \text{(un vecteur de dimension } (k \times 1)) \\
\mathbb{E}\left[n^{-1}\ell_{\alpha\alpha}\right] &= -\frac{1}{\sigma^2} \text{ (un scalaire)} \\
\mathbb{E}\left[n^{-1}\ell_{\alpha\boldsymbol{\beta}^T}\right] &= -\frac{1}{\sigma^2} \bar{\mathbf{X}}^T \text{ (un vecteur de dimension } (k \times 1)) \\
\mathbb{E}\left[n^{-1}\ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T}\right] &= -\frac{1}{\sigma^2} \overline{\mathbf{X}^T \mathbf{X}}. \text{ (une matrice de dimension } (k \times k)) && \text{(4-49)}
\end{aligned}$$

En rassemblant tous ces termes, nous obtenons :

$$-\mathbf{E} \left[n^{-1} \frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right] = -\mathbf{E} \begin{bmatrix} \ell_{\sigma\sigma^T} & \ell_{\alpha\sigma^T} & \ell_{\boldsymbol{\beta}\sigma^T} \\ \ell_{\sigma\alpha^T} & \ell_{\alpha\alpha^T} & \ell_{\boldsymbol{\beta}\alpha^T} \\ \ell_{\sigma\boldsymbol{\beta}^T} & \ell_{\alpha\boldsymbol{\beta}^T} & \ell_{\boldsymbol{\beta}\boldsymbol{\beta}^T} \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sigma^4} & -\frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} & -\frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} \bar{\mathbf{X}} \\ -\frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} & \frac{1}{\sigma^2} & \frac{1}{\sigma^2} \bar{\mathbf{X}} \\ -\frac{1}{\sigma^3} \sqrt{\frac{2}{\pi}} \bar{\mathbf{X}}^T & \frac{1}{\sigma^2} \bar{\mathbf{X}}^T & \frac{1}{\sigma^2} \overline{\mathbf{X}^T \mathbf{X}} \end{bmatrix}$$

(4-50)

qui est une matrice de dimension $(k+2) \times (k+2)$.

Bien entendu cette matrice est symétrique et il devient facile de voir que les mineurs d'ordre deux des deux dernières lignes sont tous nuls et par conséquent cette matrice est singulière. Il s'ensuit que la méthode de Cramér pour démontrer la convergence achoppe et ne peut produire le résultat escompté.

En d'autres termes, bien que nous ayons un estimateur du maximum de vraisemblance, nous ne pouvons pas utiliser la méthode de Cramér pour démontrer tant la convergence que la normalité asymptotique. La convergence est le problème le plus facile à traiter et comme dans le cas avec la loi exponentielle, nous nous en remettons à la méthode de Wald (1949) pour la démontrer. Pour la loi asymptotique, nous ne voyons d'autres solutions que de nous tourner vers les méthodes de simulation, que nous présenterons plus loin.

4-1-2-2 Démonstration de la convergence

Encore une fois, nous allons vérifier les huit conditions de Wald. Pour rappel, la vraisemblance est donnée par :

$$L(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \left(\frac{2}{\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp \left\{ -\left(\frac{1}{2\sigma^2}\right) \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right\}.$$

La log-vraisemblance est :

$$\ell(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \ln L(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \frac{n}{2} \ln \left(\frac{2}{\pi} \right) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2,$$

avec $y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta} \leq 0$ ou $y_i \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}$.

Comme dans le cas de la loi exponentielle, l'aspect crucial de la démonstration est le comportement de la fonction dans les valeurs extrêmes tant sur son domaine que dans l'espace des paramètres. Afin de réaliser cette analyse, nous utilisons ici aussi un ensemble de fonctions qui rendent explicite les valeurs extrêmes que nous devons contrôler.

Définition 4.2. Pour la fonction de densité semi-normale

$$f(y, \boldsymbol{\theta}) = f(y, \alpha, \boldsymbol{\beta}, \sigma^2) = \sqrt{\left(\frac{2}{\pi \sigma^2} \right)} \exp \left\{ \left(-1/2\sigma^2 \right) (y - \alpha - \mathbf{X} \boldsymbol{\beta})^2 \right\}, \quad \text{avec}$$

$$\boldsymbol{\theta} = \left[\sigma^2 \quad \alpha \quad \boldsymbol{\beta}^T \right], \text{ nous définissons :}$$

i. Pour un ρ donné :

$$f(y, \boldsymbol{\theta} | \rho) := \sup_{\boldsymbol{\theta}'} f(y, \boldsymbol{\theta}') = \sup_{\alpha', \boldsymbol{\beta}', \sigma'^2} \sqrt{\frac{2}{\pi \sigma'^2}} \exp \left\{ -\frac{1}{2(\sigma'^2)} (y - \alpha' - \mathbf{X}_i \boldsymbol{\beta}')^2 \right\}, \quad (4-51)$$

$$\text{avec } \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \rho, \text{ et } \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| = \left[(\sigma^2 - \sigma'^2)^2 + (\alpha - \alpha')^2 + (\boldsymbol{\beta} - \boldsymbol{\beta}')^T (\boldsymbol{\beta} - \boldsymbol{\beta}') \right]^{1/2}.$$

ii. Pour un $r > 0$ donné :

$$\varphi(y, r) := \sup_{\boldsymbol{\theta}} f(y, \boldsymbol{\theta}) = \sup_{\alpha, \boldsymbol{\beta}, \sigma^2} \sqrt{\frac{2}{\pi \sigma^2}} \left\{ -\frac{1}{2\sigma^2} (y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right\}, \quad (4-52)$$

$$\text{pour tout } \|\boldsymbol{\theta}\| \geq r, \text{ avec } \|\boldsymbol{\theta}\| = \left[\sigma^2 + \alpha^2 + \boldsymbol{\beta}^T \boldsymbol{\beta} \right]^{1/2}.$$

iii.

$$f^*(y, \boldsymbol{\theta} | \rho) = \begin{cases} f(y, \boldsymbol{\theta} | \rho) & \text{si } f(y, \boldsymbol{\theta} | \rho) > 1 \\ 1 & \text{si } f(y, \boldsymbol{\theta} | \rho) \leq 1 \end{cases} \quad (4-53)$$

iv.

$$\varphi^*(y, r) = \begin{cases} \varphi(y, r) & \text{si } \varphi(y, r) > 1 \\ 1 & \text{si } \varphi(y, r) \leq 1. \end{cases} \quad (4-54)$$

Nous pouvons maintenant aborder la démonstration. Les paramètres à estimer sont σ^2 , α et β , soit $k+2$ paramètres d'où $\Theta \subset \mathbb{R}^{k+2}$. Puisque le problème doit être défini avec une inefficacité non nulle avec probabilité un, nous supposons, comme avec le cas de la loi exponentielle, que $\sigma^2 \geq \underline{\sigma}^2 > 0$. Pour des raisons bien évidentes d'existence de la frontière de production, on suppose que α et β sont finis et aussi que la variance est finie, mais d'un point de vue statistique imposer une borne arbitraire pourrait être plus gênante qu'autre chose et nous pourrions vouloir éviter une telle hypothèse. Par contre, pour que le problème ait un sens économique, la valeur de la frontière doit être compatible avec des valeurs de production positive. Par conséquent, la quantité $\alpha + X_i\beta$ est supposée être bornée inférieurement. Une façon d'obtenir cette borne est de supposer que $\alpha \geq M > -\infty$ et que $\beta_i \geq M > -\infty$ pour tout $i=1, \dots, k$. Notons que M peut être négatif. Comme chez Wald, nous supposons que l'espace des paramètres est fermé et ajouterons les bornes précédentes sur la valeur des paramètres. Par conséquent :

Hypothèse 4.2. *L'espace des paramètres est un sous-espace fermé de \mathbb{R}^{k+2} , tel que*

$$\Theta \subseteq [M, \infty)^{k+1} \times [\underline{\sigma}, \infty) \subseteq \mathbb{R}^{k+2}, \quad \underline{\sigma} > 0.$$

Les paramètres d'une Cobb-Douglas sont nécessairement positifs. Ce n'est cependant pas le cas de la forme Translog, d'où la nécessité d'avoir un espace des paramètres qui comprend des valeurs négatives. L'Hypothèse 1 donne automatiquement la septième condition de Wald, puisque nous n'avons que réduit le domaine de la variance en forçant des valeurs strictement positives et celui des paramètres de la frontière à être bornés. L'hypothèse sur σ^2 implique, entre autres, qu'il doit y avoir de l'inefficacité car la loi de probabilité ne peut pas être dégénérée en un point.

Comme précédemment, nous procédons par étape pour démontrer que les conditions demandés par le théorème de Wald pour la convergence tiennent. Pour ce faire, nous démontrons au travers d'une suite de lemmes que les hypothèses de Wald s'appliquent au cas de l'estimateur d'Aigner et Chu avec inefficacité semi-normale.

Commençons par caractériser la fonction cumulative. En guise de notation, retenons que $F(y, \theta)$ est la fonction cumulative de la densité semi-normale $f(y, \theta)$, avec

$$F(y, \theta) = \int_y^{\alpha + X_i \beta} f(s, \theta) ds. \text{ Bien entendu } F(y, \theta) \text{ n'existe pas sous forme explicite.}$$

Lemme 4.8. *La fonction cumulative $F(y, \theta)$ est absolument continue.*

Démonstration. Ceci découle directement de la proposition 4.14 de Royden (1988) puisque f est continue pour toutes les valeurs de paramètres θ données.

CQFD

Donc la définition même du problème nous assure que la première condition du Théorème 2 de convergence de Wald (1949) est satisfaite

Lemme 4.9. *Pour ρ suffisamment petit et r suffisamment grand, les espérances*

$$\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \theta | \rho) dF(y, \theta_0) \text{ et } \int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln \varphi^*(y, r) dF(y, \theta_0) \text{ sont finies pour la mesure de}$$

probabilité vraie, c'est-à-dire pour $F(y, \theta)$ définie pour le vecteur des vrais paramètres $\theta_0 = [\sigma_0^2 \quad \alpha_0 \quad \beta_0]$.

Démonstration. Commençons par démontrer que $\int_{-\infty}^{\alpha_0 + X_i \beta_0} \ln f^*(y, \theta | \rho) dF(y, \theta_0)$ est

finie. Par définition de f^* nous avons $\log f^* \geq 0$ puisque $\ln 1 = 0$. On note que la fonction f est monotone croissante en y et bornée pour tout $\sigma^2 \geq \underline{\sigma}^2$ étant donné α et β . Maintenant, notons que pour $y < \alpha + X_i \beta \quad \forall \alpha, \beta$ admissibles, la fonction f

est monotone croissante en y jusqu'en $y = \alpha + \mathbf{X}_i \boldsymbol{\beta}$ et $f(y, \boldsymbol{\theta}) = 0$ pour $\alpha + \mathbf{X}_i \boldsymbol{\beta} \leq y < \alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0$. Comme f est monotone croissante en y pour α et $\boldsymbol{\beta}$ donnée, le domaine où f est supérieure à 1 est connecté (*i.e.*, le sous-ensemble E de Θ , $E \in \Theta$, est tel que pour $[\sigma^2 \quad \alpha \quad \boldsymbol{\beta}] \in E$, nous avons $f > 1$, est un ensemble connecté).

Définissons l'ensemble $N_\rho(\boldsymbol{\theta})$ de telle sorte que $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \rho$. Alors $N_\rho(\boldsymbol{\theta})$ ou $\Theta \cap N_\rho(\boldsymbol{\theta})$ (si $N_\rho(\boldsymbol{\theta})$ n'est pas entièrement compris dans Θ) sont un et l'autre compact, et puisque f est continue à droite en $\boldsymbol{\theta}$, f atteint son maximum dans cet ensemble. Si nous notons ce maximum e^M , alors :

$$\sup_{\boldsymbol{\theta} \in N_\rho(\boldsymbol{\theta})} f = \max_{\boldsymbol{\theta} \in N_\rho(\boldsymbol{\theta})} f = e^M, \quad (4-55)$$

et ce pour tout y , c'est-à-dire $\sup_{\boldsymbol{\theta} \in N_\rho(\boldsymbol{\theta})} f \leq e^M \quad \forall y \in (-\infty, \alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0)$. On remarque que comme dans le cas de la loi exponentielle, on aura $f(y, \boldsymbol{\theta}) > 1$ si σ^2 est suffisamment petit. Pour un tel σ^2 , nous avons alors :

$$\begin{aligned} \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln f^*(y, \boldsymbol{\theta} | \rho) dF(y, \boldsymbol{\theta}_0) &= \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln f^*(y, \boldsymbol{\theta} | \rho) f(y, \boldsymbol{\theta}_0) dy \\ &= \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln[f(y, \boldsymbol{\theta} | \rho)] I(f > 1) f(y, \boldsymbol{\theta}_0) dy \\ &\leq \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln e^M f(y, \boldsymbol{\theta}_0) dy \\ &= \ln e^M \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} f(y, \boldsymbol{\theta}_0) dy \\ &= M \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} f(y, \boldsymbol{\theta}_0) dy \\ &= M \end{aligned} \quad (4-56)$$

Ceci complète la démonstration de la première partie du lemme.

Nous devons maintenant démontrer que l'intégrale $\int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln \varphi^*(y, r) dF(y, \boldsymbol{\theta}_0)$ est finie. Comme pour la première partie, nous pouvons remplacer $dF(y, \boldsymbol{\theta})$ par $f(y, \boldsymbol{\theta}) dy$. Si $f(y, \boldsymbol{\theta}) < 1$ pour tout $\boldsymbol{\theta}$ sur le domaine de y , alors $\ln \varphi^* = 0$ et l'intégrale est finie. Supposons alors qu'il existe y tel que $f(y, \boldsymbol{\theta}) > 1$ pour $\|\boldsymbol{\theta}\| > r$. On note alors que la fonction $\sqrt{2/\pi\sigma^2} \exp\{-(y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 / 2\sigma^2\}$ s'approche de son maximum (supremum) pour $y - \alpha - \mathbf{X}_i \boldsymbol{\beta} \rightarrow 0$ et $\sigma^2 \rightarrow \underline{\sigma}^2 > 0$. Puisque la valeur de σ^2 est inférieurement bornée par $\underline{\sigma}^2$ et que $\exp\{-(y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 / 2\sigma^2\} \leq 1$ et égale à 1 pour $y = \alpha + \mathbf{X}_i \boldsymbol{\beta}$ alors $f(y, \boldsymbol{\theta})$ est bornée pour tout $\boldsymbol{\theta} \in \Theta$. De plus, puisque $f(y, \boldsymbol{\theta})$ est monotone croissante en y , nous atteignons éventuellement un point où $f(y, \boldsymbol{\theta}) \leq 1$ avec $\|\boldsymbol{\theta}\| = r$. En fait, nous n'avons qu'à poser $\sigma^2 = \underline{\sigma}^2$ et ensuite choisir α et $\boldsymbol{\beta}$ de telle sorte que $\underline{y} = \underline{\alpha} + \mathbf{X}_i \underline{\boldsymbol{\beta}}$ avec $\|\underline{\boldsymbol{\beta}}\|^2 = \|[\underline{\alpha} \quad \boldsymbol{\beta}^T]\|^2 \geq r^2 - \|\underline{\sigma}^2\|^2$. Donc, il existe un \underline{y} , tel que $\forall y < \underline{y}$, nous avons $f(y, \boldsymbol{\theta}) < 1$, d'où :

$$\begin{aligned}
\int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln \varphi^*(y, r) dF(y, \boldsymbol{\theta}_0) &= \int_{\underline{y}}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln \varphi^*(y, r) dF(y, \boldsymbol{\theta}_0) \\
&\leq \int_{\underline{y}}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} \ln e^M dF(y, \boldsymbol{\theta}_0) \\
&\leq M \int_{-\infty}^{\alpha_0 + \mathbf{X}_i \boldsymbol{\beta}_0} dF(y, \boldsymbol{\theta}_0) \\
&= M \\
&< \infty.
\end{aligned} \tag{4-57}$$

CQFD

Notons que le fait de poser $\underline{\sigma}^2 > 0$ nous donne une variable aléatoire non-dégénérée. Si $\underline{\sigma}^2 = 0$ la variable aléatoire y dégénère en un point et par conséquent il n'y a plus de problème statistique. Cette borne est donc nécessaire tant du point de vue

statistique qu'économique. Elle implique entre autres qu'il doit y avoir de l'inefficacité.

Lemme 4.10. *Si la suite $\{\theta_n\}$ converge, $\lim_{n \rightarrow \infty} \theta_n = \theta$, alors $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ pour tout y sauf possiblement pour un ensemble de mesure zéro par rapport à $F(y, \theta_0)$ qui peut dépendre de θ mais pas de la suite.*

Démonstration. Puisque θ appartient à un ensemble fermé, toute suite converge dans l'ensemble et nous avons $\lim_{n \rightarrow \infty} \theta_n = \theta$. La fonction f est continue en dans ses paramètres pour tout $\theta \in \Theta$, sauf pour les α et β tels que $y = \alpha + X\beta$. Dans ce cas nous avons $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ pour $y \leq \alpha_n + X\beta_n$ et $\lim_{n \rightarrow \infty} f(y, \theta_n) = 0$ pour $y > \alpha_n + X\beta_n$. Puisque l'ensemble les α et β tels que $y = \alpha + X\beta$ forment un ensemble de mesure zéro dans Θ , alors toute suite convergente, $\lim_{n \rightarrow \infty} \theta_n = \theta$, implique que $\lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ sauf pour un ensemble de mesure zéro qui ne dépend pas de la suite.

CQFD

On remarque que $\lim_{n \rightarrow \infty} \theta_n = \theta, \lim_{n \rightarrow \infty} f(y, \theta_n) = f(y, \theta)$ est une définition de la continuité, donc ce lemme ne nous donne pas beaucoup d'information pour une loi semi-normale.

Lemme 4.11. *Si θ_1 est un vecteur de paramètres différent de θ_0 , c'est-à-dire $\theta_1 \neq \theta_0$, alors $F(y, \theta_1) \neq F(y, \theta_0)$ pour au moins une valeur de y .*

Démonstration. Par construction, pour deux valeurs différentes du vecteur de paramètres la densité exponentielle aura un nombre fini de points pour lesquels

$f(y, \theta_1) = f(y, \theta_0)$. Par conséquent, sauf en un nombre fini de y , nous avons :

$$\begin{aligned} f(y, \theta_1) &= \sqrt{\frac{2}{\pi\sigma_1^2}} \exp\left\{-\frac{1}{2\sigma_1^2}(y - \alpha_1 - \mathbf{X}_i\boldsymbol{\beta}_1)^2\right\} \\ &\neq \sqrt{\frac{2}{\pi\sigma_0^2}} \exp\left\{-\frac{1}{2\sigma_0^2}(y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2\right\} \\ &= f(y, \theta_0). \end{aligned} \quad (4-58)$$

La démonstration procède par construction en montrant que si $F(\tilde{y}, \theta_1) = F(\tilde{y}, \theta_0)$ en un point \tilde{y} , alors il existe au moins un autre point tel que $F(y, \theta_1) \neq F(y, \theta_0)$.

Considérons maintenant la valeur de F pour $\theta_1, \theta_0 \in \Theta$ tels que $\theta_1 \neq \theta_0$ au point \tilde{y} et supposons que :

$$F(\tilde{y}, \theta_1) = \int_{\tilde{y}}^{\alpha_1 + \mathbf{X}_i\boldsymbol{\beta}_1} \ln f(s, \theta_1) ds = \int_{\tilde{y}}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \ln f(s, \theta_0) ds = F(\tilde{y}, \theta_0). \quad (4-59)$$

Considérons alors $y = \tilde{y} - \Delta y$ pour Δy arbitrairement petit et $y \neq \tilde{y}$, nous avons alors :

$$\begin{aligned} F(y, \theta_1) &= \int_y^{\alpha_1 + \mathbf{X}_i\boldsymbol{\beta}_1} f(s, \theta_1) ds \\ &= \int_{\tilde{y}}^{\alpha_1 + \mathbf{X}_i\boldsymbol{\beta}_1} f(s, \theta_1) ds + \int_y^{\tilde{y}} f(s, \theta_1) ds \\ &\approx F(\tilde{y}, \theta_1) + f(y, \theta_1)\Delta y \end{aligned}$$

et

$$\begin{aligned} F(y, \theta_0) &= \int_y^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} f(s, \theta_0) ds \\ &= \int_{\tilde{y}}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} f(s, \theta_0) ds + \int_y^{\tilde{y}} f(s, \theta_0) ds . \\ &\approx F(\tilde{y}, \theta_0) + f(y, \theta_0)\Delta y. \end{aligned}$$

Alors,

$$F(y, \theta_1) - F(y, \theta_0) \cong (f(y, \theta_1) - f(y, \theta_0)) \Delta y. \quad (4-60)$$

Et puisque $f(y, \theta_1) - f(y, \theta_0) \neq 0$ sauf en un point (auquel cas on prendra alors $\tilde{y} \neq y$), nous avons : $F(y, \theta_1) \neq F(y, \theta_0)$.

CQFD

Lemme 4.12. Si $\lim_{n \rightarrow \infty} \|\theta_n\| = \infty$, alors $\lim_{n \rightarrow \infty} f(y, \theta_n) = 0$ pour tout y .

Démonstration. Nous avons $f(y, \theta) = \sqrt{2/\pi\sigma^2} \exp\{-(y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 / 2\sigma^2\}$ et puisque $\exp\{-(y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 / 2\sigma^2\}$ est borné supérieurement pour toute combinaison admissible de $[\sigma^2 \quad \alpha \quad \boldsymbol{\beta}^T]$. Nous avons deux cas à considérer. Si

$\sigma_n / \|\alpha \quad \boldsymbol{\beta}^T\| \rightarrow \infty$, alors nous avons

$\lim_{\|\sigma_n^2 \quad \hat{\boldsymbol{\beta}}_n^T\| \rightarrow \infty} \sqrt{2/\pi\sigma_n^2} \exp\{-(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)^2 / 2\sigma_n^2\} \rightarrow 0 \quad \forall y$, puisque $\sqrt{2/\pi\sigma_n^2} \rightarrow 0$ et

$\exp\{-(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)^2 / 2\sigma_n^2\} \rightarrow 1$. Maintenant, si $\sigma_n / \|\alpha \quad \boldsymbol{\beta}^T\| \rightarrow 0$, alors

$\exp\{-(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)^2 / 2\sigma_n^2\} \rightarrow 0$ et nous avons

$\lim_{\|\sigma_n^2 \quad \hat{\boldsymbol{\beta}}_n^T\| \rightarrow \infty} \sqrt{2/\pi\sigma_n^2} \exp\{-(y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)^2 / 2\sigma_n^2\} \rightarrow 0 \quad \forall y$. Par conséquent, nous

devons avoir :

$$\lim_{\|\sigma_n^2 \quad \hat{\boldsymbol{\beta}}_n^T\| \rightarrow \infty} \sqrt{\frac{2}{\pi\sigma_n^2}} \exp\left\{-\frac{1}{2\sigma_n^2} (y - \alpha_n - \mathbf{X}_i \boldsymbol{\beta}_n)^2\right\} \rightarrow 0 \quad \forall y. \quad (4-61)$$

CQFD

Lemme 4.13. Soit θ_0 le vecteur de paramètres vrais, alors

$$\int_{-\infty}^{\infty} |\ln f(y, \theta_0)| dF(y, \theta_0) \leq M < \infty.$$

Démonstration. Nous avons $dF(y, \theta_0) = f(y, \theta_0) dy = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right\} dy$.

Maintenant, par définition, $y \leq \alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0$, d'où :

$$\begin{aligned} & \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} |\ln f(y, \theta_0)| dF(y, \theta_0) \\ &= \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \left| \ln \sqrt{\frac{2}{\pi}} - \ln \sigma^2 - \frac{1}{2\sigma^2}(y - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2 \right| \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right\} dy \quad (4-62) \\ &\leq \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \left| \ln \sqrt{\frac{2}{\pi}} \right| dF(y, \theta_0) + \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} |\ln \sigma^2| dF(y, \theta_0) + \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \frac{1}{2} \left(\frac{y - \alpha - \mathbf{X}_i\boldsymbol{\beta}}{\sigma} \right)^2 dF(y, \theta_0). \end{aligned}$$

Les deux premiers termes sont clairement bornés. Maintenant, notons que :

$$\lim_{y \rightarrow -\infty} \frac{1}{2\sigma_0^2} (y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2 \exp\left\{-\frac{1}{2\sigma_0^2}(y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2\right\} \rightarrow 0. \quad (4-63)$$

Alors que le dernier terme de notre intégrale nous donne :

$$\begin{aligned} & \int_{-\infty}^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \left| \frac{1}{2\sigma_0^2} (y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2 \right| dF(y, \theta_0) \\ &= \int_0^{\alpha_0 + \mathbf{X}_i\boldsymbol{\beta}_0} \left| \frac{1}{2\sigma_0^2} (y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2 \right| dF(y, \theta_0) \quad (4-64) \\ &+ \int_{-\infty}^0 \left| \frac{1}{2\sigma_0^2} (y - \alpha_0 - \mathbf{X}_i\boldsymbol{\beta}_0)^2 \right| \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y - \alpha - \mathbf{X}_i\boldsymbol{\beta})^2\right\} dy. \end{aligned}$$

Alors que le premier terme est clairement fini, on montre aisément par intégration que la seconde intégrale est finie.

CQFD

Lemme 4.14. $f(y, \theta | \rho)$ est une fonction mesurable de y pour tout θ et ρ .

Démonstration. On note que puisque Θ est fermé alors le sous-ensemble $N_\rho(\theta)$ défini pour tout θ' tel que $\|\theta - \theta'\| \leq \rho$ est soit un sous-ensemble strict de $N_\rho(\theta) \subseteq \Theta$ ou alors c'est le cas de $N_\rho(\theta) \cap \Theta$, mais, dans les deux cas, le résultat est un ensemble compact.

Puisque $f(y, \theta)$ est continu dans ses paramètres, sauf un θ t.q. $y = \alpha + \mathbf{X}_i \boldsymbol{\beta}$, où il y a une discontinuité du premier ordre, le maximum se confond avec le supremum et est atteint dans l'ensemble, parce que la fonction f est continue à droite (limite à gauche). Par la nature même de f , pour un σ^2 donné, le maximum de cette fonction est atteint pour $y - \alpha - \mathbf{X}_i \boldsymbol{\beta}$ le plus grand possible avec bien entendu $\alpha, \boldsymbol{\beta} \in N_\rho(\theta)$ et $y \leq \alpha + \mathbf{X}_i \boldsymbol{\beta}$. Donc, lorsqu'il existe α et $\boldsymbol{\beta}$ tels que, avec $y - \alpha - \mathbf{X}_i \boldsymbol{\beta} = 0$, la fonction $f(y, \theta)$ est constante en y et ne dépend, continument, que de σ^2 . Pour un tel choix, nous posons $\underline{\sigma}^2 = \min_{\sigma^2 \in N_\rho(\theta)} \sigma^2$ ce qui permet de trouver :

$$f(y, \underline{\theta}) = \sup_{\alpha, \boldsymbol{\beta}, \sigma^2} \sqrt{\frac{2}{\pi \sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2 \right\} = \sqrt{\frac{2}{\pi \underline{\sigma}^2}}. \quad (4-65)$$

Pour tout autre y , (pour tout $\sigma^2, \alpha, \boldsymbol{\beta} \in N_\rho(\theta)$ tel que $y - \alpha - \mathbf{X}_i \boldsymbol{\beta} \neq 0$) on pose $\alpha = \underline{\alpha}$ et $\boldsymbol{\beta} = \underline{\boldsymbol{\beta}}$, la valeur maximale de $y - \underline{\alpha} - \mathbf{X}_i \underline{\boldsymbol{\beta}} < 0$ et ces valeurs existent puisque f est monotone en y . Donc, $f(y, \theta | \rho)$ est soit constante ou décroissante (monotone) et alors les deux cas, continue d'où $f(y, \theta | \rho)$ est une fonction mesurable.

CQFD

Ce résultat est aussi une conséquence du Lemme 4.8 comme dans le cas de la loi exponentielle.

Théorème 4.2. Posons $\theta = [\sigma^2 \quad \alpha \quad \boldsymbol{\beta}^T]$ et définissons l'estimateur du maximum de vraisemblance, $\hat{\theta}_n$, comme l'argument qui maximise :

$$\ell(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \ln L(\sigma^2, \alpha, \boldsymbol{\beta} | y) = \frac{n}{2} \ln \left(\frac{2}{\pi} \right) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \alpha - \mathbf{X}_i \boldsymbol{\beta})^2$$

sous la contrainte $y_i \leq \alpha + X_i\beta$, $i = 1, \dots, n$. Cet estimateur, $\hat{\theta}_n$, est convergent, i.e.

$$\text{plim}_{n \rightarrow \infty} \hat{\theta}_n = \theta_0.$$

Démonstration. La spécification semi-normale nous permet de vérifier l'Hypothèse 1 de Wald (1949) (absolument continue par le Lemme 4.7) et notre Hypothèse 4.2 nous dit que Θ est fermé tout comme l'Hypothèse 7 de Wald (1949). Par les Lemmes 4.8 à 4.13, les six autres hypothèses de Wald sont vérifiées. Par conséquent, le résultat découle du Théorème 2 de Wald (1949).

CQFD

4-2 Inférence statistique du modèle semi-paramétrique

Dans cette section nous développons une méthodologie qui permet de faire de l'induction statistique pour des coefficients du modèle d'Aigner et Chu. Dans la section précédente, nous avons démontré que lorsque les termes d'erreurs suivent une loi de probabilité donnée, l'estimateur du modèle de frontière déterministe devient un estimateur du maximum de vraisemblance. Par contre, la caractérisation statistique des estimateurs d'Aigner et Chu demeure à ce jour incomplet. Nous avons démontré dans la section 4-1 la convergence des estimateurs d'Aigner et Chu dans les cas des programmes linéaire et quadratique. Cependant, il n'est toujours pas possible de déduire facilement la loi de probabilité asymptotique qui conduirait à une procédure d'inférence simple. Greene (1980) a démontré sous l'hypothèse que le terme d'erreur suit une loi gamma que la fonction de vraisemblance se comporte correctement en ce sens qu'elle est compatible avec les conditions de convergence asymptotique par la méthode de Cramér et ainsi que les estimateurs sont asymptotiquement de loi normale. Cependant, les erreurs de loi gamma ne correspondent à aucun modèle de programmation connu.

Comme nous l'avons vu au chapitre 3, le modèle de frontière déterministe n'est pas explicitement fondé sur des hypothèses statistiques. Les propriétés de ces estimateurs

sont donc ambiguës tant et aussi longtemps que nous ne connaissons pas le processus qui génère les données. C'est la logique d'une grande partie de la discussion de Simar et Wilson (1998, 1999) sur la façon d'obtenir les propriétés statistiques des estimateurs DEA. Les estimateurs obtenus par la résolution d'un programme linéaire ont un inconvénient évident : ils ne produisent pas directement d'écart-type pour les coefficients, donc l'inférence statistique semble exclue. Jusqu'à maintenant, le principal inconvénient est qu'en l'absence d'une spécification plus détaillée, la convergence même des estimateurs ne peut pas être vérifiée. Cependant, si on suppose qu'un processus stochastique commun à toutes les observations génère les données, comme on le retrouve dans les modèles non paramétriques (Simar (1999) et Cazals *et al.* (2002)), nous pouvons simuler le processus générateur des données empirique pour déduire les propriétés de l'estimateur. Une telle procédure requiert tout de même la convergence de l'estimateur simulé. Kneip *et al.* (1998) ont démontré cette convergence pour l'estimateur DEA, ce qui donne le fondement à l'approche de simulation proposée par Simar et Wilson (1998).

Il est donc intéressant de tenter d'appliquer cette méthodologie ici à l'estimateur d'Aigner et Chu. Comme dans le cas du DEA, nous devons supposer l'existence d'un processus générateur de données et montrer que l'estimateur est convergent dans un premier temps. C'est ce que nous avons fait dans la section précédente, où nous avons montré que nous disposions d'un estimateur convergent pour le modèle semi-paramétrique. Armé de ce résultat, nous pouvons introduire une procédure de simulation en ré-échantillonnage (de type *bootstrap*) pour obtenir une caractérisation statistique de nos estimations, et surtout, construire un intervalle de confiance pour tous les paramètres estimés.

Cette section est organisée de la façon suivante. La première sous-section exposera la méthodologie générale de simulations que nous avons adoptée pour introduire l'inférence statistique pour les paramètres estimées dans le modèle d'Aigner et Chu. Dans cette section, nous présentons le principe de base guidant les simulations et nous

donnerons les conditions d'applications. La méthode de simulation la plus adaptée à notre problème relève du *bootstrap*; nous présenterons donc certaines ramifications de cette procédure. La deuxième sous-section est consacrée à la méthode utilisée pour construire les intervalles de confiance pour les paramètres du modèle d'Aigner et Chu. La méthode utilisée repose sur le sous-échantillonnage. Toute la méthodologie introduite pour calculer les intervalles de confiance pour les estimateurs d'Aigner et Chu est une contribution originale de cette thèse.

4-2-1 Les méthodes ré-échantillonnages

Les méthodes de ré-échantillonnage sont connues dans la littérature française sous la dénomination de *bootstrap*, un terme qui n'a jamais été vraiment traduit dans la littérature française. L'appellation anglo-saxonne reflète le fait que c'est la méthode de dernière instance à laquelle on fait appel lorsque tout a échoué pour obtenir une solution à un problème (« *to pull oneself up by one's bootstraps* »). Cette appellation fait malheureusement perdre l'essence de la procédure qui consiste à simuler la distribution d'un estimateur à l'aide d'échantillons obtenus par tirage dans l'échantillon original. Depuis les fameux articles de Efron (1979, 1982) et le livre de Efron et Tibshirani (1993), la méthode *bootstrap* a été appliquée dans de nombreuses circonstances et s'est sophistiquée pour s'adapter à des situations de plus en plus complexes, dont les estimations de frontières font partie. Les références classiques pour cette méthode sont (Davison et Hinkley, 1997 ; Hall, 1992 ; Shao et Tu, 1995 ; Zoubir et Iskander, 2004). Le principe général de la méthode est d'extraire de l'échantillon original un grand nombre d'échantillons (avec remise) et de d'appliquer l'estimateur étudié à tous ces pseudo-échantillons pour en déduire les caractéristiques statistiques à partir de la distribution de cet ensemble de valeurs estimées. En d'autres termes, la distribution des valeurs estimées obtenue à partir des pseudo-échantillons devrait approximer la distribution d'échantillonnage d'un estimateur. La difficulté inhérente aux procédures de ré-échantillonnage est d'être à même de pouvoir simuler correctement le Processus Générateur des Données (PGD). Le succès rencontré par

les méthodes de ré-échantillonnage en économétrie provient d'une mise en œuvre relativement simple de la procédure. En particulier, comme nous l'avons mentionné plus haut, l'avantage de cette méthode est de fournir un outil statistique valide lorsque les outils classiques ne permettent pas de caractériser un estimateur.

La construction des pseudo-échantillons est au centre de la méthode. Pour être valide, ces pseudo-échantillons doivent être générés en conformité avec le processus générateur de données. Il existe donc un grand nombre de façons de générer ces échantillons. L'approche très générale que nous retenons pour nos travaux, proposée par Politis et Romano (1992, 1994), permet de réaliser une inférence statistique asymptotiquement valide. Leur approche consiste à extraire des sous-échantillons de l'échantillon original plutôt que d'utiliser la méthode originale suggérée par Efron qui consistait à tirer des échantillons remise de taille identique à l'échantillon original. Cette méthode a la vertu de reposer sur des conditions de cohérence simples et de permettre un plus large éventail d'applications (voir, par exemple, Bickel *et al.* (1997) et Andrews (2000)).

4-2-1-1 Concepts et notations

Nous présentons dans cette section quelques définitions et concepts fondamentaux liés à la méthode que nous mettrons en œuvre. Pour que les méthodes de simulations puissent être mise en œuvre, il faut faire une première hypothèse qui suppose que le mécanisme qui a généré les données peut être représenté mathématiquement ; cette représentation est appelée le processus générateur des données (PGD). Le point de départ est un échantillon issu d'un PGD qui demeure inconnu. Dans notre cas, nous supposons qu'il existe une frontière (fonction de production ou de coût) qui supporte les données générée et cette frontière doit donc être estimée. Par conséquent, les pseudo-échantillons ne peuvent être obtenus à partir du vrai PGD puisque celui-ci est inconnu. Par conséquent, les pseudo-échantillons sont obtenus à partir d'une valeur estimée du PGD spécifié et appliqué à l'échantillon original. Si le PGD estimé est proche du vrai PGD, alors simuler ou générer des données à l'aide de ce PGD devraient permettre

d'obtenir des quantités similaires à celles générées par le vrai PGD. Par conséquent, le processus de ré-échantillonnage agit comme si l'échantillon initial était une population et les pseudo-échantillons sont ainsi obtenus par tirage comme si l'échantillon initial était une « population ». En répétant cette procédure un grand nombre de fois, on peut obtenir un grand nombre d'estimation des caractéristiques de notre PGD postulé. Il est alors possible de rassembler toute cette information pour caractériser le comportement de l'estimateur initial.

Maintenant, la question qui se pose porte sur la façon de conduire l'inférence statistique. La construction d'intervalle de confiance à partir des estimations obtenues à partir des pseudo-échantillons doit être résolue. Pour ce faire nous introduisons quelques notations et concepts. La fonction de probabilité cumulative (FDC) d'une variable aléatoire x est notée F_x . Les paramètres inconnus, $\boldsymbol{\beta} = [\beta_1 \dots \beta_k]^T$, (qui peuvent par exemple être le vecteur de coefficients d'une régression, de dimension $(k \times 1)$), sont exprimés par une fonction de F sous la règle t :

$$\boldsymbol{\beta} = t(F). \quad (4-66)$$

Ceci est une représentation du vrai PGD. Malheureusement, nous ne connaissons pas la distribution exacte des données car F n'est bien entendu pas connue. Nous n'avons qu'un échantillon aléatoire à notre disposition, $\mathbf{X} = (X_1, X_2, \dots, X_n)$, une matrice de dimension $(n \times k)$, générée par F . Cet échantillon est utilisé pour construire la distribution empirique de F , \hat{F} . Nous devons donc estimer $\boldsymbol{\beta}$ à partir de \hat{F} , ou autrement dit, utiliser une autre fonction u , qui lie l'estimateur à l'échantillon observé :

$$\hat{\boldsymbol{\beta}} = t(\hat{F}) = u(\mathbf{X}). \quad (4-67)$$

Un intervalle de confiance pour le vecteur de paramètres $\boldsymbol{\beta}$ est généralement obtenu à partir de l'estimateur $\hat{\boldsymbol{\beta}}$, en tenant compte de la distribution de la probabilité de $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}$

. Si S_α désigne le percentile d'ordre α de la distribution de $\hat{\beta} - \beta$, alors l'intervalle de confiance pour $\hat{\beta}$ a la forme suivante :

$$P(S_{\alpha/2} \leq \hat{\beta} - \beta \leq S_{1-\alpha/2}) = 1 - \alpha. \quad (4-68)$$

Un intervalle de confiance à $(1 - \alpha)$ % est donné par :

$$S_{\alpha/2} \leq \hat{\beta} - \beta \leq S_{1-\alpha/2}. \quad (4-69)$$

La probabilité que la vraie valeur β se trouve dans cet intervalle est égale à $(1 - \alpha)$.

Donc pour obtenir cet intervalle, il est nécessaire de connaître ou être capable d'approximer la fonction de répartition de $\hat{\beta} - \beta$.

Généralement, lorsque nous connaissons la loi de probabilité de $\hat{\beta} - \beta$, nous pouvons éviter le calcul de l'intervalle de confiance. Nous construisons alors un test statistique et on calcule la probabilité marginale de rejet que nous pouvons confronter à un seuil critique, ce qui permet de déterminer si l'hypothèse testée est rejetée ou non. Mais, malheureusement, comme nous l'avons indiqué plus haut, dans le cas qui nous occupe la loi de probabilité du test est inconnue. Habituellement, on s'en remet à la théorie asymptotique pour déduire la loi de probabilité de la statistique du test. Cependant, nous savons depuis la section précédente que notre estimateur ne rencontre pas les conditions pour invoquer un théorème de la limite centrale et par conséquent nous ne pouvons pas déduire une loi pour notre test. Ceci nous amène à utiliser des méthodes de simulation pour obtenir une caractérisation statistique de $\hat{\beta} - \beta$. Dans la sous-section suivante, nous présentons les principes du ré-échantillonnage.

4-2-1-2 Principes du ré-échantillonnage

Les deux plus importants objectifs de l'application statistique sont la détermination d'un estimateur pour un paramètre donné et l'évaluation de sa précision via l'estimation de sa dispersion (écart-type ou variance), ce qui permet la construction d'un intervalle de confiance. Lorsqu'Efron (1979a) a introduit la méthode *bootstrap* il

était motivé par ces deux problèmes. Le plus important était l'estimation de l'écart-type de l'estimateur du paramètre, lorsque les autres méthodes statistiques ne fonctionnent pas (en particulier en matière de précision).

Avant de fournir une définition formelle de la procédure de ré-échantillonnage, nous donnons une explication intuitive de la méthode. Soit un échantillon de taille n , $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$, une matrice de dimension $(n \times k)$, avec $\mathbf{X}_i = [X_{i1} \ \dots \ X_{ik}]$, un vecteur ligne de dimension $(1 \times k)$ pour lequel nous voulons estimer un paramètre (écart-type, intervalle de confiance, etc.). À partir de l'échantillon nous pouvons construire la fonction de probabilité cumulative empirique. Pour ce faire, à chaque valeur de \mathbf{X}_i dans notre échantillon, nous attribuons une probabilité de tirage de $1/n$. Le principe du *bootstrap* à la Efron est de simplement remplacer la vraie fonction cumulative inconnue de la population par sa contrepartie empirique qui est observée par construction. À partir de cette loi empirique, nous générons un grand nombre de pseudo-échantillons (tirage avec remise à partir de la loi empirique calculée) pour lesquelles nous pouvons à nouveau calculer les paramètres d'intérêt. Les propriétés de l'estimateur sont alors déterminées à partir de la loi empirique des paramètres générés à partir des estimations sur les pseudo-échantillons.

Maintenant, considérons cette méthode plus en détail. Supposons que nous ayons un échantillon de n vecteurs aléatoires de loi identique et indépendante (*i.i.d.*), $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ et un estimateur du vecteur de paramètres de dimension $(k \times 1)$: $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$. Une procédure pour évaluer la précision de $\hat{\boldsymbol{\beta}}$ est définie à partir de la fonction de répartition empirique F_n de la façon suivante. La fonction de répartition empirique attribue une probabilité de $1/n$ à chaque élément observé du vecteur aléatoire, \mathbf{X}_i pour $i = 1, 2, \dots, n$. La fonction de répartition empirique est un estimateur du maximum de vraisemblance de la vraie fonction et cet estimateur est

convergent. L'approximation *bootstrap* de la répartition de $\hat{\beta} - \beta$ est donnée par la répartition empirique de $\hat{\beta}^b - \hat{\beta}$ où $\hat{\beta}^b$ est obtenu par un ré-échantillonnage indépendant avec remise à partir de la fonction de probabilité empirique F_n . Le résultat essentiel ici est que la loi de $\hat{\beta} - \beta$ dans l'univers réel se comporte comme la loi de $\hat{\beta}^b - \hat{\beta}$ dans l'univers des *bootstrap*. En répliquant de nombreuses fois $\hat{\beta}^b - \hat{\beta}$ nous obtenons alors un ensemble de pseudo-observations dont la densité (ou fonction cumulative) permet d'approcher la vraie loi non observée. L'application de cette méthode nécessite habituellement la génération de très nombreux pseudo-échantillons. En d'autres termes, nous devons ré-échantillonner notre base de donnée B fois (selon la nature du problème ce nombre peut aller de quelques centaines à plusieurs milliers, voire plus). Somme toute, l'algorithme demeure assez simple :

1. Générer un échantillon avec remplacement à partir de la loi de répartition empirique.
2. Obtenir $\hat{\beta}^b$, l'estimateur du vecteur de paramètres β en utilisant le pseudo-échantillon.
3. Répéter les étapes 1 et 2 B fois pour obtenir B valeurs estimées de l'estimateur $\hat{\beta}^b, b = 1, \dots, B$.
4. Utiliser les B estimateurs pour construire l'écart-type, l'intervalle de confiance, etc.

En répétant B fois les étapes 1 et 2 nous obtenons une approximation Monte-Carlo de la répartition de $\hat{\beta}^b$, on notera alors que l'écart-type de $\hat{\beta}^b$ est une estimation Monte-Carlo de l'écart-type de $\hat{\beta}$.

Maintenant, rappelons que nous sommes intéressé tout particulièrement par le comportement de $\hat{\beta} - \beta$ mais ce que nous pouvons obtenir par la simulation est une approximation de la répartition de $\hat{\beta}^b - \hat{\beta}$. Comme nous l'avons mentionné plus haut,

à la base du *bootstrap* réside l'idée que, pour un N suffisamment grand, nous nous attendons à ce que ces deux distributions deviennent à peu près identiques lorsque B tend vers l'infini. Puisque les pseudo-échantillons sont pour l'essentiel tous différents les uns des autres, puisque certaines observations sont omises et que d'autres apparaissent plusieurs fois, la valeur de $\hat{\beta}^b$ varie d'une itération à l'autre. Il est important de se rappeler que le concept essentiel de la procédure *bootstrap* est que la variation de $\hat{\beta}^b$ (basée sur la répartition empirique F_n) autour de $\hat{\beta}$ est similaire à la variation de $\hat{\beta}$ autour du vrai paramètre β , cette dernière variation représentant la répartition vraie dans population. Il y a de bonnes raisons de croire que ce principe est vérifié. Lorsque la taille de l'échantillon est assez grande, la répartition empirique F_n va se rapprocher de la vraie répartition de la population, $F_n \sim F$. Par conséquent, les pseudo-échantillons générés par F_N seront proches des échantillons aléatoires obtenus à partir de F . En fait, le théorème de Glivenko et Cantelli (Chung, 1974) affirme que la distribution empirique converge en probabilité vers la distribution réelle uniformément lorsque les observations sont générées en échantillonnage. C'est ce résultat qui donne toute sa crédibilité à l'approche de ré-échantillonnage. Par conséquent, les deux ingrédients fondamentaux de la procédure *bootstrap* sont un estimateur convergent de β et une modélisation convergente du PGD.

4-2-1-3 Sous-échantillonnage

Le problème de mise en œuvre que nous rencontrons lorsque nous désirons utiliser le *bootstrap* classique dans les problèmes de frontière est la contrainte imposée sur l'observabilité des données imposée par cette frontière. En fait, les observations sont bornées par des paramètres que nous cherchons, donc il prend en considération que des observations qui s'ajouteraient pourraient se trouver de l'autre côté de la frontière calculée initialement. En d'autres termes, si nous nous limitons aux données observées pour conduire notre inférence, nous risquons de biaiser nos résultats en ne

donnant pas la possibilité à la frontière de se situer au-delà de la frontière calculée. Dans le cas du DEA, la première procédure qui a été suggérée est le *bootstrap* lisse, les données sont réfléchies de chaque côté de la frontière estimée, ce qui permet de créer ce bruit sur la frontière. Dans le cas qui nous occupe, cette procédure est extrêmement complexe, voire impossible à mettre en œuvre. Une façon de contourner le problème est d'injecter le bruit en travaillant à l'intérieur de l'ensemble de données. En prenant des sous-échantillons de l'ensemble de données initial, nous pouvons générer des frontières qui ont la capacité de se trouver de part et d'autres de la frontière initiale lorsque nous « déplaçons » le bruit généré pour le coller à la frontière estimée initialement. Cette section a pour but de présenter une alternative au ré-échantillonnage *bootstrap*, c'est-à-dire le sous-échantillonnage.

Lorsque le *bootstrap* non paramétrique a été introduit, Efron a proposé de prendre la même taille des pseudo-échantillons que l'échantillon initial. Cela semblait raisonnable et fonctionne très bien pour de nombreuses applications. Cependant, il était clair que le choix de la taille des pseudo-échantillons pouvait très bien être de plus petites tailles que l'échantillon initial. Il est donc permis d'avoir $m < n$ (où m est la taille du pseudo-échantillon et n la taille de l'échantillon initial). La taille des pseudo-échantillons doit cependant satisfaire deux conditions, soit que $m \rightarrow \infty$ et $m/n \rightarrow 0$ lorsque $n \rightarrow \infty$ pour qu'il y ait convergence asymptotique de la procédure. Contrairement au *bootstrap* classique qui exige que l'on tire les pseudo-échantillons avec remise, cette règle n'est pas nécessaire pour le sous-échantillonnage. Par conséquent, si le tirage est avec remise, nous parlons de *m-bootstrap* alors que si le tirage est sans remise, nous avons le sous-échantillonnage. Ces méthodes d'échantillonnage, avec ou sans remise, ont été examinées par Bickel et Freedman (1981), Bretagnolle (1983), Swanepoel (1986), Beran et Ducharme (1991), et Politis et Romano (1994, 1999).

La technique du sous-échantillonnage remonte à Hartigan (1969). Ce dernier a développé une méthode pour calculer des intervalles de confiance en utilisant le sous-

échantillonnage aléatoire en appuyant sa procédure sur le théorème de la valeur typique. Ce théorème dit que la vraie valeur de β a une probabilité de $1/B$ d'être dans un intervalle, noté I_i , défini de la façon suivante : $I_1 = [-\infty, a_1), I_2 = [a_1, a_2), \dots, I_{B-1} = [a_{B-2}, a_{B-1})$ et $I_B = [a_{B-1}, \infty)$, pour $i = 1, \dots, B$. Les intervalles sont définis tel que a_1 est le plus petit $\hat{\beta}^b$ obtenu à partir d'un sous-échantillon, a_2 est le second plus petit $\hat{\beta}^b$ obtenu à partir d'un autre pseudo-échantillon, ainsi de suite. Ce théorème fournit la probabilité que chaque intervalle contienne β dans une procédure de sous-échantillonnage aléatoire. En d'autres termes, nous pouvons construire une région de confiance de $100(j/B)\%$ pour $1 \leq j \leq B-1$, simplement en combinant tous les j intervalles. Ce résultat est à la base de la méthode de calcul des intervalles de confiance par le percentile proposée par Efron.

Politis, Romano et Wolf (1999) résument les résultats concernant le sous-échantillonnage et les comparent au *bootstrap* classique. Ils démontrent que les conditions de convergence de la méthode de sous-échantillonnage sont plus faibles que celle du *bootstrap*. En fait, la méthode de sous-échantillonnage fonctionne toujours lorsque les conditions d'application du *bootstrap* sont rencontrées, la précision de sous-échantillonnage pourrait cependant être plus faible que celle du *bootstrap*. Le plus grand avantage du sous-échantillonnage est que les conditions de convergence sont plus faibles que pour le *bootstrap*, ce qui implique que le sous-échantillonnage peut être utilisé sans danger dans certaines situations où les conditions de régularité sont violées pour la méthode *bootstrap*. En effet, la méthode de sous-échantillonnage fournit un cadre d'application suffisamment étendu pour couvrir diverses méthodes d'estimation et de modélisation. Comme nous l'avons mentionné au début de cette section, le *bootstrap* naïf ne fonctionne pas dans notre cas. Par contre, suivant Simar et Wilson (2011), une méthode de sous-échantillonnage

peut nous donner une inférence convergente même si la frontière dépend des paramètres estimés.

4-2-2 La correction du biais

Bien que l'estimateur d'Aigner et Chu soit convergent, il est aussi biaisé en échantillon fini. Rappelons que le biais d'un estimateur est donné par $Biais[\hat{\beta}_i] = E(\hat{\beta}_i) - \beta$. La correction pour le biais peut être calculée par analogie à l'aide de l'estimateur obtenu par ré-échantillonnage, $\hat{\beta}_i^b$, $b = 1, \dots, B$ (où B est le nombre d'échantillons tirés, de la façon suivante :

$$\widehat{biais}_B[\hat{\beta}_i] = B^{-1} \sum_{b=1}^B \hat{\beta}_{bi}^* - \hat{\beta}_i, \quad (4-70)$$

et l'estimateur corrigé pour le biais est :

$$\hat{\beta}_i = \hat{\beta}_i - Bias[\hat{\beta}_i] = 2\hat{\beta}_i - B^{-1} \sum_{b=1}^B \hat{\beta}_{bi}^*. \quad (4-71)$$

4-2-3 Estimation des intervalles de confiance

Dans cette sous-section, nous nous concentrons sur l'estimation par intervalle de confiance des paramètres par la méthode de simulation. Tout d'abord, nous présentons brièvement l'estimation par intervalle de confiance pour ensuite passer à leur estimation en utilisant le ré-échantillonnage. Finalement, nous exposons la méthode utilisée pour construire les intervalles de confiance pour les paramètres du modèle semi-paramétrique.

4-2-3-1 La définition de l'intervalle de confiance

À partir des pseudo-échantillons il est possible de construire des intervalles de confiance pour les paramètres estimés. Hall (1992a), Efron et Tibshirani (1993), proposent plusieurs méthodes pour construire des intervalles de confiance par ré-échantillonnage.

Supposons que nous ayons un vecteur de paramètres β appartenant à un espace euclidien de dimension n , $\beta \in \mathbb{R}^n$. Un ensemble de confiance de taille $1 - \alpha$ est un sous-ensemble de \mathbb{R}^n avec la propriété que si un nombre infini d'échantillons était généré dans les mêmes conditions et que le paramètre β était estimé pour chaque échantillon, alors il se retrouverait dans l'ensemble de confiance $100 * (1 - \alpha) \%$ du temps. Dans le cas le plus simple où le paramètre est de dimension un, un ensemble de confiance est alors appelé un intervalle de confiance.

4-2-3-2 La méthode basée sur l'écart-type

Il existe plusieurs façons de construire un intervalle de confiance. Une première possibilité est de construire un intervalle de confiance en utilisant l'écart-type de la façon suivante :

$$\left[\hat{\beta} - \mu_{(1-\alpha/2)} \hat{\sigma}_{\hat{\beta}}, \hat{\beta} + \mu_{(1-\alpha/2)} \hat{\sigma}_{\hat{\beta}} \right], \quad (4-72)$$

avec $\mu_{1-\alpha/2}$ correspondant au percentile $1 - \alpha / 2$ de la loi de répartition normale réduite et $1 - \alpha$ correspond le niveau de confiance. Cette méthode repose sur trois hypothèses principales, soit que l'échantillon est approximativement normal, que l'estimateur est non biaisé et que $\hat{\sigma}_{\hat{\beta}}$ est une bonne estimation de l'écart-type du paramètre estimé. Cette méthode n'est pas possible dans notre cas puisque nous ne connaissons pas la loi de répartition des paramètres, mais nous savons qu'elle n'est certainement pas normale en petit échantillon.

4-2-3-3 La méthode basée sur le percentile simple

La méthode utilisant le percentile est la plus évidente pour construire un intervalle de confiance pour un paramètre ré-échantillonné. Supposons que les pseudo-échantillons sont de taille n et notons $\hat{\beta}_b$ la valeur estimée du paramètre à la $b^{ième}$ itération de la procédure. Pour obtenir l'intervalle de confiance nous commençons par ordonner en

ordre croissant les B valeurs estimées $\hat{\beta}_b$ puis on élimine les $(\alpha/2)\%$ des $\hat{\beta}_b$ aux deux extrémités de la distribution. La première valeur restante et la dernière, $\hat{\beta}_l^b$ et $\hat{\beta}_s^b$ forment les bornes inférieure et supérieure de l'intervalle de confiance à $(1-\alpha)\%$ pour β . C'est ce qu'on appelle la méthode percentile à la Efron pour générer un intervalle de confiance à l'aide de la méthode de ré-échantillonnage (*bootstrap*). Ce serait un intervalle exact comme un intervalle construit aux estimations de sous-échantillons aléatoires.

Cette méthode tient pour des échantillons *bootstrap* alors que nous utilisons le sous-échantillonnage aléatoire. La différence fondamentale repose sur le fait que les pseudo-échantillons *bootstrap* sont tous de taille n identique à celle de l'échantillon original et généré par un tirage avec remise alors que pour le sous-échantillonnage nous avons des pseudo-échantillons de taille inférieure à n et générée par tirage sans remise. Cependant, lorsque la taille des échantillons augmente, la différence entre la répartition des estimations *bootstraps* et par sous-échantillonnages devient arbitrairement petite. Par conséquent, nous nous attendons à ce que les deux méthodes de ré-échantillonnage génèrent des intervalles de confiance plus ou moins identiques. Pour résumer, pour un ensemble de pseudo-valeurs obtenu par ré-échantillonnage, $\{\hat{\beta}^b, b = 1, 2, \dots, B\}$, l'intervalle de confiance de niveau $1 - \alpha$ pour le vrai paramètre β est donné par :

$$\left[\hat{\beta}_{(\alpha/2)}^*, \hat{\beta}_{(1-(\alpha/2))}^* \right]. \quad (4-73)$$

C'est la méthode de calcul que nous utilisons dans ce chapitre.

4-2-3-4 La méthode de calcul des intervalles de confiance

Comme nous l'avons mentionné plus haut, le concept essentiel soutenant l'approche empirique pour calculer les intervalles de confiance est que la loi de répartition inconnue de $\hat{\beta}(X) - \beta(X)$ pourrait être approchée par celle de $\hat{\beta}^b(X) - \hat{\beta}(X)$. Si

nous connaissons la loi de répartition de $\hat{\beta}(\mathbf{X}) - \beta(\mathbf{X})$, il serait trivial de trouver des valeurs, a_α et b_α telles que :

$$\Pr(-b_\alpha \leq \hat{\beta}(\mathbf{X}) - \beta(\mathbf{X}) \leq a_\alpha) = 1 - \alpha, \quad (4-74)$$

avec $1 - \alpha$ correspondant au niveau de confiance que nous voulons attribuer à cet intervalle. Les trois seuils classiques que l'on reporte en général sont pour des valeurs de $1 - \alpha$ égales à 90 %, 95 % et 99 %. Puisque nous ne connaissons pas la forme que prend la loi de probabilité régissant l'intervalle, nous utilisons les valeurs simulées pour trouver les seuils a_α^* et b_α^* de la façon suivante. Tout d'abord, ordonnons par ordre croissant les $\hat{\beta}_b^*(\mathbf{X}) - \hat{\beta}(\mathbf{X})$ pour $b = 1, 2, \dots, B$. Par la suite, nous éliminons $(\alpha/2) \times 100\%$ des éléments aux deux extrémités de la liste tirée. Les valeurs de $-a_\alpha^*$ et $-b_\alpha^*$ correspondent aux extrémités inférieure et supérieure respectivement de la série tronquée. Lorsque B tend vers l'infini, l'équation (4-73) est approchée avec probabilité un par l'intervalle de confiance empirique générée par la procédure *bootstrap* que nous venons de décrire. Sachant que F et F_n sont respectivement la vraie fonction de répartition et la fonction de répartition empirique, nous avons :

$$\left[\hat{\beta}(\mathbf{X}) - \beta(\mathbf{X}) \right] | F \sim \left[\hat{\beta}^*(\mathbf{X}) - \hat{\beta}(\mathbf{X}) \right] | F_n. \quad (4-75)$$

nous obtenons une estimation de l'équation (4-74) en remplaçant a_α et b_α par a_α^* et b_α^* , $\left[\hat{\beta}(\mathbf{X}) - \beta(\mathbf{X}) \right]$ par $\left[\hat{\beta}^*(\mathbf{X}) - \hat{\beta}(\mathbf{X}) \right]$, ce qui donne :

$$\text{Prob}(-b_\alpha^* \leq \hat{\beta}(\mathbf{X}) - \beta(\mathbf{X}) \leq a_\alpha^* | \mathbf{X}) \approx 1 - \alpha. \quad (4-76)$$

Notre intervalle de confiance de niveau $1 - \alpha$ est ainsi défini par :

$$\hat{\beta}(\mathbf{X}) + a_\alpha^* \leq \beta(\mathbf{X}) \leq \hat{\beta}(\mathbf{X}) + b_\alpha^*. \quad (4-77)$$

4-2-4 Algorithme pour calculer l'inférence statistique

Dans cette section, nous présentons un algorithme de la méthode simulation en sous-échantillonnage qui nous permettra d'obtenir les intervalles de confiance pour les huit modèles estimés par la méthode d'Aigner et Chu.

La taille des sous-échantillons est une fraction de la taille de l'échantillon initial, soit $m = \phi n$ avec $\phi \in [0,1]$ et m la taille des sous-échantillons, n la taille de l'échantillon initial et ϕ est la proportion de l'échantillon initial que nous retenons pour le sous-échantillonnage. Nous proposons maintenant un algorithme pour la procédure de sous-échantillonnage pour obtenir des intervalles de confiance valides.

Etape 1 : Estimation des paramètres ⁸ $\hat{\beta}_i$ du modèle à partir l'échantillon d'observations originales S_n .

Etape 2 : Génération d'un sous-échantillon $S_m^b = \{X_i\}_{i=1}^m$ par tirage au sort (de manière indépendante, uniforme et sans remise) de m observations à partir de l'échantillon original S_n .

Etape 3 : Estimation des paramètres $\hat{\beta}^b$ pour le modèle de l'étape 1 en utilisant le sous-échantillon S_m^b .

Etape 4 : Répéter les étapes 2 et 3 B fois, $b = 1, \dots, B$ pour obtenir B valeurs estimées $\left\{ \left(\hat{\beta}_i^b \right) \text{ pour } b = 1, 2, \dots, B \right\}_{i=1}^I$.

Etape 5 : À partir de $\hat{\beta}_i$ et $\hat{\beta}_i^b$, la distribution empirique de $\hat{\beta} - \beta$ est déduite. Les valeurs de a_α^* et b_α^* sont obtenues à partir de la méthode présentée plus haut.

⁸ Pour simplifier l'écriture, ici $\hat{\beta}_i$, $i = 1, \dots, k$, représente les paramètres pour tous des modèles du chapitre 4.

Etape 6 : Calcul de $\left[\hat{\beta}_i(X_i) + a_\alpha^*, \hat{\beta}_i(X_i) + b_\alpha^* \right]$, un intervalle de confiance de taille $1 - \alpha$ pour $\hat{\beta}_i$.

Conclusion

Dans ce chapitre, nous avons montré que les estimateurs du modèle d'Aigner et Chu sont bien des estimateurs du maximum de vraisemblance. Plus important, en utilisant la méthode de Wald (1949), nous pouvons contourner la défaillance de la méthode classique pour montrer la convergence de l'estimateur. Finalement, nous avons introduit une méthode de simulation, le sous-échantillonnage, pour calculer des intervalles de confiance pour les paramètres estimés.

Chapitre 5. L'indicateur d'infrastructure

Le processus de production dans l'industrie du transport de personnes utilise, d'une part, des facteurs de production sous le contrôle de l'entreprise comme le travail, les cars et autobus ou encore l'essence et, d'autre part, un facteur de production implicite dans les analyses et hors du contrôle des entreprises, soit le réseau routier. Dans le contexte qui nous occupe ici, nous voulons mesurer l'impact de ce dernier sur le fonctionnement des entreprises et en particulier nous voulons comprendre de quelle façon il affecte l'efficacité et la productivité des entreprises de cette industrie. Par conséquent, nous devons obtenir une mesure de ce facteur de production pour pouvoir l'introduire comme input dans le processus de production.

La question qui se pose alors est de trouver une façon de mesurer cet input pour les différentes entreprises de l'industrie. Il est clair que pour des entreprises localisées à des endroits différents, l'accès et la disponibilité du réseau sont potentiellement différents. Par conséquent, nous avons besoin d'une mesure d'infrastructures spécifique à chacune des entreprises. Il faut aussi que notre mesure capte la disponibilité du réseau à partir du point où l'entreprise se trouve par rapport aux parcours et routes qui devront être empruntés par les véhicules de l'entreprise. Donc, il y a une considération locale et une considération nationale. C'est-à-dire que nous devons mesurer les infrastructures disponibles autour de l'entreprise et celles qui correspondent aux différentes destinations où elle devra transporter ses passagers et clients. Par conséquent, il est clair qu'une mesure des infrastructures est une réflexion du service offert par le réseau, ce qu'il est maintenant habituel d'appeler l'accessibilité.

Dans la littérature, la plupart des études ont pris le stock du capital public comme approximation pour les infrastructures. Dans ce contexte, il n'y a pas de distinction entre les différents secteurs et le stock d'infrastructures est identique pour toutes les entreprises. Puisque nous nous concentrons sur les entreprises du transport de

voyageurs, pour avoir une bonne mesure des infrastructures, il faut capter l'influence des infrastructures routières sur le processus de production. Pour ce faire, notre idée est de mesurer les infrastructures en nous basant sur un indicateur d'accessibilité et de l'utiliser comme un input quasi-fixe dans nos modèles. À notre connaissance, il n'y a que Martin, Gutiérrez et Roman (2004) qui utilisent un indicateur d'accessibilité dans une application de DEA. Par contre ils ont construit l'indicateur d'accessibilité simple et l'utiliser comme un output dans leur mesure.

Dans ce chapitre, nous présentons la méthode de calcul de l'indicateur d'infrastructures afin de permettre l'inclusion des infrastructures dans notre modèle empirique. L'organisation est la suivante : premièrement, nous expliquerons pourquoi l'accessibilité peut être utilisée comme mesure des infrastructures. Sur cette base, nous présentons la méthode retenue pour construire une mesure des infrastructures à partir de l'accessibilité. Finalement, nous exposerons la méthode spécifique utilisée pour obtenir une mesure quantitative des infrastructures spécifiques à chaque entreprise.

5-1 Une mesure de la disponibilité des infrastructures

Le réseau routier est effectivement un input utilisé par les entreprises de transport. Cependant, celui-ci offre une disponibilité relative et c'est cette disponibilité qui constitue le réel input à utiliser dans le processus de production. Par conséquent, nous ne pouvons pas directement utiliser une mesure toute simple, par exemple le nombre de kilomètres de route, comme input. D'une part, la nature de la route est un aspect qu'il faut considérer et inclure. La vitesse permise et la largeur de la route modifient la nature même de l'infrastructure. Par conséquent, il faut trouver le moyen de combiner la longueur et le volume de route. La capacité du système routier sera aussi pertinent, la congestion jouant un rôle non négligeable. Bien entendu, la localisation dans ce réseau joue aussi un rôle, la capacité de transport et la congestion varient en général d'un point à un autre et offre donc une nuance sur le volume disponible d'infrastructures pour l'entreprise. En d'autres termes, pour bien capter les

infrastructures utilisées par les entreprises de transport il faut générer une mesure de la disponibilité du réseau routier pour chacune des entreprises. La mesure qui est généralement utilisée en économie du transport pour agréger toute cette information est un indicateur d'accessibilité. Dans cette section, nous présentons les raisons pour lesquelles nous pouvons l'utiliser comme représentation des infrastructures à l'échelle de l'entreprise. Ceci nous amènera à développer une mesure spécifique des infrastructures pour chaque entreprise.

La notion d'accessibilité est très variable. Cependant, on s'entend sur le fait que le terme accessibilité exprime la facilité avec laquelle les activités peuvent être atteintes à partir d'une localisation donnée en utilisant un système de transport (Morris *et al.*, 1978), de même que les opportunités offertes aux particuliers et aux entreprises pour atteindre les lieux de leurs activités (Linneker et Spence, 1992) et la simplicité avec laquelle les activités de la société peuvent être atteintes. En fait, l'accessibilité est le produit principal d'un système de transport. Il détermine l'avantage de la localisation d'une région par rapport à toutes les autres régions. Du point de vue théorique, il existe une relation bien connue entre les infrastructures de transport et l'accessibilité. L'amélioration des infrastructures produit une augmentation de l'accessibilité et, par conséquent, influence la productivité des entreprises de transport. Dans une perspective régionale, les nouveaux investissements en infrastructures de transport peuvent accroître l'accessibilité régionale et, par conséquent, fournir des conditions qui favorisent les entreprises locales. Bien entendu, dans le cas du transport il y aura une externalité positive pour les entreprises situées hors de la zone en question. De nouvelles entreprises sont également créées dans la région, car il existe de nouvelles opportunités économiques (Banister et Berechman, 2000). Par conséquent, l'accessibilité peut être considérée comme la représentation d'un ensemble d'effets des infrastructures de transport.

5-2 Les fondements de l'accessibilité.

Depuis longtemps, l'accessibilité a été utilisée par les politiciens et les planificateurs comme un objectif de planification, sans avoir pour objectif de l'utiliser comme facteur de production dans un processus productif. Cette inclusion demeure cependant notre objectif et nous devons comprendre de quelle façon s'articulent les mesures d'accessibilité pour pouvoir les adapter à notre objectif. Par conséquent, ce dont nous avons besoin est une traduction du concept d'accessibilité en une mesure de sorte qu'elle puisse être utilisée pour évaluer le service offert par les infrastructures. Ici, nous allons présenter quelques méthodes existantes dans la littérature qui sont à la base de la méthode utilisée pour calculer notre indicateur d'infrastructures.

5-2-1 Les indicateurs simples

La mesure basée sur les indicateurs simples, comme le nombre de voyages, d'embouteillages ou la vitesse moyenne sur le réseau routier joue un rôle important dans les études sur les politiques de transport. Geurts et van Eck (2001) ont fait une synthèse des études portant sur l'accessibilité en utilisant les indicateurs simples dans l'Union Européenne. Par exemple, le plan de transport en Angleterre en 2010 est évalué en utilisant les embouteillages et le nombre d'heures d'embouteillage évitées comme indicateurs d'accessibilité ; aux Pays-Bas, on utilise souvent la vitesse de déplacements, le nombre de train en retard, *etc.* comme les indicateurs d'accessibilité. Le problème pour ce type de mesures est qu'il ne prend pas en considération la destination, qui peut être loin de la zone d'intérêt et l'activité. Dans notre cas, ce type de mesure est clairement insuffisant en particulier parce qu'il ne prend en considération qu'une toute petite fraction des infrastructures physiques et bien entendu ignore les destinations.

5-2-2 La mesure basée sur l'activité

Les mesures d'accessibilité basées sur l'activité se déclinent en différentes sous-méthodes : mesure de distance, mesure de contour, mesure de potentielle, mesure de type spatio-temporelle géographique. Regardons-les tour à tour.

La mesure de distance

La mesure de distance la plus simple est « l'accessibilité relative » qui ne tient compte que de la distance entre l'origine et la destination. L'accessibilité relative entre deux points i et j est donnée par la distance entre deux points d_{ij} . Il est parfois possible d'utiliser le temps ou le coût de déplacement de i à j . L'accessibilité intégrale est alors définie comme le degré d'interconnexion pour un point i avec tous les autres points :

$$A_i = \sum_j d_{ij} . \quad (5-1)$$

Ces mesures sont fondées sur l'hypothèse que la distance explique à elle seule l'accessibilité. Des limitations sont bien entendu que cette mesure ne prend pas vraiment en considération le nombre de destinations potentielles ni l'effet de la congestion, par exemple.

La mesure de contour

Cette mesure est souvent utilisée dans le planning urbain et les études géographiques. La mesure indique le nombre d'opportunités accessibles dans un temps ou une distance donné. Cette mesure indique que l'accessibilité va augmenter s'il y a plus d'opportunités mises à disposition; une diminution du temps de voyage, une augmentation des destinations, etc. représentent cette augmentation des opportunités. L'inconvénient de ce type de mesure est que la pondération est la même pour toutes les destinations potentielles. On suppose donc implicitement que toutes les opportunités ont un potentiel identique, sans tenir compte de la distance ou du temps pour y accéder. Il n'est pas clair que l'intérêt d'une entreprise de transport inter-

régional pour les trajets longues distances et donc l'accessibilité à des régions éloignées est le même que celui d'une entreprise de transport intra régional.

La mesure de potentiel

Hansen (1959) fut le premier à utiliser le concept de potentiel pour décrire l'accessibilité en termes d'opportunités. Il définit l'accessibilité comme « le potentiel d'opportunités pour l'interaction ». Cette mesure est obtenue en pondérant les opportunités des destinations par une mesure d'influence négative du temps de transport. Mais aucun consensus ne se dégage pour sa formulation quantitative qui reste très variable. Ingram (1971) a joué un rôle central en donnant à l'accessibilité une forme opérationnelle : il sépare l'accessibilité relative (entre deux points) et l'accessibilité intégrale (d'un point par rapport à l'ensemble des autres points). Même si les formulations théoriques sont différentes, la mesure prend la forme analytique suivante :

$$A_i = \sum_j E_j f(d_{ij}), \quad (5-2)$$

où E_j qui mesure les opportunités de destinations (postes vacants d'une zone d'emploi, nombre d'hôpitaux ou même, le niveau de bien-être en zone j), A_i représente l'accessibilité pour la zone i pour tous les « attracteurs » (la possibilité de trouver un emploi, la possibilité de recevoir un traitement médical, *etc.*) dans la zone j , d_{ij} est la distance ou le temps de déplacement entre la zone i et j ; et f représente une fonction de résistance, qui est décroissante en d_{ij} . La fonction de résistance indique le degré d'accessibilité d'une opportunité située à la distance d_{ij} de l'utilisateur, ou encore la facilité que les agents de l'origine ressentent pour accéder à cette opportunité. Dans notre formulation de l'accessibilité, cette fonction servira à mesurer l'influence de la distance et du temps de congestion sur le service que les infrastructures peuvent proposer aux entreprises de transport. Clairement, plus il est difficile pour un autobus ou un autocar de se déplacer entre l'origine et la destination

(plus longue est la distance ou plus important est le temps de congestion), moins les entreprises peuvent transporter de voyageurs en utilisant les infrastructures routières, d'où un service des infrastructures disponible en plus petite quantité. Si nous poursuivons cette idée, cette fonction étant décroissante dans son argument, elle symbolise une résistance à l'accessibilité.

L'application de cette mesure est assez répandue. Les économistes utilisent cet indicateur pour mesurer l'accessibilité aux emplois (Linneker et Spence, 1992), aux populations (Patton et Clark, 1970), à l'éducation (Pacione, 1989). Puisque nous étudions le déplacement des véhicules des entreprises de transport, la facilité avec laquelle ces entreprises peuvent opérer pour accéder aux destinations devient un enjeu dans notre recherche. En pratique, on a donné à la fonction de résistance des formes diverses. On retrouve souvent une forme de type $d_{ij}^{-\alpha}$, où α est un paramètre pour capter l'importance de la distance sur l'accessibilité. On rencontre aussi dans la littérature la fonction exponentielle négative, la fonction Gaussienne ou encore la fonction logistique, etc.

Ces trois formes d'indicateurs d'accessibilité captent toutes des caractéristiques importantes des infrastructures, mais ont toutes des limitations pour nos applications. En fait, il y a toujours une dimension importante qui est omise. Dans la suite, nous présenterons un indicateur qui rassemble ces caractéristiques importantes des infrastructures, que nous voulons bien entendu capter.

5-3 Une mesure des infrastructures

Notre objectif ici est de construire un indicateur d'infrastructures qui reflète le service que tire l'entreprise de transport des infrastructures routières. Le service à capter est comme nous l'avons expliqué plus haut, l'accessibilité aux différentes destinations potentielles des entreprises de transport de personne. Si ce qui importe comme input pour les entreprises est le service offert par les infrastructures et que ce service est de l'accessibilité, notre objectif est de construire un indice d'accessibilité pour

représenter l'input infrastructures dans le processus de production de l'entreprise. Ensuite, nous pourrions utiliser cet indicateur pour représenter les infrastructures et il pourra ainsi agir comme input quasi-fixe utilisé dans le processus de production de l'output transport. L'indicateur que nous élaborons dans cette section enrichit les indicateurs proposés plus haut, entre autres en prenant en considération le potentiel de destination (qui est bien entendu important pour l'entreprise de transport).

5-3-1 La forme de l'indicateur

Voyons de quelle façon construire l'indicateur. Un indicateur d'infrastructures sera la combinaison de la quantité d'infrastructures dans la zone considérée et du potentiel d'accès aux différentes interconnexions entre les régions. Il faut cependant prendre en considération certains points importants pour chacune des composantes. Premièrement, bien qu'il faille inclure la quantité d'infrastructures, comme par exemple la longueur des routes, cette quantité de route doit être rapportée à la taille de la région considérée, car pour une quantité d'infrastructures données, plus la taille de la zone est grande, moins les entreprises obtiennent de service de ces infrastructures. Ceci implique donc un ajustement de notre première composante de l'indicateur. Ensuite, lorsque nous mesurons le potentiel ou l'accessibilité potentielle, il faut prendre en considération que l'éloignement joue un rôle. En fait, plus la distance entre deux zones d'emploi est grande moins il est facile pour une entreprise de transporter des voyageurs entre ces deux endroits, donc moins la valeur de l'indicateur devra être élevée. C'est la même chose pour le temps de congestion, plus grand est le temps de congestion entre l'origine et la destination, moins les entreprises peuvent utiliser le service fourni par les infrastructures, et donc plus petite est la valeur prise par l'indicateur. Pour cette raison, la mesure de potentiel suggère de prendre une fonction de résistance qui peut refléter l'influence de la distance et le temps de congestion sur la facilité qu'ont les entreprises pour transporter des voyageurs. Nous supposons que la fonction de résistance f est de forme exponentielle.

Si nous assemblons ces deux composantes, nous obtenons un indicateur qui prendra en considération les caractéristiques des infrastructures de l'origine, R_{i1} , et une fonction de la distance entre l'origine i et la destination j incluant le temps passé en déplacement, et la quantité des infrastructures de la destination, R_{i2} :

$$Inf_i = R_{i1}R_{i2}, \quad (5-3)$$

où Inf_i est l'indicateur d'infrastructures, R_{i1} est une mesure locale (à l'origine) de la capacité de transport et R_{i2} de l'accessibilité aux infrastructures de transport.

Les deux termes composant Inf_i sont développés de la façon suivante. Tout d'abord l'indicateur du volume d'infrastructures :

$$R_{i1} = L_i V t_i / ST_i, \quad (5-4)$$

où L_i représente la longueur totale de route, $V t_i$ la vitesse moyenne et ST_i est la surface totale de la zone d'emploi i . Donc, la quantité d'infrastructures est bien rapportée à la taille géographique de la région. Par exemple une grande région avec la même quantité de route qu'une petite région est considérée avoir moins d'infrastructures. Plus d'autoroutes veut aussi dire plus de vitesse et ceci influe positivement sur le volume d'infrastructures.

La deuxième composante capte le potentiel, c'est à dire les relations avec les autres zones que l'on ajuste pour inclure le temps de transport et la distance. Elle est définie comme suit :

$$R_{i2} = \sum_j L_j \exp\{-\nu d_{ij} - \gamma t_{ij}^{con}\}, \quad (5-5)$$

où L_j représente la longueur totale des routes dans la zone d'emploi j , ν représente le « déflateur » de l'indicateur d'accessibilité par rapport à la distance et d_{ij} est la distance entre la zone d'emploi i et la zone d'emploi j , γ représente le « déflateur » de l'indicateur d'accessibilité par rapport au temps perdu à cause de congestion et t_{ij}^{con}

est le temps moyen perdu à cause de la congestion entre les zones d'emploi i et j , avec $i \neq j$. Bien entendu, ν et γ sont deux coefficients à estimer. Nous discuterons de l'estimation de ces paramètres plus loin. Il est clair que plus la distance est grande et plus le temps perdu à cause de la congestion est important, plus l'indicateur doit être petit. Ceci correspond bien à moins de services fournis par les infrastructures.

Notre indicateur d'accessibilité pour les firmes qui se trouvent dans la zone i est donc égal à :

$$Inf_i = \frac{L_i V t_i}{ST_i} \sum_j L_j \exp\{-\nu d_{ij} - \gamma t_{ij}^{con}\}. \quad (5-6).$$

Pour calculer l'indicateur, il nous faut une matrice de distance (elle donne les distances entre les zones i et j), la longueur totale des routes, la vitesse moyenne, la surface totale d'une zone d'emploi, le temps de congestion et les deux déflateurs. Puisque nous pouvons obtenir toutes les variables sauf les deux déflateurs à partir de notre base de données, il ne nous reste que deux déflateurs (ν pour la distance et γ pour le temps) à estimer.

5-3-2 Déflateur à la distance et déflateur au temps

Dans cette sous-section, nous présenterons la méthode classique pour estimer le déflateur pour la distance ν et le déflateur pour le temps γ . Ils sont habituellement estimés à l'aide d'un modèle gravitaire des échanges.

5-3-2-1 Modèle gravitaire simple et ses limites

Le modèle gravitaire établit une relation d'interaction entre deux zones géographiques sur la base de leur masse respective et de la distance qui les sépare. Un simple modèle gravitaire peut être représentés par la formule suivante :

$$F_{ij} = A M_i M_j d_{ij}^{-\alpha}, \quad (5-7)$$

où F_{ij} représente le flux de marchandise entre les zones d'emploi i et j , A est une constante, M_i et M_j représentent respectivement les masses des zones d'emploi i et j

(elles peuvent représenter la population des zone d'emploi, leur nombre d'emploi, le nombre total de déplacements émis par l'origine ou le nombre total de déplacements reçus par la destination, etc.), d_{ij} représente la distance entre les zones d'emploi i et j et α est un paramètre de sensibilité à la distance, positif, qui vient atténuer l'effet d'une augmentation de la distance. Rappelons que $(-\alpha)$ est l'élasticité distance du flux de marchandises :

$$\frac{\partial \log(F_{ij})}{\partial \log(d_{ij})} = -\alpha. \quad (5-8)$$

On parle alors de modèle gravitaire, car le flux entre deux zones i et j est proportionnel au produit des masses M_i et M_j et à une fonction décroissante de la distance d_{ij} .

Ce modèle est généralement utilisé pour réaliser des prévisions de flux de marchandises entre une origine et sa destination, ce qui n'est clairement pas notre objectif. Bien que le principe qui nous intéresse soit là, ce modèle n'est pas tout à fait adapté à notre objectif pour la mesure des deux freins. Pour nous, il est important de pouvoir mesurer l'apport de la distance et du temps de congestion à la structuration des échanges entre les zones d'emplois. Une autre limite des modèles non contraints, tels que celui présenté plus haut est que "*ces derniers accordent une place trop importante au rôle des capacités d'émission et réception*" (Dujardin, 2001) qui ne font pas l'objet de cette étude. Par conséquent, nous nous tournons maintenant vers une évolution de ce modèle qui palie à ces problèmes. Ce sont les modèles à double contrainte.

5-3-2-2 Modèle gravitaire double contrainte et son ajustement

Pour estimer les paramètres dont nous avons besoin, nous voulons retenir l'idée du modèle gravitaire simple, mais le corriger pour ses défauts les plus apparents. La solution proposée est un développement de ce modèle, soit le modèle gravitaire à double contrainte. Dans ce type de modèle, les seules variables introduites sont la

distance et le temps de congestions entre les zones d'emplois. L'émission provenant des origines et l'attractivité des destinations sont quant à elles fixées et constituent des contraintes de la modélisation. Le modèle de double contrainte permet dès lors d'étudier l'effet de la séparation spatiale, toutes choses égales par ailleurs, quant aux capacités d'émission et de réception des lieux. Il est donc possible de résumer l'effet de la distance et du temps à l'aide d'un « frein » et de mesurer l'apport spécifique de la distance et du temps de congestion au pouvoir explicatif du modèle. La formulation du modèle est la suivante :

$$F_{ij} = a_i \cdot O_i \cdot b_j \cdot D_j \cdot C_{ij}, \quad (5-9)$$

où F_{ij} est le flux de marchandise entre i et j par le modèle à double contrainte, O_i est l'émission de flux total observé de la zone d'emploi d'origine i , soit $O_i = \sum_j F_{ij}$, D_j est la réception de flux total observé de la zone d'emploi de destination j , soit $D_j = \sum_i F_{ij}$. $a_i = 1 / \sum_j b_j D_j C_{ij}$ et $b_j = 1 / \sum_i a_i O_i C_{ij}$ sont les coefficients d'ajustement associés respectivement à O_i et D_j , $C_{ij} = d_{ij}^{-\nu} (t_{ij}^{con})^{-\gamma}$ est la fonction de frein, qui reflète l'influence de la distance et du temps de congestion sur le flux de marchandises, avec d_{ij} la distance entre l'origine et la destination, et t_{ij}^{con} le temps de congestion entre l'origine et la destination. Bien entendu ν et γ sont les coefficients à estimer, souvent nommés « frein de la distance » et « frein du temps ».

Nous pouvons bien sûr estimer les paramètres ν et γ par les modèles présentés ci-haut en forme logarithmique. Par contre, un ajustement du modèle log-linéaire pose plusieurs problèmes. Premièrement, l'utilisation de la transformation logarithmique affecte la nature des estimations. Ce modèle va sous-estimer les flux importants et les flux totaux (Haworth et Vincent, 1979). Le deuxième problème découle de l'utilisation de la transformation logarithmique lorsque certains flux sont nuls. Le logarithme de zéro ne peut pas être calculé et entrer dans le modèle. Donc un petit nombre positif (par exemple, 0,5) est généralement ajouté à toutes les observations.

Quand le nombre de zéros est assez important, le choix de ce nombre peut avoir un impact considérable sur les coefficients du modèle.

En considérant tous les problèmes exprimés ci-dessus, nous ne pouvons pas vraiment estimer directement ce modèle, cependant nous voulons en garder les principes car nous croyons que c'est la version la plus adaptée pour calculer les paramètres dont nous avons besoin.

Une alternative pour estimer ce modèle est de passer par une régression de Poisson. En fait, c'est la méthode généralement utilisée pour estimer l'influence des variables environnementales (la distance, le temps de congestion, etc.) sur le flux de marchandise dans un modèle gravitaire. Cette méthode est largement éprouvée (D'Aubigny *et al.*, 2000). Elle offre de nombreux avantages statistiques. Elle permet notamment de prendre en considération les flux nuls, tout en conservant les contraintes sur les origines et les destinations du modèle à double contrainte. De plus, elle accorde une place plus mesurée aux petits flux dont le rôle est souvent exagéré par les méthodes d'ajustement log-linéaires, alors même qu'ils sont entachés d'une forte incertitude. Le plus important, par la maximisation de la log vraisemblance, une régression de Poisson peut garantir que les contraintes de production et d'attraction sont respectées pour toutes les zones d'emplois.

Supposons que la probabilité que le flux de marchandise entre les zones d'emploi i et j soit égal à F_{ij} soit donnée par :

$$\Pr[F_{ij}] = \frac{\lambda_{ij}^{F_{ij}}}{F_{ij}!} e^{-\lambda_{ij}}, \quad (5-10)$$

où F_{ij} est le flux de marchandises entre les zones d'emplois i et j et λ_{ij} est l'espérance mathématique de F_{ij} .

Alors, la log-vraisemblance d'une série de flux observés où chaque flux est le résultat d'un processus de Poisson particulier est donnée par l'équation suivante :

$$L^* = \sum_i \sum_j \ln \left(\exp(-\lambda_{ij}) \lambda_{ij}^{F_{ij}} / F_{ij}! \right) = \sum_i \sum_j \left[F_{ij} \ln \lambda_{ij} - \lambda_{ij} - \ln(F_{ij}!) \right]. \quad (5-11)$$

Maximiser cette fonction est équivalent à maximiser le numérateur puisque le dénominateur est une constante. Par conséquent, nous maximisons :

$$Z = \sum_i \sum_j \left[F_{ij} \ln \lambda_{ij} - \lambda_{ij} \right] \quad (5-12)$$

par choix de λ_{ij} . Avant de nous tourner vers la spécification des λ_{ij} , qui sont bien entendu la moyenne du processus de Poisson pour les relations entre i et j , nous devons expliciter certaines conditions de cohérence interne au modèle. Pour que la maximisation de Z soit assurée, les conditions de premier ordre implique que le système d'équations suivant doit être satisfait :

$$\frac{\partial Z}{\partial \lambda_{ij}} = \frac{F_{ij}}{\lambda_{ij}} - 1 = 0. \quad (5-13)$$

En prenant la somme sur les j et les i respectivement, nous obtenons :

$$\begin{aligned} \sum_j F_{ij} &= \sum_j \lambda_{ij} \quad \forall i \\ \sum_i F_{ij} &= \sum_i \lambda_{ij} \quad \forall j. \end{aligned} \quad (5-14)$$

Puisque par définition $O_i = \sum_j F_{ij}$ et $D_j = \sum_i F_{ij}$, la première série de conditions garantit que $O_i = \sum_i \lambda_{ij} \quad \forall i$ alors que la seconde série garantit que $D_j = \sum_j \lambda_{ij} \quad \forall j$.

Maintenant, tournons-nous vers la spécification de l'espérance mathématique de F_{ij} .

Nous supposons que λ_{ij} est de la forme suivante :

$$\lambda_{ij} = g \left(a_i O_i, b_j D_j, d_{ij}, t_{ij}^{con} \right), \quad (5-15)$$

où d_{ij} est la distance entre les zones d'emploi i et j , t_{ij}^{con} est le temps perdu à cause de congestion entre les zones d'emploi i et j et g une fonction à valeurs réelles.

Pour rendre opérationnelle cette fonction nous devons spécifier explicitement g . Il est habituel de retenir la forme exponentielle. Nous obtenons alors :

$$g(a_i O_i, b_j D_j, d_{ij}, t_{ij}^{con}) = \exp(\ln a_i O_i + \ln b_j D_j - \nu \ln d_{ij} - \gamma \ln t_{ij}^{con}), \quad (5-16)$$

où $\ln a_i O_i$ et $\ln b_j D_j$ sont des variables d'origine et de destination respectivement. En fait, ces deux variables reflètent l'influence de taille de la zone d'emploi (au sens de la capacité à recevoir et à envoyer des marchandises) sur le flux entre les deux zones. Elles peuvent aussi capter l'influence de toutes les variables explicatives qui ne sont pas mesurables précisément de manière exogène. En fait, lors de l'estimation, $\ln a_i O_i$ et $\ln b_j D_j$ sont spécifiés comme des variables dichotomiques pour représenter respectivement l'origine i et la destination j .

Nous avons le modèle suivant à estimer :

$$\lambda_{ij} = \exp(\beta_i I_i + \beta_j J_j - \nu \ln d_{ij} - \gamma \ln t_{ij}^{con}), \quad (5-17)$$

où I_i et J_j sont des variables dichotomiques pour capter l'émission et la réception de marchandises. En maximisant le log vraisemblance, nous pouvons obtenir les paramètres de frein de distance et de frein de temps. Une fois ces deux paramètres obtenus, nous pouvons les utiliser dans l'équation (5-6) pour calculer notre indicateur d'infrastructure.

Conclusion

Notre contribution méthodologique se trouve principalement dans cette partie. Premièrement, en se basant sur le modèle Aigner et Chu (1968), nous développons le modèle semi-paramétrique pour la fonction de coût, cela peut compléter l'aspect de recherche quand nous étudions la performance d'une entreprise ou d'un secteur. Deuxièmement, nous avons démontré la convergence des estimateurs obtenus et fourni une façon de construire l'inférence statistique. Avec ces démonstrations, nous avons complété la partie statistique du modèle semi-paramétrique, et donné une méthode pertinente pour calculer l'intervalle de confiance des paramètres estimés, donc nous pouvons maintenant tester la significativité du modèle de Aigner et Chu. Finalement, nous avons construit un indicateur d'accessibilité. L'utilisation de cet indicateur pour estimer l'influence des infrastructures au niveau microéconomique est une nouveauté et nous permet tester quantitativement l'influence des infrastructures sur la performance des entreprises avec un modèle semi-paramétrique.

Pour conclure, dans cette partie, nous avons présenté le modèle non paramétrique et le modèle semi-paramétrique pour mesurer la performance des entreprises. Nous avons également démontré les propriétés statistiques du modèle semi-paramétrique et introduit une façon de mesurer les infrastructures routières. Le contenu de cette partie nous fournit la base théorique pour lancer une application de cette méthode au secteur du transport routier de voyageurs dans la partie suivante.

Troisième partie :

**Impact des infrastructures sur la
productivité des entreprises de transport
routier de voyageurs en France**

Introduction

L'expansion urbaine, un phénomène d'agglomération, met en difficultés les efforts de planification des autorités. Les gens ont besoin de prendre un moyen de transport adéquat leur permettant de se rendre à leur travail, de commercer, d'aller à l'école, aux centres culturels ou à l'hôpital. Ainsi, un système de transports publics doit assurer la mobilité et l'accessibilité par un moyen rapide, sûr et régulier à un coût raisonnable. Malheureusement, il n'est pas facile d'assurer l'ensemble de ces caractéristiques. Cela découle de la complexité des arrangements institutionnels entre l'État et plusieurs municipalités. Ainsi, une première étape consiste à travailler à un accord entre toutes les institutions politiques concernées. En particulier, se posent des questions telles que la division du pouvoir entre les autorités, la coordination administrative, le financement, la sélection et le fonctionnement pour exploiter les services fournis par plusieurs intervenants (bus, autocars, camionnettes, etc.). Du côté empirique, notre objectif est de mesurer les caractéristiques technologiques du secteur du transport de voyageurs en utilisant les modèles théoriques présentés dans les chapitres précédents. Nous pourrions ainsi tester la pertinence des modèles et, nous l'espérons, obtenir des résultats utiles pour les dirigeants politiques.

Une première analyse d'enveloppement des données (DEA) est adoptée pour sélectionner l'efficacité des systèmes et leurs caractéristiques, surtout la nature des rendements d'échelle. Cela nous fournit une base pour faire une comparaison entre les résultats obtenus par le modèle non-paramétrique et le modèle semi-paramétrique.

Le transport routier de voyageurs peut être défini comme une activité qui « englobe les transports individuels (par taxi ou voiture particulière) et collectifs (par autocar et autobus) de voyageurs par route tant à courte qu'à longue distance »⁹. Il est l'une des plus importantes composantes d'une économie. Dans la plupart des pays, le transport

⁹ « Concurrence dans le transport routier » par Darryl Biggar, *Revue de l'OCDE sur le droit et la politique de la concurrence* 2002-4 (Volume 4) page 40.

routier de voyageurs est assuré par les services publics, des entreprises privées ou des entreprises mixtes dans un environnement fortement réglementé. De plus, certaines des composantes les plus importantes du transport sont en général des biens publics. Ceci suppose qu'il y ait des bonnes raisons pour qu'il existe des défaillances de marché dans ce secteur : par exemple le manque de concurrence, des externalités imposées à l'environnement, l'offre de routes comme un bien public, des indivisibilités, etc. L'industrie du transport routier de voyageurs en France assure les déplacements de la population pour des prix inférieurs aux coûts unitaires encourus pour sa fourniture. Par conséquent, les gouvernements locaux servent habituellement à financer les déficits des firmes dans ce secteur. Toutefois, les autorités locales doivent contrôler les dépenses publiques en vertu des règles établies par le droit du transport national (Loi d'orientation sur les transports intérieurs) et le traité européen, ce qui oblige les firmes financées par le public à effectuer leurs tâches dans un environnement où le coût est contrôlé. Ensuite, étant donné que ces entreprises exercent leurs activités exclusivement dans une zone économique, elles conservent un certain pouvoir de monopole, qu'elles ne peuvent pas exploiter en raison de la fixation des prix qu'elles sont tenues de respecter. La firme peut cependant transférer une partie des coûts du capital vers les coûts d'exploitation afin de demander et obtenir une subvention plus élevée ou encore elles peuvent tarder à améliorer l'efficacité ce qui force une augmentation de tarif. Donc, évaluer l'efficacité objective dans ce sous-secteur des transports peut aider à diminuer les asymétries d'information entre les autorités locales et les firmes et aider les décideurs publics à mieux prendre leurs décisions.

Dans cette partie, nous présentons une analyse de l'efficacité des entreprises de transports routiers de voyageurs en France métropolitaine pour la période 2000-2004. Nous introduisons la notion d'indicateur d'accessibilité pour capter l'influence des infrastructures routières sur la performance des entreprises afin de capter l'effet des infrastructures. L'intérêt de cette application empirique repose sur les deux points

suivants : premièrement, dans ce secteur, les firmes utilisent les infrastructures comme input sans avoir de contrôle sur la quantité disponible conjointement avec d'autres facteurs de production pour produire un output, le transport des voyageurs. Donc une analyse de l'efficacité des firmes de ce secteur utilisant les modèles théoriques de la deuxième partie peut nous donner un test adéquat de l'influence des infrastructures sur la performance des firmes. Deuxièmement, du point de vue des politiques économiques, une estimation de l'efficacité des entreprises de ce secteur, très subventionné par l'état, peut offrir l'information nécessaire aux décideurs publics pour les aider dans la planification économique.

Dans un premier temps, nous présentons la base de données puis dans un deuxième temps, une analyse d'enveloppement des données (DEA) pour évaluer l'efficacité de l'industrie et en donner les caractéristiques. La nature des rendements d'échelle des entreprises est analysée et un lien est fait avec la taille de celles-ci. Finalement, le modèle semi-paramétrique en fonction de production et en fonction de coût est utilisé pour calculer l'efficacité des DMU et caractériser leur technologie. L'accent est mis sur l'impact des infrastructures sur la performance des firmes. L'analyse de nos résultats permettra de mettre en évidence les caractéristiques technologiques du secteur de transport des voyageurs. Ces résultats sont bien entendu appuyés sur les propriétés statistiques des estimateurs du modèle semi-paramétrique.

Chapitre 6. Les données

Ce chapitre se concentre sur l'organisation des données. Nous définissons les inputs et les outputs, de même que la construction de l'indicateur d'infrastructures.

6-1 Source des données

Sans aucun doute, la qualité des données détermine la qualité des résultats, donc il faut d'abord chercher les données les plus fiables et les plus précises possibles. Nos données proviennent de deux sources. Pour les données concernant les entreprises, elles proviennent de *L'enquête annuelle d'entreprises (EAE) dans le secteur des transports*. Cette enquête est la seule enquête nationale sur le sujet. Elle est réalisée en métropole (il existe une enquête spécifique pour les entreprises qui se trouvent dans des départements d'outre-mer (DOM)) dans tous les secteurs d'activité par l'Institut National de la Statistique et des Études Économiques (INSEE) en partenariat avec les services statistiques de tous les ministères. Elle décrit la structure des entreprises (siège) et concerne, pour le transport, les entreprises dont l'activité principale (nomenclature d'activités française (NAF)) relève des transports terrestres, des transports par eau, des transports aériens (hors transports spatiaux) et des services auxiliaires des transports (hors entreposage frigorifique et agences de voyage). L'enquête est effectuée en $t+1$ pour l'année t et comporte un questionnaire général pour les entreprises de six salariés et plus et un questionnaire simplifié pour les petites entreprises. Les questionnaires sont différenciés selon les secteurs. Sont recueillies des données communes à toutes les enquêtes (compte de résultats, effectifs, investissements), des données communes à toutes les entreprises du secteur transport (personnel sédentaire, roulant ou navigant, sous-traitance, etc.) et des données propres à certaines activités de transport (transport routier de marchandise, transport routier de voyageurs, taxis, transport maritime, fluvial, remontées mécaniques)¹⁰. Le transport

¹⁰ La source de présentation : l'INSEE, « Enquête annuelle d'entreprise de transport (EAE) ».

routier de marchandises est pris en compte dans le secteur « Transports routiers de marchandises », qui se décompose lui-même en « Transports routiers de marchandises de proximité » (602L), « Transports routiers de marchandises interurbains » (602M). Le transport routier de voyageurs est inclus dans le secteur « transports urbains et routiers de voyageurs », qui se décompose lui-même en « transports urbains de voyageurs » (602A), « transports routiers réguliers de voyageurs » (602B), « transports de voyageurs par taxi » (602E) et « autres transports routiers de voyageurs » (602G). Les entreprises sont classées selon leur activité principale. Les données recueillies, détaillées par taille d'entreprises et par sous-secteur, portent sur l'activité principale exercée par l'entreprise au lancement de l'enquête, soit 602A, 602B, 602E ou 602G. Elles incluent les effectifs non-salariés et salariés (dont salariés roulants) ; le chiffre d'affaires ; les charges (sous-traitance, achats de carburants, loyers et charges locatives, entretiens et réparations, primes d'assurances, frais de personnel) ; les kilomètres parcourus ; le parc routier (le nombre d'autocars, d'autobus, le nombre de camions utilisés), etc.

Le transport routier de voyageurs se décompose en « transports urbains de voyageurs » (602A), « transports routiers réguliers de voyageurs » (602B), « transports de voyageurs par taxi » (602E) et « autres transports routiers de voyageurs » (602G). Les entreprises sont classées selon leur activité principale. Notre application empirique porte sur un seul type de transport de voyageurs, soit le transport routier de voyageurs. Pour le transport de voyageurs, nous éliminons d'abord les entreprises de 602E car les données de taxi sont incomplètes. Comme notre base de données porte sur les infrastructures interurbaines, ce sont les entreprises dans la section 602B qui nous intéressent.

Notre deuxième source concerne les données sur les infrastructures routières. Puisque nous nous concentrons sur les infrastructures routières, les variables concernant les caractéristiques des autoroutes seront exigées. Pour intégrer les infrastructures dans notre analyse, nous avons décidé d'utiliser un indicateur d'accessibilité dans notre

analyse pour capter le service fourni (nous exposerons les détails nécessaires dans la section suivante). Les données dont nous avons besoin sont une carte du réseau français, la longueur du réseau autoroutier, la vitesse moyenne des véhicules, le temps de congestion, la distance entre chaque zone étudiée (la région, le département, la zone d'emploi, les communes, etc.), et les flux de marchandises entre ces zones. Bien entendu, ce genre de données est très difficile à recueillir et à trouver. Nous avons eu la chance d'obtenir l'aide de M. Jean Dominique Blardone du *Commissariat général au développement durable* (CGDD) et MVA Consultancy Limited. Monsieur Blardone travaille dans le domaine du transport pour le gouvernement français avec son modèle MODEV. Les données qu'il nous a fournies, au format Mapinfo, sont les suivantes :

1. La matrice des flux annuels en 2002 entre zones d'emplois. Ces flux sont calculés par le modèle en fonction des emplois par type d'activité et de la population de chaque zone d'emploi.
2. La carte des routes principales à l'horizon 2002.
3. La vitesse moyenne en kilomètre par heure pour les véhicules circulant sur chaque route.
4. Le temps de congestion pour chaque route.

Notre analyse empirique mesure l'impact des infrastructures sur l'efficacité des entreprises dans le secteur de transport routier de voyageurs pour la France, pour la période s'étendant de 2000 à 2004. Nous devons connaître le processus de production des entreprises. Ensuite, quels sont des inputs qu'elles utilisent pour produire quel genre d'output ? Quel est le coût associé ? Comment mesurer la main d'œuvre, le capital, les types d'énergie, les charges d'exploitation ? Comment mesurer les outputs pour des entreprises (chiffres d'affaires ou kilomètres parcourus, ou les deux), *etc.* Tous ces choix dépendent clairement de la disponibilité des données, et nous devons, en plus, vérifier la cohérence des données dans notre base. Pour estimer l'efficacité des DMU, nous avons besoin des quantités d'inputs utilisés et des outputs produits

par les entreprises. En plus, afin de travailler dans l'espace dual avec le modèle d'Aigner et Chu, nous aurons besoin d'information sur le prix des inputs. Donc pour chaque input, nous avons besoin de connaître la dépense totale, la quantité et le prix. Malheureusement ces informations ne sont pas toutes disponibles, il n'y a que la dépense totale dans la plupart du cas. Dans cette circonstance, nous avons recours à d'autres moyens pour déterminer les prix et les quantités. Ainsi, nous utilisons les indices de prix publiés par INSEE et pour certaines variables nous utilisons des indices de Laspeyres afin de procéder à des agrégations.

6-2 Sélection des variables

Tout d'abord, nous présentons les variables d'inputs et d'outputs. Nous utilisons les données comptables de l'enquête « EAE » pour obtenir les quantités d'inputs et d'outputs des entreprises de transport routier pour la période de 2000-2004. Pour chaque année, les douze tableaux les plus importants sont : « Données des entreprises QS et QG »¹¹, « Ventilation des chiffres d'affaires », « Données sur les groupes », « Données sur les établissements », « Données sur le transport fluvial », « Données sur la manutention portuaire », « Données sur le transport maritime », « Données sur les remontées mécaniques », « Données sur les restructurations », « Données sur les taxis », « Données sur le transport routier de marchandises » et « Données sur les transports de personnes ».

Puisque nous nous intéressons aux entreprises œuvrant dans le transport de personnes, les variables d'inputs et d'outputs proviennent des tableaux suivants : « Données des entreprises QS et QG », « Ventilation des chiffres d'affaires », « Données sur le transport routier de marchandises » et « Données sur les transports de personnes »¹².

¹¹ QS : questionnaire simplifié ; QG : questionnaire générale.

¹² En fait, les variables dans les tableaux « Données sur les groupes » et « Données sur les établissements » sont incluses dans le tableau « Données des entreprises QS et QG ».

Le tableau « Données des entreprises QS et QG » contient des variables générales sur toutes les entreprises de transport. Sont disponibles dans ce tableau : le numéro SIREN de l'entreprise enquêtée¹³, les éléments de localisation géographique des entreprises statistiques, les investissements corporels, les achats de carburants, le nombre d'employés, etc. Ce tableau nous donne une partie des inputs. Ensuite le tableau « Ventilation des chiffres d'affaires » nous renseigne sur le chiffre d'affaires des entreprises, qui peut nous servir pour les variables d'outputs. Finalement, le tableau « Données sur les transports de personnes » nous donne les autres variables spécifiques pour chaque secteur : le nombre d'autobus, le nombre d'autocars, les kilomètres parcourus par véhicule, etc.

Tel que mentionné, dans notre modèle, nous aurons besoin des quantités de facteurs de production et d'outputs des entreprises. Pour les facteurs de production, il y aura deux types d'inputs : les inputs variables (ces inputs peuvent être contrôlés parfaitement par les entreprises) et les inputs quasi-fixes (ces inputs ne peuvent pas être contrôlés par des firmes à court terme). Puisque notre modèle empirique repose à la fois sur la fonction de production et la fonction de coût, l'information sur les prix des inputs variables seront également requis.

6-2-1 Inputs variables

Les inputs variables sont des inputs pour lesquels la direction des DMU exerce un contrôle direct et par conséquent elles peuvent ajuster la quantité de ces inputs à court terme. Ces inputs sont la main d'œuvre, L , l'énergie, E , les dépenses de réparation et entretien des véhicules, ENT , et les dépenses en matériels et fournitures, M . Nous définissons $\mathbf{X} = [L \ E \ ENT \ M]$, le vecteur des inputs variables. Pour chacun des

¹³ En fait, c'est le numéro SIRET qui est plus intéressant pour nous, car le numéro SIRET est un numéro d'établissement, qui correspond mieux notre objet d'étude. Mais la variable SIRET est manquante pour la plupart des entreprises. Par contre, nous avons les données correspondent l'adresse de l'unité statistique dans l'enquête, donc SIREN est satisfaisant.

inputs, nous présentons maintenant la source et le détail du calcul menant à l'obtention leur quantité et leur prix.

Main d'œuvre

Quantité de main d'œuvre (L)

Pour une entreprise de transport, idéalement, la main d'œuvres devrait comprendre les salariés cols bleus et cols blancs. Dans le tableau « Données des entreprises QS et QG », nous avons plusieurs variables qui concernent les effectifs des entreprises. Malheureusement, plusieurs variables sont incomplètes ou manquantes¹⁴. En conséquence nous avons choisi une variable agrégée, *TOTEFF*, qui est égale à la somme des effectifs salariés annuel en équivalent temps complet (EF50) et des effectifs non-salariés, pour représenter la quantité de main d'œuvre : $L = TOTEFF$.

Prix de main d'œuvre (w_L)

Pour mesurer le prix de la main d'œuvre, nous aurons besoin d'information sur la masse salariale. Dans notre base de données, nous avons deux variables correspondantes respectivement aux deux types d'effectifs. Pour *EF50*, nous avons la variable *REMUN* (Frais de personnels, qui est la somme des salaires et les charges sociales) ; pour la dépense salariale sur les effectif non-salariés, nous avons la variable *RCH26* (Autres achats et charges externes : Personnel extérieur à l'entreprise). La somme des variables *REMUN* et *RCH26* nous donne la masse salariale (*MS*). Le taux de salaire est égale à la masse salariale divisée par la quantité de travail :

$$w_L = MS/L = (REMUN + RCH26) / TOTEFF .$$

Énergie

Quantité d'énergie (E)

¹⁴ *EF50ROU* : Conducteurs routiers et personnel d'accompagnement, avec 88,12 % de données manquantes. *EF50SED* : personnes sédentaires, qui est égale la somme des personnels dirigeant et cols blancs, avec 89,86 % de données manquantes.

Pour la dépense en énergie, nous avons la variable *ACHACARB* (Achats de carburants). Pour obtenir la quantité d'énergie utilisée, nous la divisons par le prix de carburant: $E = ACHACARB / w_E$.

Prix de l'énergie (w_E)

Pour isoler la quantité d'énergie, nous devons utiliser le prix de l'énergie. Comme ce prix ne peut pas être déduit d'une façon ou d'une autre à partir de la base de données, le prix de l'énergie que nous utilisons est l'indice de prix des carburants donné par l'IPC-Ensemble des ménages - France métropolitaine – par fonction de consommation – Carburants¹⁵.

Entretien et réparation

Quantité d'entretien et réparation (ENT)

Les dépenses en l'entretien et réparation correspondent à la variable *RCH28S2* (Autres achats et charges externes : Entretiens et réparations) dans notre base de donnée. Pour isoler la quantité, nous la divisons par son prix : $R = RCH28S2 / w_{ENT}$.

Prix de l'entretien et réparation (w_{ENT})

Pour l'entretien et réparation, comme celui de carburant, le prix que nous utilisons est l'IPC-Ensemble des ménages-France métropolitaine- Réparation Automobile donné par l'INSEE¹⁶.

Matériels et fournitures

Quantité des matériaux (M)

¹⁵ Il n'existe pas en France d'indice de prix industriel pour le carburant. Par conséquent, nous nous en remettons au prix à la consommation, car c'est l'indice le plus proche de la composition d'énergie utilisée par les entreprises de transport.

¹⁶ La situation est identique au cas de l'énergie. Il n'existe pas en France d'indice de prix industriel pour l'entretien et réparation. Pour la même raison, nous utiliserons le prix à la consommation.

Pour les inputs matériels, les choses sont plus compliquées, car il y a plusieurs catégories de dépenses dans le tableau qui peuvent correspondre aux inputs matériels. Certaines variables sont comprises dans les autres et certaines variables sont croisées avec d'autres.

La dépense en matières premières, les dépenses de restauration et d'hôtellerie, la dépense en fournitures sont également très importantes pour ces entreprises. Il faut donc les extraire de la base de données. Cependant, la dépense totale des charges d'exploitation (*TOTCHEXP*), qui serait la variable à exploiter, est égale à la somme de cinq variables : la consommation intermédiaires (*CI*) ; impôts, taxes et versements assimilés (*IMPOIND*) ; frais de personnels (qui comprend la somme des salaires et les charges sociales) (*REMUN*) ; Provisions, dotations d'exploitation (*RCH14*) et autres charges d'exploitation (*AUTCHARG*) et ne peut être exploitée directement pour les raisons qui suivent. Tout d'abord, nous noterons que *IMPOIND* correspond aux versements obligatoires sans contrepartie aux administrations publiques ou aux institutions de l'Union européenne, frappant la production et l'importation de biens et services, la propriété ou l'utilisation de facteurs de production ; ces impôts sont dus indépendamment de la réalisation de bénéfices ou de la perception de revenus (ce poste inclut notamment la taxe professionnelle). Donc nous ne la considérons pas dans les inputs. Bien entendu *REMUN* entre plutôt dans la catégorie des inputs de main d'œuvre. Ensuite, la variable *RCH14* mesure le coût d'utilisation des biens d'équipement engagés dans la production ainsi que de la perte de valeur subie par ces biens du fait de leur usure et de l'obsolescence (amortissements), enregistrement des réductions de valeur d'éléments d'actifs, liées à des circonstances extérieures telles que les mouvements de prix du marché des éléments patrimoniaux, la défaillance d'un débiteur, etc. (provisions pour dépréciation qui anticipent une perte de valeur d'actif) ; prévision de dépense ou d'engagement futur (provisions pour risques et charges et assimilées, qui anticipent un accroissement de passif). Il est clair que ces items ne font pas parti des matériaux et ne sont pas inclus dans *M* tout comme les deux premières.

Finalement, *CI* comprend l'achat non stocké de matières et fournitures (comptabilité 606 française), qui elle-même comprend l'achat de carburant, donc il y a un croisement entre *CI* et *ACHACARB*. Donc nous ne pouvons pas utiliser *CI* directement. Par contre, il existe d'autres termes dans *CI* qui sont intéressants pour nous, comme les loyers et les charges locatives (*LOYCHLOC*), l'entretien et réparation (*REPARTOT*), les primes d'assurances (*PCH28S3*), les dépenses en poste et télécommunications (*RCH288*), etc. Donc selon la disponibilité des données, nous définissons les matériaux comme les achats de marchandise, matières premières et autres approvisionnements. Finalement, les dépenses correspondant aux matériaux sont égales à $RCH701 + RCH71 + RCH25 + RCH28S4 - ACHACARB - RCH2$. Pour obtenir la quantité, nous divisons cette nouvelle variable par le prix de matériaux $M = (RCH701 + RCH71 + RCH25 + RCH28S4 - ACHACARB - RCH26) / w_M$.

Prix de matériaux (w_M)

L'INSÉE ne produit pas d'indices de prix spécifiques pour les matériaux, surtout dans la configuration qui nous intéresse. Alors nous devons construire un indice de prix correspondant. Nous utiliserons un indice de prix de Laspeyres en nous basant sur le tableau des entrée-sorties.

Pour construire un l'indice de prix correspondant, nous utilisons le tableau entrées-sorties (TES) de l'économie française et cherchons les secteurs qui interagissent le plus le transport routier de voyageur et le transport routier de marchandise. Le TES est un tableau à double entrée, avec deux lectures possibles : une en ligne et une en colonne. Ce tableau rassemble dans un même cadre comptable les comptes de biens et services par produit et les comptes de production et d'exploitation des branches. Il décrit, produit par produit, l'ensemble des transactions de l'économie au cours d'une période, et la manière dont ces ressources ont été utilisées.

Depuis 1995, le TES s'appuie sur la nomenclature d'activité française (NAF). Et il y a cinq niveaux de détail : niveau D avec cinq branches, niveau E seize branches, niveau

F avec 39 branches, niveau G avec 118 branches et niveau H avec 472 branches (niveau très fin). Selon l'INSEE, la production est à peu près correctement estimée au niveau G de la nomenclature. Par contre, au niveau H, les données sont moins fiables, donc nous allons utiliser le TES au niveau G. La formule classique pour l'indice des prix de Laspeyres est le suivant :

$$W_m = \frac{\sum_{i=1}^n w_t^i \cdot x_0^i}{\sum_{i=1}^n w_0^i \cdot x_0^i} = \frac{\sum_{i=1}^n w_0^i \cdot x_0^i \cdot \frac{w_t^i}{w_0^i}}{\sum_{j=1}^n w_0^j \cdot x_0^j} = \sum_{i=1}^n \frac{w_0^i \cdot x_0^i}{\sum_{j=1}^n w_0^j \cdot x_0^j} \cdot \frac{w_t^i}{w_0^i} \quad (6-1)$$

Pour construire notre indice de Laspeyres, nous utilisons les tableaux entrées-sorties pour les années de 2000 à 2004. Nous utilisons les entrées dans la branche du transport de voyageur et nous retenons les produits énumérés au tableau 6-2 pour construire notre indice de Laspeyres. C'est cet indice que nous utiliserons pour déflater les autres matériaux du processus de production¹⁷.

Tableau 6-1 : L'indice des prix des matériaux

2000	2001	2002	2003	2004
100	101,645	103,998	105,159	106,493

Tableau 6-2 : Les différentes entrées utilisées pour construire l'indice de Laspeyres

Secteur	Branches
	1. FABRICATION DE PRODUITS CERAMIQUES ET DE MATERIAUX DE CONSTRUCTION
	2. INDUSTRIE DU CAOUTCHOUC
	3. SERVICES INDUSTRIELS DU TRAVAIL DES METAUX
	4. FABRICATION DE PRODUITS METALLIQUES
	5. COMMERCE ET REPARATION AUTOMOBILE
	6. COMMERCE DE GROS INTERMEDIAIRES

¹⁷ Le détail du TES est présenté dans l'annexe.

Services	7. MANUTENTION ENTREPOSAGE ET GESTION D'INFRASTRUCTURES
auxiliaires de	8. AGENCES DE VOYAGE
transports	9. ORGANISATION DU TRANSPORT DE FRET ET TRANSPORT SPATIAL
<hr/>	
Assurance	10. INTERMEDIATION FINANCIERE
	11. ASSURANCES
	12. LOCATION IMMOBILIERE
	13. TELECOMMUNICATION, POSTE ET COURRIER
	14. ACTIVITES INFORMATIQUES
	15. SERVICES PROFESSIONNELS
	16. ADMINISTRATION D ENTREPRISES
Services	17. LOCATION SANS OPERATEUR
opérationnels	18. SELECTION ET FOURNITURE DE PERSONNEL
	19. SECURITE NETTOYAGE ET SERVICES DIVERS AUX ENTREPRISES
	20. HOTELS ET RESTAURANTS
	21. EDUCATION MARCHANDE
<hr/>	

6-2-2 Inputs quasi-fixes

Dans la littérature qui porte sur la mesure de l'efficacité et la productivité des DMU, les auteurs supposent le plus souvent que les inputs sont sous le contrôle immédiat des DMU. En réalité, il y a des facteurs de production exogènes à la DMU qui entrent dans son processus de production sans être sous son contrôle immédiat. L'inefficacité des DMU peut être partiellement attribuée à ces inputs, mais on ne peut pas raisonnablement les faire entrer directement dans le calcul car ils ne sont pas variables à court terme, c'est-à-dire au moment de la décision de production. Dans la gestion courante, les inputs quasi-fixes agissent comme une contrainte supplémentaire sur l'allocation des ressources. Parce qu'une modification ces inputs est trop coûteuse ou impossible, leurs ajustements se fait à long terme ou ne se fait pas. Il faut prendre ce genre d'inputs en considération dans une analyse de l'efficacité des DMU. Dans notre thèse, les inputs quasi-fixes sont les véhicules et les infrastructures car à court terme, acheter les nouveaux véhicules est très coûteux pour les entreprises, et une fois la localisation d'une entreprise choisie, il est très compliqué d'en changer.

Véhicules

Pour les inputs quasi-fixes, nous devons utiliser cette fois le tableau « Données sur les transports de personnes ». Dans ce tableau, nous trouvons les variables concernant différents types de véhicules utilisées par les entreprises : le nombre d'autobus (*BUS*), le nombre d'autobus articulés (*BUSA*), le nombre total d'autobus en propriété (*BUSP11*), le nombre total d'autobus en crédit-bail (*BUSP12*), le nombre total d'autobus (*BUSP10*), le nombre total d'autocars, selon la taille: neuf places ou moins (*CAR10*), 10 à 29 places (*CAR20*),..., 60 places ou plus (*CAR50*).

Cependant, dans les faits la base de données est incomplète et ou incohérente. Il existe un grand nombre de cas pour lesquels la variable *BUS* est nulle mais pas *BUSP10*, ou inversement. Pour essayer d'avoir plus d'information concernant cette variable, nous adoptons la stratégie suivante : nous prenons la variable *BUS* comme la variable de base, car elle est quand même plus complète que *BUSP10*. Pour les entreprises avec *BUS* égal à zéro nous avons vérifié la valeur de *BUSP10* : si *BUSP10* n'est pas nul, nous utilisons *BUSP10* pour le nombre d'autobus, sinon nous considérons que le nombre d'autobus est nul pour l'entreprise.

Bien entendu, la taille des véhicules joue un rôle important, car un autobus ou un autocar de dix places n'a pas la même capacité qu'un autocar ou un autobus de 60 places pour transporter des voyageurs. Par conséquent, nous pondérerons le nombre d'autocars ou d'autobus par leur taille pour obtenir une nouvelle variable qui représente le nombre de places de véhicule pour une DMU.

Les infrastructures

Nous utilisons un indicateur d'accessibilité pour représenter les services fournis par les infrastructures dans notre fonction de production et fonction de coût. La méthodologie a été présentée dans le chapitre 5. Dans cette sous-section, nous présenterons notre calcul de l'indicateur d'infrastructures en utilisant l'équation 5-6 du chapitre précédent.

Pour calculer l'indicateur d'accessibilité, nous avons utilisé le logiciel de système d'information géographique MAPINFO. Ce logiciel nous permet de localiser notre DMU sur une carte française et de calculer la distance entre chaque zone d'emploi en France. Notre indicateur d'accessibilité est un indicateur individuel pour chaque zone d'emploi. Nous n'avons pas essayé de chercher un indicateur pour chacune entreprise ni pour chacune commune française pour les raisons suivantes :

1. La géolocalisation de chacune entreprise est un travail très lourd, surtout quand nous devons géo-localiser les DMU à la main (nous pouvons auto géo-localiser les objectifs par Mapinfo, mais ce service est payant et nécessite une base de données de haute qualité). Comme il nous était impossible d'obtenir une carte de la France à l'échelle des villes (avec les rues représentées), nous nous contentons d'une représentation communale de la France. En plus, dans notre base de données, l'adresse géographique des entreprises n'est pas complète, donc ce n'est pas une option de géo-localiser les entreprise au niveau de la rue.
2. Malheureusement, nous ne pouvons non plus géo-localiser les entreprises au niveau communal. Pour calculer l'indicateur d'accessibilité, nous devons calculer une matrice de distance, c'est-à-dire une matrice qui peut représenter la distance entre chaque commune de France. Au 1^{er} janvier 2010, il existe 36 682 communes en France. En excluant les communes d'outre-mer, il nous reste tout de même 36 560 communes, donc la matrice de distance sera une matrice de $36\,560 \times 36\,560$, et pour calculer cette distance ceci requiert une capacité de calcul à laquelle nous n'avons pas accès. Et même si nous pouvions trouver un ordinateur suffisamment performant pour lancer ce processus, un autre problème fait surface : le manque des données pour les étapes suivantes. Pour obtenir notre indicateur d'accessibilité, nous devons calculer deux déflateurs : le déflateur par rapport à la distance, ν , et représente le déflateur par rapport au temps, γ . Pour estimer ces deux déflateurs, nous avons besoin de lancer une estimation économétrique, dans laquelle la

variable dépendante est le flux de marchandise entre chaque zone, mais la qualité des données par commune n'est pas satisfaisante.

Pour toutes ces raisons, nous avons décidé de limiter notre analyse au niveau zone d'emploi. Ceci est possible car nous avons une carte fournie par CGDD au niveau des zones d'emploi et nous avons des données en bonne qualité pour ce niveau géographique. De plus, la zone d'emploi est plus fine que le département, et comme nous avons 371 zones d'emploi total, cela nous permet de lancer une analyse suffisamment précise et nous permet de géo-localiser des entreprises.

Pour estimer les paramètres ν et γ de l'équation 5-6, nous avons besoin des variables suivantes : L_i la longueur des routes totales pour chaque zone d'emploi, Vt_i , la vitesse moyenne pour chaque zone d'emploi, ST_i , la surface totale de chaque zone d'emploi, d_{ij} , la distance entre chaque zone d'emploi, t_{ij}^{con} le temps de congestion dans chaque zone d'emploi.

Dans la base de données fournie, nous avons les variables suivantes : le trafic longue distance des véhicules légers mesuré par le nombre de véhicules (V), le trafic longue distance des véhicules lourds mesuré par le nombre des véhicules poids lourd (P), la vitesse de chaque route pour les véhicules légers en km/h (VL), la vitesse de chaque route pour les véhicules du poids lourd en km/h (PL), le temps perdu pour cause de congestion (t^{con}), le flux de marchandise entre chaque zone d'emploi (F_{ij}) et la longueur de chaque autoroute en kilomètres (L_r). i et j représentent respectivement les zones d'emploi i et j , avec $i \neq j$.

Pour trouver la vitesse moyenne, nous pondérons notre variable par le trafic passé pour chaque autoroute, donc la vitesse moyenne pour la route r est égale à :

$$Vt_r = \frac{V_r}{V_r + P_r} * VL_r + \frac{P_r}{V_r + P_r} * PL_r. \quad (6-2)$$

où Vt_r est la vitesse moyenne pour la route r , V_r est le nombre de véhicules légers pour la route r , P_r est le nombre de véhicules poids lourds pour la route r et $r = 1, 2, \dots, N$ représente la route. Ensuite, pour chaque zone d'emploi i , nous calculons la moyenne des vitesse moyennes de chaque route qui se trouve dans la zone d'emploi i qui sera en définitive la vitesse moyenne pour cette zone d'emploi i :

$$Vt_i = N^{-1} \sum_{r=1}^N Vt_{ir}, \text{ où } N \text{ est le nombre de routes pour zone d'emploi } i. \text{ Nous mesurons}$$

également la longueur de toutes les routes dans une zone d'emploi i , qui est représenté par la variable L_i

Maintenant, il nous reste à expliquer ce que nous avons fait pour les variables qui demandent le traitement le plus difficile, soit la matrice de distances (une matrice de 371×371), le déflateur de la distance ν et le déflateur des temps perdus à cause de la congestion γ .

Pour chercher la matrice de distances, nous avons besoin du logiciel SIG et une carte de la France au niveau des zones d'emploi. La figure 6-1 nous montre la carte de la France au niveau zone d'emploi.



Figure 6-1 : La carte de la France au niveau des zones d'emplois

Pour calculer la distance entre chaque zone d'emploi, nous utilisons la méthode classique en économie spatiale qui consiste à mesurer la distance entre le centroïde de chacune zone d'emploi. Pour calculer la distance de centroïde à centroïde, nous avons identifié tous les centroïdes à partir du tableau des zones d'emploi. Pour ce faire, nous pouvons utiliser l'outil Mapbasic de Mapinfo. Nous obtenons la représentation de la Figure 6-2. C'est à partir de cette carte que nous mesurons les distances entre zones d'emploi.

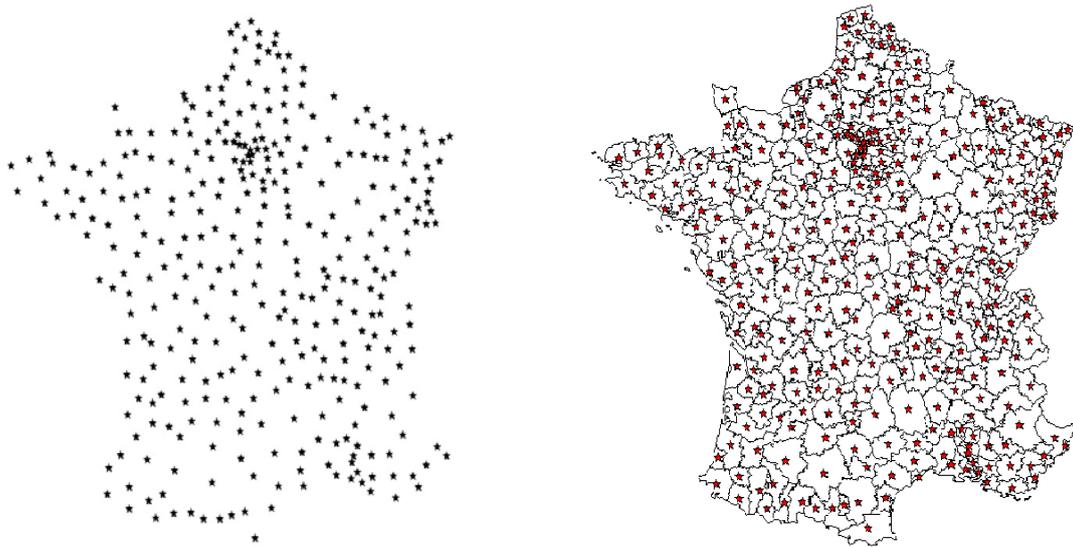


Figure 6-2 : La carte de centroïdes au niveau zone d'emploi

Pour obtenir la longueur totale des routes dans une zone d'emploi, nous devons croiser deux cartes de fond de Mapinfo, soit la carte de la France au niveau zone d'emploi et la carte des routes (qui contient a priori toutes les routes principales à l'horizon 2002) :

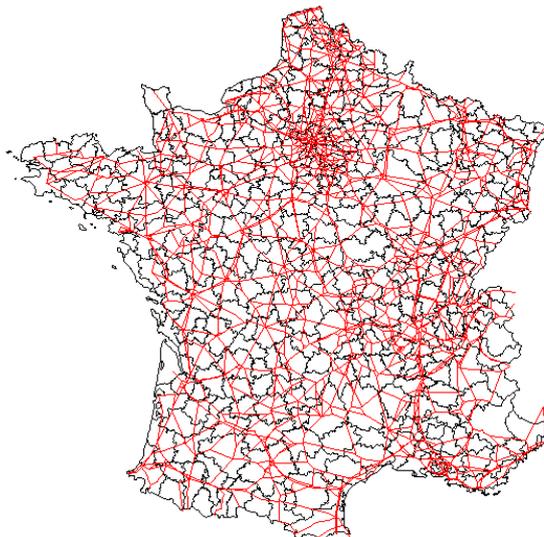


Figure 6-3 : La carte croisée (les lignes rouges représentent les routes principales)

Après avoir calculé la distance entre les zones d'emploi i et j (d_{ij}) et, la longueur totale des routes pour une zone d'emploi i (L_i), il nous reste à estimer le déflateur de distance (ν) et le déflateur de temps perdus à cause de la congestion (γ). Nous estimons notre déflateur par SAS selon le mécanisme présenté dans chapitre 5 et le résultat est présenté au tableau 6-3:

Tableau 6-3 : Déflateurs de distance et du temps perdus à cause de la congestion

Variable	Valeur estimée	Valeur p
ν	-0,217	< 0,001
γ	-0,139	< 0,001

A l'aide de ces variables et paramètres, nous pouvons calculer ensuite l'indicateur d'infrastructures pour chaque zone d'emploi. La figure 6-4 nous donne la distribution spatiale de l'indicateur d'infrastructures, chaque couleur correspond un niveau de l'indicateur, plus la couleur est foncée, plus l'indicateur est grand et plus le service fournis par les infrastructures est grand.

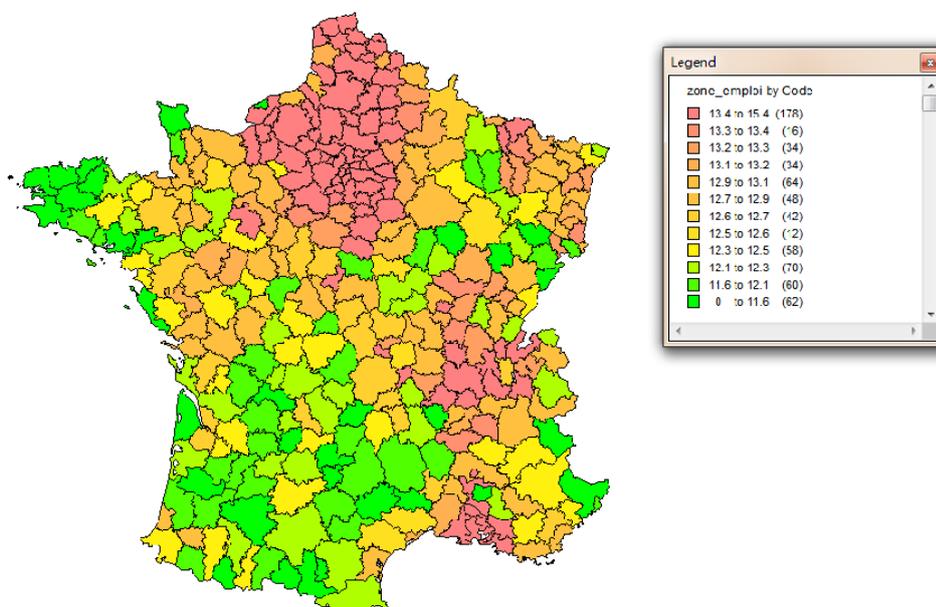


Figure 6-4 : La distribution d'indicateur des infrastructures par zone d'emploi

6-2-3 Outputs et coût variable

L'idéal pour mesurer l'output des entreprises serait d'avoir les kilomètres-véhicules parcourus et les kilomètres-passagers parcourus. Toutefois, la banque de données ne nous permet que de calculer le nombre de kilomètres parcourus pour l'ensemble des clientèles. Nous avons pris par conséquent cette variable pour mesurer l'output des entreprises, car dans le cas français, les DMU dans ce secteur sont souvent subventionnés par le gouvernement dans le but de satisfaire le besoin de déplacement des individus, donc la mesure des kilomètres parcourus pour l'ensemble des clientèles est satisfaisant comme proxy de l'output. Pour les kilomètres parcourus, nous avons retenu la variable *KMTURV10* (kilomètres parcourus, y compris à vide, total des activités de transport collectif de voyageurs).

Pour la fonction de coût dans le modèle semi-paramétrique, nous avons besoin d'information sur le coût total pour une DMU. Dans notre cas, il est égal à la somme des dépenses en inputs variables : main d'œuvre, énergie, entretien et réparation et matériaux.

6-3 Synthèse des données.

Les tableaux 6-4, 6-5 et 6-6 offrent une synthèse de la base de données. Au Tableau 6-4, nous donnons la source exacte et la description de chaque variable utilisée. Aux Tableaux 6-5 et 6-6, nous présentons les statistiques descriptives des variables utilisées pour les modèles primal et dual respectivement.

Tableau 6-4 : Sources et descriptions des variables

	Variable	Symbole	Données	Description	Sources
Output	Kilomètre	Y	$KMTURV10$	Kilomètres parcourus, y compris à vide, total des activités de transport collectif de voyageurs	EAE
Inputs variables	Travail	Quantité (L)	$TOTEFF$	Effectif annuel total en équivalent temps complet	EAE
		Prix (w_L)	D_L/L	Prix de main d'œuvre pour un employeur	Calculé
		Dépenses (D_L)	$RCH31+RCH32+RCH26$	La somme des salaires, des charges sociales et les frais de personnel extérieur à l'entreprise	EAE
	Carburant	Quantité (E)	D_E/W_E	La quantité de carburant par les entreprises	Calculé
		Prix (w_E)	Indice de prix	IPC-Ensemble des ménages-France métropolitaine-par fonction de consommation-Carburants	INSEE
		Dépenses (D_E)	$ACHACARB$	Achats de carburant	EAE
	Entretien et Répartition	Quantité (ENT)	D_{ENT}/W_{ENT}	La quantité sur l'entretien et répartition pour des entreprises	Calculé
		Prix (w_{ENT})	Indice de prix	IPC-Ensemble des ménages-France métropolitaine-Commerce et Répartition Automobile	INSEE
	Matériaux et Fournitures	Dépenses (D_{ENT})	$RCH28S2(REPARTOT)$	Autres achats et charges externes : entretiens et réparations	EAE
		Quantité (M)	D_M/W_M	Quantité de matériaux et de fournitures utilisés par les entreprises	Calculé
Prix (w_M)		Indice de prix de Laspeyres	Indice de prix de Laspeyres calculé à partir du tableau entrées-sorties	Calculé	
		Dépenses (D_M)	$RCH701+RCH71+RCH25+RCH28S4-ACHACARB-RCH26$	Dépense des entreprises sur les matériaux et fournitures	EAE
Inputs quasi-	Quantité de bus	K_B	$BUS+BUS_A$	Nombre d'autobus utilisé par les entreprises	EAE

fixes	Quantité de cars	K_{C1}	$CAR10$	Nombre total de minicars, 9 places ou moins	EAE
		K_{C2}	$CAR20$	Nombre total d'autocars, 10 à 29 places	EAE
		K_{C3}	$CAR30$	Nombre total d'autocars, 30 à 49 places	EAE
		K_{C4}	$CAR40$	Nombre total d'autocars, 50 à 59 places	EAE
		K_{C5}	$CAR50$	Nombre total d'autocars, 60 places ou plus	EAE
Infrastructures	Vitesse moyenne	Vt_i	$Vt_i = N^{-1} \sum_{r=1}^N Vt_{ir}$	Vitesse moyenne pour la zone d'emploi i	Calculé
	Longueur totale des autoroutes	L_i	$L_i = \sum_{r=1}^N L_{ir}$	Longueur des routes totales pour la zone d'emploi i	CGDD
	Surface	ST_i	SUP	Surface de la zone d'emploi i	INSEE
	Distance	d_{ij}	Carte réseau de Mapinfo	Distance entre le centroïde de chaque zone d'emploi i et j	Calculé
	Temps de congestion	t_{con}	$TEMPS_j$	Temps perdu à cause de congestion pour la route r	CGDD
	Flux de marchandises	F_{ij}	F_{ij}	Flux annuels en 2002 entre les zones d'emplois i et j	CGDD
	Indicateur d'infrastructures	Inf_i	$Inf_i = \left(\frac{L_i Vt_i}{ST_i} \right) * \sum_j L_j \exp\{-\nu d_{ij} - \gamma t_{ij}^{con}\}$	Indicateur d'accessibilité pour la zone d'emploi i	Calculé

Tableau 6-5 : Statistiques descriptives des variables du modèle primal

	Variable	Symbole	N	Maximum	Minimum	Écart-type	Moyenne	Médiane
Output	Kilomètre	Y	2 554	45 040 389	10 886	3 060 574,48	2 097 411,79	1 183 424,50
	Dépense variable	D_{COUT}	2 554	144 424	16,86	7 531,98	3 349,87	1594
Input variables		L	2 554	1 957	2	121,37	67,89	36
	Travail	w_L	2 554	80,86	0,50	7,42	26,19	26,13
		D_L	2 554	94433	1	4 444,43	1 931,42	916,50
		E	2 554	8 968,19	1,05	687,30	451,83	248,32
	Carburant	w_E	2 554	102,02	92,34	3,71	96,87	94,92
		D_E	2 554	9149	1	667,07	437,87	240
		ENT	2 554	3 997,19	1,78	251,84	141,56	69,47
	Entretien et Répartition	w_{ENT}	2 554	112,28	100	4,72	105,10	103,77
		D_{ENT}	2 554	4 488	2	267,11	149,16	73
	Matériaux et Fournitures	M	2 554	47 847,68	2,82	2 420,73	802,80	314,67
		w_M	2 554	106,49	100	2,37	103,46	104
		D_M	2 554	4 8635	3	2 501,16	831,42	323
Inputs quasi-fixes	Quantité de bus	K_B	2 554	716	0	31,52	7,74	0
		K_{car_p}	2 554	43 020	0	3 341,86	2 393,96	1 395
	Quantité de cars	K_{C1}	2 554	65	0	4,31	1,13	0
		K_{C2}	2 554	127	0	6,20	2,05	0
		K_{C3}	2 554	127	0	9,65	2,99	0
		K_{C4}	2 554	682	0	37,74	14,65	0

		K_{CS}	2 554	171	0	6,73	1,34	0
	Nombre de places des véhicules	KV	2 554	67 270	5	4 088,05	2 742,05	1 575
	Indicateur d'infrastructures	Inf_i	337	15,31	7,81	1,01	12,82	12,85
	Temps de congestion	t_{con}	337	120,43	5,85	11,40	20,02	17,42
Infrastructures	Vitesse moyenne	Vt_i	337	116,41	37,74	12,54	69,01	68,67
	Longueur	L_i	337	856,20	11,37	152,27	285,14	255,71
	Distance	d_{ij}	115 940	1 044,28	8,61	191,81	395,60	386,20
	Surface	ST_i	337	6 256,05	46,36	998,52	1 591,49	1 448,19
	Flux de marchandises	F_{ij}	116 281	4 280 771,20	1	29019,03	2318,28	79,50

6-4 Ajustement des variables de prix pour la fonction de coût

Dans notre base de données, les prix du carburant, de l'entretien et des réparations, et des matériaux et fournitures sont représentés par des indices de prix calculés, qui restent identiques pour toutes les entreprises pendant une année. Par contre, lors de notre estimation du modèle avec fonction de coût, nous avons besoin de variations de prix pour toutes les DMU sinon il n'y a pas de solution réalisable au problème d'optimisation. Pour rendre la solution réalisable, nous devons agréger les prix du carburant, de l'entretien et des réparations, des matériaux et fournitures en se servant des dépenses comme base de pondération. Nous obtenons une nouvelle variable de prix, qui correspond au prix des matériaux utilisé dans le processus de production pour une entreprise :

$$w_K = w_E * \frac{D_E}{D_E + D_{ENT} + D_F} + w_{ENT} * \frac{D_{ENT}}{D_E + D_{ENT} + D_F} + w_M * \frac{D_M}{D_M + D_M + D_M} . \quad (6-3)$$

Ceci entraîne aussi un ajustement pour un certain nombre de variables que nous utiliserons dans les modèles empiriques. Les nouvelles variables sont résumées aux Tableau 6-6 et 6-7 et les statistiques descriptives sont présentées au Tableau 6-8.

Tableau 6-6 : Variables utilisées dans le modèle paramétrique et la fonction primal du modèle semi-paramétrique

Variable	Symbole	Description
Output	<i>Y</i>	Kilomètres parcourus
Input variable	<i>L</i>	Quantité de main d'œuvre
	<i>E</i>	Quantité de carburant
	<i>ENT</i>	Quantité d'entretien et répartition
	<i>M</i>	Quantité des matériaux
Input quasi-fixe	<i>V</i>	Nombre de places de véhicules transportés
	<i>Inf</i>	Indicateur d'infrastructures

Tableau 6-7 : Variables utilisées dans la fonction duale du modèle semi-paramétrique

Variables	Symboles	Description
Coût	C	Coût variable
Input variable	w_L	Taux de salaire
	w_K	Prix de la variable agrégée (énergie, entretien, matériaux)
Output	Y	Kilomètres parcourus
Input quasi-fixe	V	Nombre de places des véhicules
	Inf_i	Indicateur d'infrastructures de la zone d'emploi i

Tableau 6-8 : Statistiques descriptives des variables utilisées dans le modèle dual

	Variable	Symbole	Maximum	Minimum	Écart-type	Moyenne	Médiane
Output	Kilomètre	Y	45 040 389	10 886	3 060	2 097	1 183
Coût	Dépense variable	D_{COUT}	144 424	16,86	7 531,98	3 349,87	1 594
Input variables	Travail	L	1 957	2	121,37	67,89	36
		w_L	80,86	0,50	7,42	26,19	26,13
		D_L	94 433	1	4 444,43	1 931,42	916,50
	Matériaux (variable agrégée)	K	58 388,10	11,98	3 187,14	1 394,28	664
		w_K	108,26	95,91	2,65	101,30	100
		D_K	5 8795	12	3 263,48	1 418,45	676
Inputs quasi-fixes	Véhicules	K_V	6 7270	5	4 088,05	2 742,05	1575
Infrastructures	Indicateur	Inf_i	15,31	7,81	1,01	12,90	12,86

Chapitre 7 : Application au transport routier de voyageurs avec un modèle non paramétrique DEA standard

7-1 Introduction

Dans ce chapitre nous utilisons la méthode DEA standard afin d'analyser l'efficacité du transport routier de voyageurs pour la France métropolitaine sur la période 2000 à 2004. Cette analyse se fera en introduisant deux inputs quasi-fixes dont l'indicateur d'infrastructures pour capter l'influence des infrastructures routières sur la performance des entreprises. Nous testons la présence d'efficacité d'échelle des DMU dans ce secteur et nous regarderons la relation entre la taille des entreprises et l'efficacité des entreprises. Nous utilisons le modèle non paramétrique et les algorithmes définis au chapitre 3. Nous terminerons par une analyse des résultats.

7-2 Applications DEA au secteur du transport

Les applications en DEA pour évaluer l'efficacité dans le secteur des transports sont devenues courantes. Les applications récentes incluent Hirschhausen et Cullmann (2008), Odeck (2008), Cullinae *et al.* (2006), Juan *et al.* (2003), Pacheco et Fernandes (2003), Karlaftis (2003), Viton (1997), Kerstens (1996) et plusieurs autres. Toutefois, à notre connaissance, il n'y a que Martin, Gutiérrez et Roman (2004) qui utilisent un indicateur d'accessibilité dans une application de DEA.

Hirschhausen et Cullmann (2008) ont utilisé la méthode DEA avec bootstrap pour étudier l'efficacité de 179 entreprises du transport par autobus en Allemagne pour la période 1990-2004. Ils ont utilisé les kilomètres parcourus par bus comme output, le nombre d'employeurs (temps plein et temps partiel) et le nombre de bus comme inputs. Ils ont trouvé une efficacité assez basse en moyenne (39,5 % en CCR et 42,8 % en BCC) et les petites entreprises montrent une inefficacité d'échelle. Viton (1997) a étudié l'efficacité des bus multifonctions aux États-Unis. Il a utilisé la méthode

DEA avec un ensemble de données qui comprend 217 DMU. Les variables d'output qu'il utilise sont les kilomètres parcourus et le nombre de voyages de passagers. Les variables d'inputs sont la vitesse moyenne (qui peut être considéré comme proxy de l'accessibilité en utilisant la mesure simple, mais comme nous avons mentionné dans le chapitre 5, cette méthode ne prend pas en considération les autres facteurs pertinents), l'âge moyen des bus, la vitesse moyenne des bus, les heures de travail de maintenance, les heures de travail administratif, etc. Dans son étude, Viton a trouvé un faible degré d'inefficacité et 80 % des DMU étudiés étaient techniquement efficaces. Cowie et Asenova (1999) ont étudié la performance de l'industrie du transport par autocar britannique à la lumière des réformes fondamentales sur la propriété et la réglementation qui a eu lieu depuis 1980. Ils ont utilisé la méthode DEA avec des données de 1995/1996. Les revenus d'exploitation ont été utilisés comme output, en notant que la variable idéale pour mesurer l'output aurait été les kilomètres parcourus par passager. Leurs variables d'inputs prennent en considération à la fois les éléments de la main-d'œuvre et le capital. La main-d'œuvre a été mesurée par le nombre d'employés (à la fois de gestion et d'exploitation), tandis que le capital a été mesuré comme le nombre d'autobus exploités divisés en autobus avec moins de 35 sièges et les autobus de 35 sièges ou plus. Les résultats de leur étude ont montré un haut degré d'inefficacité technique dans l'industrie britannique du transport par autobus. Les auteurs concluent que cela pourrait être causé en partie par la structure de marché oligopolistique qui a émergé à la suite de la privatisation qui a commencé dans les années 80. Ils ont également trouvé que les entreprises privées sont plus efficaces que les entreprises publiques et ils ont identifié des rendements d'échelles croissants pour les petites entreprises et des rendements d'échelle décroissants pour les grandes entreprises.

Jorgensen *et al.* (1997) ont utilisé la méthode de frontière stochastique (*SFA*) pour analyser l'efficacité de 170 entreprises norvégiennes de transport par autobus subventionnées par le gouvernement. Ils ont trouvé une inefficacité de 7 à 14 % en

moyenne. Odeck et Alkadi (2001) ont également examiné la performance des entreprises d'autobus norvégiennes, mais ils ont utilisé la méthode DEA. Ils ont découvert un potentiel d'amélioration de l'efficacité d'environ 28 %. Kersten (1996) a aussi lancé une étude pour le secteur de transport avec les données françaises. Il a trouvé une efficacité de 75 % à 88 % selon la méthode utilisée (DEA ou FDH).

7-3 Modèle et résultats

Les modèles empiriques que nous utilisons sont des suivants :

$$\theta_i^{CCR} = \min_{\theta_0} \left\{ \theta_0 \mid \sum_{j=1}^N \lambda_j y_j \geq y_0, \sum_{j=1}^N \lambda_j X_j \leq \theta_0 X_0, \sum_{j=1}^N \lambda_j K_j \leq K_0 \right\} \quad (7-1)$$

$$\theta_i^{BCC} = \min_{\theta_0} \left\{ \theta_0 \mid \sum_{j=1}^N \lambda_j y_j \geq y_0, \sum_{j=1}^N \lambda_j X_j \leq \theta_0 X_0, \sum_{j=1}^N \lambda_j K_j \leq K_0, \sum_{j=1}^N \lambda_j = 1 \right\}, \quad (7-2)$$

où $X = [L \ E \ ENT \ M]$ représente le vecteur d'inputs variables, et $K = [V \ Inf]$ représente le vecteur d'inputs quasi-fixes. θ_i^{CCR} est le coefficient d'efficacité pour la DMU i estimé par le modèle CCR, θ_i^{BCC} est le coefficient d'efficacité pour la DMU i estimé par le modèle BCC, y est l'output mesuré par les kilomètres parcourus, L est la quantité de main-d'œuvre, E est la quantité d'énergie, ENT est la quantité d'entretien et réparation, M est la quantité de matériaux et fournitures, V est le nombre de véhicules de transport, et inf est l'indicateur d'infrastructures.

Notre intérêt majeur réside dans l'inférence concernant les caractéristiques des rendements d'échelle des entreprises pour les différentes régions. De plus, nous désirons lier la nature des rendements d'échelle des entreprises à l'influence des différentes caractéristiques des zones d'emploi et de l'indice d'accessibilité.

Aux tableaux 7-1 et 7-2, nous présentons la moyenne pondérés par le coût variable de

chaque région des coefficients d'efficacité technique : $\theta_j = \sum_{i=1}^n \theta_{ij} \cdot (CV_{ij} / \sum_{i=1}^n CV_{ij})$, où

θ_j est le coefficient d'efficacité moyen de la région j , θ_{ij} est le coefficient d'efficacité technique de l'entreprise i qui se trouve dans la région j , CV_{ij} est le coût variable de l'entreprise i qui se trouve dans la région j .

Tableau 7-1 : Statistique des coefficients d'efficacité pondéré par le coût (CCR)

Région	Nombre d'observations	Moyenne	Écart-type
Alsace	68	0,830	0,148
Picardie	79	0,828	0,115
Languedoc Roussillon	125	0,819	0,110
Rhône Alpes	253	0,813	0,109
Franche Comté	79	0,810	0,118
Haute Normandie	81	0,808	0,106
Basse Normandie	63	0,803	0,102
Poitou Charentes	69	0,800	0,108
Cotes d'Azur	215	0,798	0,101
Pays de la Loire	181	0,796	0,131
Bretagne	103	0,794	0,110
Lorraine	110	0,793	0,098
Nord Pas de Calais	172	0,790	0,108
Aquitaine	168	0,788	0,138
Centre	94	0,785	0,105
Champagne Ardenne	41	0,783	0,139
Limousin	77	0,779	0,239
Midi Pyrénées	188	0,776	0,089
Auvergne	63	0,769	0,105
Bourgogne	71	0,765	0,120
Ile de France	254	0,761	0,125
Total	2554	0,799	0,117

Tableau 7-2 : Statistique des coefficients d'efficacité pondéré par le coût (BCC)

Région	Nombre d'observations	Moyenne	Écart-type
Alsace	68	0,838	0,143
Franche Comté	79	0,832	0,122
Languedoc Roussillon	125	0,828	0,109
Picardie	79	0,826	0,108
Bretagne	103	0,820	0,113
Rhône Alpes	253	0,818	0,120
Haute Normandie	81	0,817	0,122
Basse Normandie	63	0,815	0,102
Cotes d'Azur	215	0,814	0,124
Poitou Charentes	69	0,812	0,110
Pays de la Loire	181	0,808	0,108
Midi Pyrénées	188	0,800	0,102
Lorraine	110	0,798	0,110
Aquitaine	168	0,796	0,106
Champagne Ardenne	41	0,795	0,115
Nord Pas de Calais	172	0,793	0,232
Centre	94	0,790	0,117
Limousin	77	0,785	0,125
Auvergne	63	0,783	0,103
Bourgogne	71	0,778	0,125
Ile de France	254	0,769	0,123
Total	2554	0,812	0,117

Tableau 7-3 : Statistique des coefficients d'efficacité pondéré par le coût par année

Région	Nombre d'observations	Moyenne	Écart-type
CCR			
2000	496	0,793	0,120
2001	508	0,785	0,135
2002	507	0,808	0,114
2003	516	0,798	0,110
2004	527	0,802	0,130
Total			
BCC			
2000	496	0,804	0,118
2001	508	0,803	0,132
2002	507	0,819	0,125
2003	516	0,804	0,110
2004	527	0,816	0,113
Total	2 554	0,812	0,117

Commençons par analyser les résultats par région. Nous estimons une frontière inter-temporelle, plus précisément, une frontière avec une base de données commune pour toutes les années. Ici, chaque observation est une entreprise sans tenir compte de la structure en panel des données. Pour le modèle CCR (tableau 7-1), nous obtenons une efficacité technique en moyenne de 79,9 % dans l'hypothèse REC où nous supposons une taille optimale des entreprises. Puisque dans le modèle BCC, nous relâchons l'hypothèse de rendements d'échelles constants, les coefficients BCC (tableau 7-2) ne représentent que l'inefficacité technique pure des entreprises de transport. Par conséquent, nous éliminons l'effet d'échelle et nous comparons les entreprises de taille similaire. Nous voyons maintenant qu'il y a une légère augmentation de l'efficacité pour les entreprises et nous pouvons obtenir en moyenne un coefficient de 81,2 %.

Nous avons également reporté les moyennes pour les années retenues à partir des résultats sur toutes les années (tableau 7-3). Nous remarquons que les coefficients

sont relativement stables au cours de ces cinq années, avec une augmentation de l'efficacité de 1,59 % pour modèle CCR et 1,49 % pour modèle BCC.

Ce sont les régions d'Alsace et de Franche Comté qui ont les coefficients d'efficacité les plus élevés (avec une moyenne de 83,82 %) et les entreprises de Bourgogne, d'Auvergne et d'Ile de France sont les moins efficaces (avec une moyenne de 76,90 %).

La figure 7-1 nous montre la différence entre le coefficient d'efficacité calculé par les modèles BCC et CCR sur les entreprises regroupées selon le chiffre d'affaires. Nous pouvons identifier une tendance assez claire : la plus grande partie des petites entreprises montrent une différence assez significative. Nous avons tout d'abord regroupé les entreprises de part et d'autres de la médiane du chiffre d'affaires, puis nous avons calculé la moyenne de $(\theta_{bcc} - \theta_{ccr})$ pour ces deux groupes. Nous avons trouvé une différence d'efficacité de 0,67 % et 1,57 % pour les grandes et les petites entreprises respectivement. Nous avons refait cet exercice, mais cette fois avec le premier et le dernier décile. Nous trouvons alors des différences d'efficacité de 1,57% et 3,91% pour les plus grandes et les plus petites entreprises respectivement. Cela indique clairement une inefficacité plus marquée pour les petites entreprises. Ces résultats sont une indication d'inefficacité d'échelle et nous pouvons immédiatement conclure que les petites entreprises de transport de voyageurs ont un désavantage d'échelle.

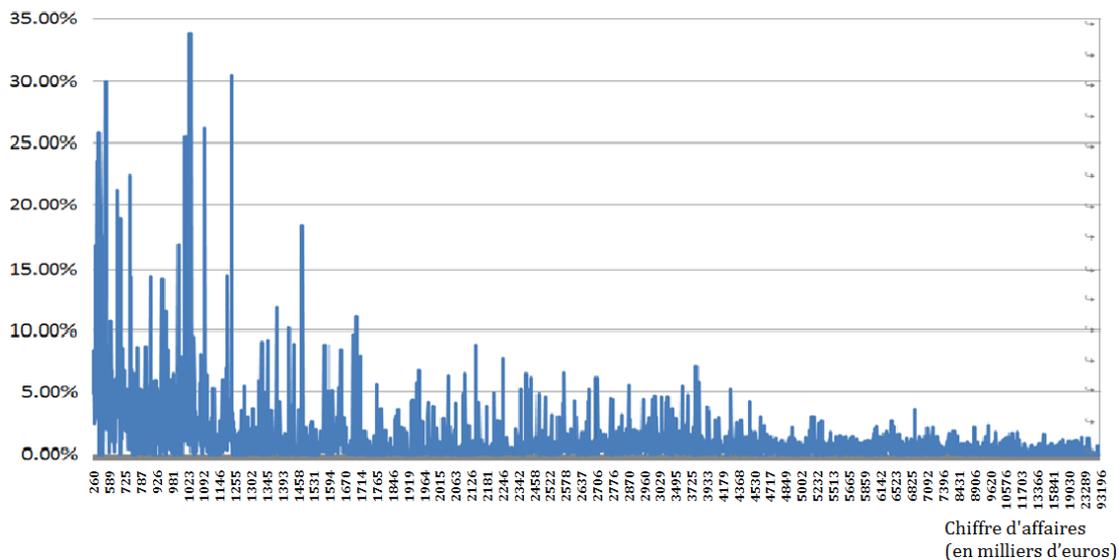


Figure 7-1 : Différences entre les coefficients des modèles BCC et CCR

Maintenant, nous enlevons l'indicateur d'accessibilité comme input pour regarder en quoi cela modifie nos résultats. Les résultats sont présentés au tableau 7-4. Si nous enlevons l'indicateur d'accessibilité dans les inputs, nous constatons qu'en moyenne le coefficient d'efficacité diminue légèrement (pour le BCC, de 81,23 % à 80,01 %). La différence n'est pas très significative (une augmentation de 1,6 %), cela peut s'expliquer par le fait que le modèle sans infrastructures est mal spécifié et biaisé dans le contexte où les firmes utilisent les infrastructures comme inputs. Ces infrastructures sont données aux entreprises comme des inputs qu'elles ne peuvent pas contrôler, au minimum à court terme, puisque une fois que les entreprises ont choisi leur localisation, les infrastructures fournis aux entreprises sont fixées. D'autre part, l'augmentation du nombre de dimensions se traduira par une perte de DMU comparables et ainsi on se rapproche de la frontière pour le DEA. Toutefois, nous pouvons analyser les effets distributifs de cette spécification du modèle en comparant les scores d'efficacité individuelle. Dans la figure 7-2, les entreprises sont ordonnées selon les indicateurs d'infrastructures. Nous pouvons voir clairement que les entreprises dans la région où l'accessibilité aux infrastructures est plus faible profitent le plus de l'intégration de la variable environnementale.

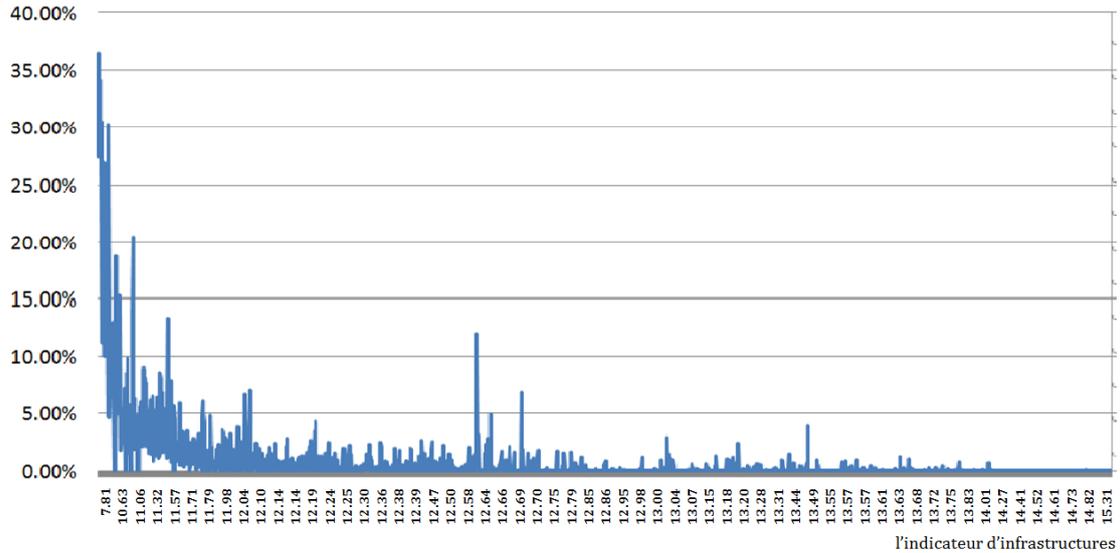


Figure 7-2 : Différence d'efficacité du modèle BCC avec et sans indicateur– classé par ordre croissant de l'indicateur d'infrastructures.

Nous avons utilisé l'équation 3-6 pour calculer les rendements d'échelles des DMU. Les résultats sont présentés dans le tableau 7-4. Comme nous le voyons, les petites entreprises sont moins efficaces que les grandes et toutes les entreprises montrent des rendements d'échelles croissants quand on inclut les infrastructures. Finalement, nous présentons le coefficient d'efficacité calculé avec et sans infrastructures au tableau 7-5. Nous interprétons ces résultats en disant que si nous ne considérons pas les infrastructures, nous surestimons l'inefficacité des entreprises.

Tableau 7-4 : Efficacité et les rendements d'échelles selon la taille des entreprises

Modèle	Échantillon	Avec infrastructures		Sans infrastructures	
		Efficacité moyenne	Médiane des rendements d'échelle	Efficacité moyenne	Médiane des rendements d'échelle
CCR	Échantillon complet	0,799		0,779	
	Petites entreprises	0,695		0,683	
	Entreprises de taille moyenne	0,764		0,762	
	Grandes entreprises	0,811		0,804	
BCC	Échantillon complet	0,812	1,118	0,801	0,886
	Petites entreprises	0,711	1,167	0,701	0,916
	Entreprises de taille moyenne	0,786	1,133	0,775	0,882
	Grandes entreprises	0,817	1,060	0,810	0,846

Tableau 7-5 : Comparaison des efficacités moyennes

Année	Avec infrastructures	Sans infrastructures
CCR		
2000	0,793	0,775
2001	0,785	0,766
2002	0,808	0,789
2003	0,798	0,778
2004	0,802	0,787
BCC		
2000	0,804	0,794
2001	0,803	0,796
2002	0,819	0,808
2003	0,804	0,798
2004	0,816	0,805

Conclusion

Nous avons utilisé la méthode DEA habituelle pour analyser l'efficacité des entreprises du transport routier de voyageurs pour la France métropolitaine pour la période 2000-2004. Nous avons utilisé deux modèles, le CCR et le BCC, avec et sans infrastructures, pour mesurer la performance des entreprises et la nature des rendements d'échelle.

Nos résultats indiquent que ce secteur en France est relativement efficace, mais la plupart des petites entreprises souffrent d'inefficacités d'échelle. Les petites entreprises qui se trouvent dans une zone d'emploi dans laquelle le niveau des infrastructures est plus faible profitent le plus de l'intégration de la variable environnementale. En plus, toutes les entreprises montrent des rendements d'échelle croissants dans ce secteur. Pour augmenter l'efficacité productive et donc diminuer le coût moyen dans ce secteur, le gouvernement doit orienter les entreprises de ce secteur à chercher des synergies avec les firmes proches (mettre en commun les bus et les garages, etc.) et essayer de fusionner les entreprises.

Chapitre 8. Application empirique avec le modèle semi-paramétrique

Le chapitre précédent présentait une application d'un modèle non paramétrique (DEA) le plus couramment utilisé aux DMU de notre échantillon d'entreprises du secteur du transport routier de voyageurs en France. Le présent chapitre se concentrera sur l'application du modèle semi-paramétrique d'Aigner et Chu (1968) aux entreprises de ce secteur. Dans un premier temps, nous présenterons les formes fonctionnelles que nous avons choisies pour le modèle empirique. Ensuite, nous présenterons séparément les résultats obtenus avec le modèle primal (fonction de production) et le modèle dual (fonction de coût). Nous présenterons également le résultat de simulation qui nous donne l'aspect statistique des paramètres obtenu. Nous analyserons les résultats obtenus pour déterminer l'impact des infrastructures sur la performance des entreprises, puis calculerons et analyserons les résultats pour les rendements d'échelle, le taux de substitution entre les inputs, etc. Ensuite, nous effectuerons une comparaison entre les résultats du modèle semi-paramétrique et ceux du modèle non-paramétrique de DEA. Nous terminerons ce chapitre avec une comparaison des estimations du modèle semi-paramétrique avec et sans les infrastructures comme variable explicative.

8-1 La spécification du modèle : forme fonctionnelle

Avant d'appliquer le modèle d'Aigner et Chu (1968) à notre base de données, il nous faut choisir une forme fonctionnelle pour représenter la technologie du secteur de transport de passagers. À la jonction entre la théorie du producteur et les applications il existe une littérature qui consiste spécifiquement à suggérer des formes fonctionnelles pour les fonctions de production ou de coût. L'intérêt des économistes dans ce domaine remonte à Cobb et Douglas (1928) et Hotelling (1932). Le choix de la forme fonctionnelle, du fait qu'elle peut imposer des restrictions sur la nature des

résultats, n'est pas sans conséquence. Par exemple, la forme Cobb-Douglas impose une élasticité de substitution unitaire entre les facteurs de production. Ces restrictions ont incité les économistes à suggérer d'autres formes fonctionnelles qui sont moins contraignantes *a priori*. Diverses formes ont été proposées afin de relâcher les contraintes de la forme Cobb-Douglas. Bien que parfois intéressantes (par exemple, la forme CES proposée par Arrow *et al.*, 1961), ces formes ne résultaient pas d'un processus systématique. Il faudra attendre les travaux de Diewert (1974) pour entrevoir une démarche formelle basée sur la théorie de l'approximation des fonctions analytiques. En particulier, un ensemble de formes reposant sur des approximations de Taylor de deuxième ordre, appelées formes flexibles, a été proposé. La notion de formes flexibles a depuis été généralisée au cas d'autres types d'approximations. On remarquera entre autres les formes fonctionnelles reposant sur les développements limités de Laurent (Barnett, 1983) et de Fourier (Gallant, 1981).

Aigner et Chu utilisent la programmation linéaire pour calibrer des technologies sous la contrainte que les propriétés découlant de la théorie économique soient respectées. La forme fonctionnelle qu'ils ont utilisée est une fonction de production Cobb-Douglas. Le théorème de Stone-Weierstrass rend explicite le fait que toute fonction continue peut être approchée de façon arbitrairement proche par un polynôme d'ordre suffisamment grand. Parmi les polynômes possibles, nous avons choisi celui proposé par Taylor, dans lequel les coefficients sont les dérivées de la fonction prise en un point. Avec les variables que nous avons choisies dans le chapitre 5, nous allons utiliser deux formes fonctionnelles dans notre modèle empirique, soit la forme Cobb-Douglas, une approximation au premier ordre et la forme Translog, une approximation de deuxième ordre, toutes deux après transformation logarithmique des variables. Ci-après, nous allons présenter la forme Translog, car la forme Cobb-Douglas est simplement un cas particulier de la forme Translog pour laquelle les termes de deuxième ordre sont nuls. Pour la fonction de production, nous aurons (Christensen, Jorgenson et Lau, 1971 et 1973) :

$$\begin{aligned} \ln y = & \alpha_o + \sum_{p=1}^P \alpha_{x_p} \ln x_p + \sum_{s=1}^S \alpha_{k_s} \ln k_s \\ & + \frac{1}{2} \sum_{p=1}^P \sum_{p'=1}^P \beta_{pp'} \ln x_p \ln x_{p'} + \frac{1}{2} \sum_{s=1}^S \sum_{s'=1}^S \beta_{ss'} \ln k_s \ln k_{s'} + \sum_{p=1}^P \sum_{s=1}^S \beta_{ps} \ln x_p \ln k_s, \end{aligned} \quad (8-1)$$

où α_o représente la constante ; $\mathbf{x} = (L, E, ENT, M)$ correspond aux inputs variables (main d'œuvre, L , énergie, E , entretien et réparations, ENT , et matériaux, M) ; et $\mathbf{k} = (V, Inf)$ représente les inputs quasi-fixes (nombre de places des véhicules, V , et indicateur d'infrastructures, Inf). Les α et β sont des paramètres à estimer. Le théorème de Clairaut-Schwarz sur les fonctions deux fois continûment différentiable implique que la hessienne est symétrique et ainsi $\beta_{pp'} = \beta_{p'p}$, $\beta_{ss'} = \beta_{s's}$ et $\beta_{ps} = \beta_{sp}$.

Les mêmes formes fonctionnelles seront utilisées pour la fonction de coût. Encore une fois, nous présentons la forme Translog de la fonction de coût variable :

$$\begin{aligned} \ln C = & \varphi_o + \sum_{p=1}^P \varphi_{w_p} \ln w_p + \sum_{s=1}^S \varphi_{k_s} \ln k_s + \varphi_y \ln y \\ & + \frac{1}{2} \sum_{p=1}^P \sum_{p'=1}^P \gamma_{pp'} \ln w_p \ln w_{p'} + \sum_{p=1}^P \sum_{s=1}^S \gamma_{ps} \ln w_p \ln k_s + \sum_{p=1}^P \gamma_{py} \ln w_p \ln y \\ & + \frac{1}{2} \sum_{s=1}^S \sum_{s'=1}^S \gamma_{ss'} \ln k_s \ln k_{s'} + \sum_{s=1}^S \gamma_{sy} \ln k_s \ln y + \frac{1}{2} \gamma_{yy} (\ln y)^2, \end{aligned} \quad (8-2)$$

où φ_o représente la constante et $\mathbf{w} = (w_L, w_E, w_{ENT}, w_M)$ correspond au vecteur de prix des inputs variables. Pour les mêmes raisons que précédemment, les paramètres de second ordre respecterons des propriétés de symétrie : $\gamma_{pp'} = \gamma_{p'p}$, $\gamma_{ss'} = \gamma_{s's}$ et $\gamma_{ps} = \gamma_{sp}$. Les φ et γ sont des paramètres à estimer.

Pour forcer la frontière de production (ou la frontière de coût dans l'espace dual) à envelopper les observations le plus étroitement possible, nous estimons ces paramètres par l'approche exposée dans le chapitre 3, que ce soit avec les équations (3-9) et (3-10) pour la fonction de production ou (3-29) et (3-30) pour la fonction de coût. Les deux critères retenus sont la minimisation de la somme des erreurs ou la

minimisation de la somme des carrés des erreurs. Les modèles sont présentés dans le tableau 8-1, $\ln y_i$ et $\ln C_i$ correspondent à $\ln y$ et $\ln C$ de la DMU i , qui sont définis respectivement par les équations (8-1) et (8-2); $\ln y_i^{obs}$ et $\ln C_i^{obs}$ correspondent respectivement à l'output observé et au coût variable observés pour la DMU i ¹⁸.

Tableau 8-1 : Les modèles empiriques

	Fonction de production	Fonction de coût
Minimisation de la somme des erreurs	$\min \sum_{i=1}^N \varepsilon_i = \min \sum_{i=1}^N (\ln y_i - \ln y_i^{obs})$ $s.c : \ln y_i \geq \ln y_i^{obs}$	$\min \sum_{i=1}^N \varepsilon_i = \min \sum_{i=1}^N (\ln C_i^{obs} - \ln C_i)$ $s.c : \ln C_i^{obs} > \ln C_i$
Minimisation de la somme du carré des erreurs	$\min \sum_{i=1}^N \varepsilon_i^2 = \min \sum_{i=1}^N (\ln y_i - \ln y_i^{obs})^2$ $s.c : \ln y_i \geq \ln y_i^{obs}$	$\min \sum_{i=1}^N \varepsilon_i^2 = \min \sum_{i=1}^N (\ln C_i^{obs} - \ln C_i)^2$ $s.c : \ln C_i^{obs} > \ln C_i$

8-2 Ajustement du modèle

Une fonction de production et une fonction de coût doivent respecter un ensemble d'exigence. Pour une fonction de production avec inputs quasi-fixes, $y = F(\mathbf{x}, \mathbf{k})$, les conditions classiques sont qu'elle est deux fois continument différentiable ($F \in C^2$); monotone croissante en \mathbf{x} et \mathbf{k} ($\nabla F > 0$) et fortement quasi-concave en \mathbf{x} et \mathbf{k} , ($\phi^T \nabla^2 F \phi < 0$ pour tout $\phi \neq 0$ tel que $\phi^T \nabla F = 0$). Ces conditions n'impliquent aucune restriction sur la nature des rendements d'échelle. Elles sont suffisantes pour garantir l'existence de la fonction de coût variable et en donner une caractérisation minimale. Cette fonction de coût est obtenue par la résolution du problème suivant :

¹⁸ Dans le tableau, nous n'avons pas distingué la forme fonctionnelle. Puisque nous avons deux formes fonctionnelles, donc nous avons au total huit modèles à estimer.

$$C(\mathbf{w}, \mathbf{k}, y) = \min_x \{ \mathbf{w}' \mathbf{x} / y \leq F(\mathbf{x}, \mathbf{k}) \} = \mathbf{w}' \mathbf{x}(\mathbf{w}, \mathbf{k}, y)$$

avec $\mathbf{x}(\mathbf{w}, \mathbf{k}, y)$ le vecteur des demandes conditionnelles de facteurs variables. Ce problème de minimisation indique clairement que les propriétés de la fonction de coût, définie dans l'espace dual (des prix des facteurs variables) sont reliées à celles de la fonction de production. En fait, il est facile de démontrer la proposition suivante¹⁹: si la fonction de production est caractérisée par les propriétés que nous venons d'énoncer, alors la fonction de coût sera deux fois continûment différentiable ($C \in C^2$); monotone croissante en \mathbf{w} et y , décroissante en \mathbf{k} ($C_w > 0, C_y > 0, C_k < 0$); homogène de degré un en \mathbf{w} ($C_w' \mathbf{w} = C$); concave en \mathbf{w} ($\phi^T C_{ww} \phi < 0$ pour $\phi \neq \alpha \mathbf{w}, \alpha \in \mathbb{R}$) et convexe en \mathbf{k} ($C_{kk} > 0$). Par le théorème de l'enveloppe, nous obtenons $C_w = \mathbf{x}$ et $C_k = -z$, où z est le prix implicite des inputs quasi-fixes. Ce dernier résultat est connu sous le nom de lemme de Shephard. Ces conditions se manifestent par des restrictions sur la valeur des paramètres et le domaine des variables.

Pour la fonction de production, il faut que les élasticités de l'output par rapport aux inputs variables (ξ_p) et par rapport aux inputs quasi-fixes (η_s) soient toujours positives, avec :

$$\xi_p = \frac{\partial \ln y}{\partial \ln x_p} = \alpha_{x_p} + \sum_{p'} \beta_{pp'} \ln x_{p'} + \sum_s \beta_{ps} \ln k_s > 0, \forall p, \quad (8-3)$$

$$\eta_s = \frac{\partial \ln y}{\partial \ln k_s} = \alpha_{k_s} + \sum_{s'} \beta_{ss'} \ln k_{s'} + \sum_p \beta_{ps} \ln x_p > 0, \forall s, \quad (8-4)$$

avec \mathbf{p} et $\mathbf{p}' = (L, E, ENT, M)$, \mathbf{s} et $\mathbf{s}' = (V, Inf)$.

¹⁹ Si nous supposons la stricte concavité de la fonction de production au lieu de la forte quasi-concavité, nous aurons, en plus, la convexité de la fonction de coût en y .

Pour la fonction de coût, l'homogénéité de degré un dans les prix implique :

$$\begin{aligned}
 \sum_{p=1}^P \varphi_p &= 1 \\
 \sum_{p=1}^P \gamma_{pp'} &= 0, \forall p' \\
 \sum_{p=1}^P \gamma_{py} &= 0 \\
 \sum_{p=1}^P \gamma_{ps} &= 0, \forall s.
 \end{aligned}
 \tag{8-5}$$

En appliquant le lemme de Shephard, l'élasticité du coût par rapport aux prix d'un input variable nous donne la part de cet input variable dans le coût variable et elle peut être calculée comme suit :

$$S_p = \frac{\partial \ln C}{\partial \ln w_p} = \varphi_{w_p} + \sum_{p'} \gamma_{pp'} \ln w_{p'} + \sum_s \gamma_{ps} \ln k_s + \gamma_{py} \ln y > 0, \forall p.
 \tag{8-6}$$

Ces parts doivent être comprises entre 0 et 1 et leurs somme doit être égale à 1.

L'élasticité du coût par rapport à l'output, aussi appelée élasticité d'échelle, doit être positive et est donnée par :

$$\xi_{C_y} = \frac{\partial \ln C}{\partial \ln y} = \varphi_y + \sum_{p=1}^P \gamma_{py} \ln w_p + \sum_s \gamma_{sy} \ln k_s + \gamma_{yy} \ln y > 0.
 \tag{8-7}$$

L'élasticité du coût par rapport aux inputs quasi-fixes (incluant les infrastructures) doit être négative et s'écrit :

$$\xi_{C_{k_s}} = \frac{\partial \ln C}{\partial \ln k_s} = \varphi_{k_s} + \sum_{p=1}^P \gamma_{ps} \ln w_p + \sum_{s'=1}^S \gamma_{ss'} \ln k_{s'} + \sum_{s'=1}^S \gamma_{sy} \ln y < 0, \forall s.
 \tag{8-8}$$

Notons que l'équation (8-8) est l'expression de la part implicite des dépenses en inputs quasi-fixes dans le coût variable. Si l'input quasi-fixes « infrastructures » est productif, sa part implicite reflète la réduction en pourcentage du coût variable suite à une augmentation en pourcentage des infrastructures.

Donc pour la fonction de coût sous forme Translog, en plus des contraintes sous forme d'égalités données par l'équation (8-5), nous ajoutons les contraintes que les paramètres doivent être tels que les équations (8-6) et (8-7) sont positives et que les équations (8-8) sont négatives.

8-3 Mesure de la technologie

La fonction Translog est une approximation de Taylor au 2^{ème} ordre après transformation logarithmique des arguments d'une fonction de production ou de coût non observée. Elle n'impose pas de restrictions sur le degré de substituabilité entre les facteurs de production ni sur la nature des rendements d'échelle, sauf lorsqu'on impose certaines contraintes découlant de la théorie (Lau, 1986). Ces derniers peuvent par ailleurs être reconstitués à partir des coefficients de la fonction estimée. En fait, toutes les mesures technologiques sont exprimées en termes de la jacobienne et de la hessienne de la fonction de production pour les mesures primales et de la fonction de coût pour les mesures duales. Un avantage du modèle semi-paramétrique est de pouvoir calculer les dérivées de la fonction de production ou de la fonction de coût. En plus de pouvoir estimer directement l'efficacité des DMU, nous pouvons également calculer les caractéristiques la technologie des entreprises, comme l'élasticité des inputs, la nature des rendements d'échelle, le taux de substitution entre les inputs, etc. Puisque notre modèle empirique repose sur la fonction de production et la fonction de coût, dans la section suivante, nous présentons les principales quantités qui ont un intérêt économique et que nous mesurerons.

8-3-1 Fonction de production

Efficacités technique

Comme l'approche paramétrique, l'approche semi-paramétrique nous permet d'obtenir pour chaque DMU, l'efficacité technique. Le coefficient d'efficacité est obtenu en calculant le rapport entre l'output efficient et l'output observé :

$$\theta = \sum_{i=1}^n \theta_i \times \left(y_i^{obs} / \sum_{j=1}^n y_j^{obs} \right), \quad \text{avec } \theta_i = \frac{y_i^{obs}}{\hat{y}_i}, \quad (8-9)$$

où θ mesure la moyenne des coefficients d'efficacité technique des firmes d'un secteur pondérés par la part de l'output de cette firme dans la somme des outputs de ce secteur, θ_i représente le coefficient d'efficacité technique de la DMU i , y_i^{obs} est l'output observé et \hat{y}_i est l'output estimé (l'exponentielle de $\ln y$ dans l'équation 8-1). Cela représente le pourcentage de l'augmentation des outputs que l'on peut réaliser sans ajouter la quantité d'input.

Élasticité des inputs

Étant donnée une fonction de production, l'élasticité de l'output par rapport à un input donné représente la variation en pourcentage de l'output suite à une variation de 1 % de cet input. Dans le cas de la fonction de production que nous utiliserons, ces élasticités sont données par : $\xi_p = \partial \ln y / \partial \ln x_p, \forall p ; \eta_s = \partial \ln y / \partial \ln k_s, \forall s$.

Rendements d'échelle

Une mesure très importante de la technologie des DMU est celle des rendements d'échelles. Les rendements d'échelle permettent, entre autres, de vérifier si l'échelle de production des DMU est optimale. Ils mesurent l'impact d'une hausse proportionnelle simultanée de tous les inputs sur l'output. Nous pouvons les mesurer sous la forme d'une élasticité. La mesure des rendements d'échelle est le rapport entre le taux de variation de l'output par rapport au taux de changement proportionnel des inputs. Il s'agit d'une mesure locale des rendements d'échelle, en un point donné (ou dans son voisinage) de la frontière de production. Dans le cas à un seul output avec un seul input, la productivité moyenne augmente (diminue) lorsque la technologie présente des rendements d'échelle croissants (décroissants). Dans le cas de multiples inputs et outputs, nous pouvons illustrer le concept de rendements d'échelles à l'aide d'une fonction de production. Supposons qu'un vecteur d'inputs variables

$\mathbf{x} = (x_1, x_2, \dots, x_p)$, un vecteur d'inputs quasi-fixes $\mathbf{k} = (k_1, k_2, \dots, k_s)$ et un d'output y correspondent à un point sur la frontière de production $y = f(\mathbf{x}, \mathbf{k})$. Le long de la

frontière nous avons $\frac{dy}{y} = \sum_{p=1}^P \frac{\partial f}{\partial x_p} \frac{x_p}{y} \frac{dx_p}{x_p} + \sum_{s=1}^S \frac{\partial f}{\partial k_s} \frac{k_s}{y} \frac{dk_s}{k_s}$. Supposons que tous les

inputs augmentent dans une même proportion β et notons le taux de croissance de

l'output α , alors $\alpha/\beta = \sum_{p=1}^P \frac{\partial f}{\partial x_p} \frac{x_p}{y} + \sum_{s=1}^S \frac{\partial f}{\partial k_s} \frac{k_s}{y} = \sum_{p=1}^P \xi_p + \sum_{s=1}^S \eta_s$ est une mesure locale

des rendements d'échelles.

Comme nous avons relâché l'hypothèse que les entreprises sont efficaces, cela signifie qu'il est possible que des entreprises fonctionnent sous de la frontière de production. Par contre, nous ne pouvons mesurer les rendements d'échelle que sur la frontière. Comme nous avons une frontière de production commune à toutes les entreprises, nous pouvons projeter toutes les entreprise sur la frontière et calculons les rendements d'échelle pour l'entreprises j , $j = 1, \dots, n$ notés η_j^E , :

$$\begin{aligned} \eta_j^E &= \sum_{p=1}^P \frac{\partial f}{\partial x_{pj}} \frac{x_{pj}}{y_j} + \sum_{s=1}^S \frac{\partial f}{\partial k_{sj}} \frac{k_{sj}}{y_j} \\ &= \sum_{p=1}^P \xi_{pj} + \sum_{s=1}^S \eta_{sj}, \forall j. \end{aligned} \tag{8-10}$$

8-3-2 Fonction de coût

Efficacités de coût

Comme le modèle de la fonction de production, le modèle de la fonction de coût nous a permis d'obtenir pour chaque DMU, le coût efficient. Le coefficient d'efficacité de coût est obtenu en faisant le rapport entre le coût observé et le coût efficient :

$$\phi = \sum_{i=1}^N \phi_i \times \left(C_i^{obs} / \sum_{j=1}^n C_j^{obs} \right) \text{ avec } \phi_i = \frac{C_i^{obs}}{\hat{C}_i}, \tag{8-11}$$

où ϕ mesure la moyenne de l'efficacité de l'ensemble de secteur pondérée par le coût, ϕ_i représente le coefficient d'efficacité de coût de la DMU i , C_i^{obs} est le coût observé et \hat{C}_i est le coût estimé.

Impact des inputs quasi-fixes sur les inputs variables

Nous pouvons mesurer l'impact d'une variation en pourcentage de la quantité d'un input quasi-fixe sur la part des dépenses en inputs variables. À partir de l'équation (8-6), les ajustements des équations de part des inputs variables peuvent être estimés par la formule suivante :

$$\tau_{S_p k_s} = \frac{\partial \ln S_p}{\partial \ln k_s} = \frac{\gamma_{ps}}{S_p}, \forall (p, s). \quad (8-12)$$

L'équation (8-12) est une mesure de la variation de la part des inputs variables dans le coût variable suite à une variation en pourcentage des inputs quasi-fixes. Les inputs quasi-fixes introduisent un biais en faveur de l'utilisation de l'input p si $\tau_{S_p k_s} > 0$, en défaveur de l'input p si $\tau_{S_p k_s} < 0$ et est neutre si $\tau_{S_p k_s} = 0$.

Récupération des rendements d'échelle à partir de la fonction de coût variable

Puisque dans notre modèle, nous avons des inputs variables, qui sont ajustables instantanément et des inputs quasi-fixes qui ne peuvent pas être ajustés à court terme, nous pouvons distinguer les mesures de rendements d'échelle de court terme, où seuls varient les inputs variables, que nous noterons η_{ct}^E , ainsi que la mesure de long terme, où les inputs quasi-fixes sont autorisés à varier, que nous noterons η_{lt}^E .

Sous l'hypothèse de minimisation des coûts variables, les conditions de premier ordre et le théorème de l'enveloppe permettent de déterminer que :

$$\frac{\partial f}{\partial x_p} = \frac{w_p}{\partial C / \partial y}, \forall p \Leftrightarrow \frac{\partial f}{\partial x_p} \frac{x_p}{y} = \frac{w_p x_p / C}{\frac{\partial C}{\partial y} \frac{y}{C}} = \frac{\xi_{Cw_p}}{\xi_{Cy}}, \forall p$$

$$\frac{\partial f}{\partial k_s} = -\frac{\partial C / \partial k_s}{\partial C / \partial y}, \forall s \Leftrightarrow \frac{\partial f}{\partial k_s} \frac{k_s}{y} = -\frac{\frac{\partial C}{\partial k_s} \frac{k_s}{C}}{\frac{\partial C}{\partial y} \frac{y}{C}} = -\frac{\xi_{Ck_s}}{\xi_{Cy}}, \forall s.$$

La substitution de ces résultats dans la définition des rendements d'échelle implique :

$$\eta_{ct}^E = \left(\frac{\partial \ln C}{\partial \ln y} \right)^{-1} \quad (8-13)$$

et

$$\eta_{lt}^E = \left(\frac{\partial \ln C}{\partial \ln y} \right)^{-1} \left(1 - \sum_{s=1}^S \xi_{Ck_s} \right) = \eta_{ct} \times \left(1 - \sum_{s=1}^S \xi_{Ck_s} \right). \quad (8-14)$$

8-3-3 Les mesures de la technologie dans le cas Translog

Dans cette section, nous présentons les diverses mesures présentées dans les sections 8-3-1 et 8-3-2 dans le cas de la forme Translog. Pour le cas Cobb-Douglas, il suffit de mettre à zéro tous les coefficients de deuxième ordre. Nous commençons par présenter les mesures de la technologie à partir de la fonction de production.

Les élasticités de l'output par rapport aux inputs variables (ξ_p) et quasi-fixes (η_s) sont données par les équations suivantes :

$$\xi_p = \frac{\partial \ln y}{\partial \ln x_p} = \alpha_{x_p} + \sum_{p'} \beta_{pp'} \ln x_{p'} + \sum_s \beta_{ps} \ln k_s \quad (8-15)$$

$$\eta_s = \frac{\partial \ln y}{\partial \ln k_s} = \alpha_{k_s} + \sum_{s'} \beta_{ss'} \ln k_{s'} + \sum_p \beta_{ps} \ln x_p, \quad (8-16)$$

avec $\mathbf{p} = (L, E, ENT, M)$ et $\mathbf{s} = (V, Inf)$. La mesure des rendements d'échelle est :

$$\eta^E = \sum_{p=1}^P \left[\alpha_{x_p} + \sum_{p'} \beta_{pp'} \ln x_{p'} + \sum_s \beta_{ps} \ln k_s \right] + \sum_{s=1}^S \left[\alpha_{k_s} + \sum_{s'} \beta_{ss'} \ln k_{s'} + \sum_p \beta_{ps} \ln x_p \right]. \quad (8-17)$$

Nous regardons maintenant la partie duale. Comme le cas primal, nous commençons par donner les formules pour calculer des élasticités. Premièrement, pour mesurer l'ampleur de la variation de l'output suite à une variation des inputs quasi-fixes, nous prenons l'équation suivante :

$$\xi_{yk_s} = \frac{\partial \ln y}{\partial \ln k_s} = -\frac{\xi_{Ck_s}}{\xi_{Cy}} = -\frac{\varphi_{k_s} + \sum_{p=1}^P \gamma_{ps} \ln w_p + \sum_{s'}^S \gamma_{ss'} \ln k_{s'} + \gamma_{yy} \ln y}{\varphi_y + \sum_{p=1}^P \gamma_{py} \ln w_p + \sum_s^S \gamma_{sy} \ln k_s + \gamma_{yy} \ln y}, \forall s. \quad (8-18)$$

Ensuite, l'impact des inputs quasis-fixes sur les inputs variables est donnée par²⁰ :

$$\xi_{ps} = \frac{\partial \ln x_p}{\partial \ln k_s} = \frac{\partial \ln S_p}{\partial \ln k_s} + \frac{\partial \ln C}{\partial \ln k_s} = \frac{\gamma_{ps}}{S_p} + \xi_{Ck_s} = \tau_{S_p k_s} + \xi_{Ck_s}, \forall p, \quad (8-19)$$

où le premier terme du membre de droite de l'équation (8-19) de cette expression est l'élasticité de la part du coût des inputs variables par rapport aux inputs quasi-fixes et le second terme est l'effet productif, représenté par le prix implicite des inputs quasi-fixes.

Finalement, nous présentons les formules pour calculer les rendements d'échelles. Pour la forme Translog, ils dépendent du niveau des inputs pour chaque entreprise, les rendements d'échelles peuvent s'écrire :

$$\eta_{ct}^E = \left(\frac{\partial \ln C}{\partial \ln y} \right)^{-1} = \left(\varphi_y + \sum_{p=1}^P \gamma_{py} \ln w_p + \sum_s^S \gamma_{sy} \ln k_s + \gamma_{yy} \ln y \right)^{-1}$$

$$\eta_{lt}^E = \left(\varphi_y + \sum_{p=1}^P \gamma_{py} \ln w_p + \sum_s^S \gamma_{sy} \ln k_s + \gamma_{yy} \ln y \right)^{-1} \left[1 - \sum_{s=1}^S \left(\varphi_{k_s} + \sum_{p=1}^P \gamma_{ps} \ln w_p + \sum_{s'=1}^S \gamma_{ss'} \ln k_{s'} + \gamma_{sy} \ln y \right) \right]. \quad (8-20)$$

Si $\eta_{ct}^E < 1$ (respectivement, >1), alors il faut réduire (augmenter) les inputs variables afin d'augmenter la productivité moyenne ou réduire (augmenter) le coût moyen. De même, si $\eta_{lt}^E < 1$ (respectivement, >1), alors il est préférable de réduire (augmenter)

²⁰ Nadiri et Mamouneas, 1994 sont des premiers qui utilisent cette formule pour mesurer cet impact.

les inputs variables et quasi-fixes. Lorsque $\eta_{it}^E = 1$, l'échelle de production est optimale²¹.

8-4 Résultats du modèle et analyse des résultats

Dans cette section, nous présenterons nos résultats d'estimation et les analyses pour obtenir une vue globale des résultats et la nature de la technologie du secteur du transport routier de voyageurs en France. Les points importants dans notre analyse seront l'efficacité des entreprises, la nature des rendements d'échelle, l'élasticité des facteurs de production et bien sûr l'influence des infrastructures sur les entreprises. Nous commençons par présenter les résultats pour la fonction de production. Ensuite nous analyserons les résultats de la fonction de coût et ferons une comparaison entre les deux.

Nos calculs sont faits à l'aide du logiciel **SAS**. Selon le critère de minimisation retenu (minimiser la somme des erreurs, ou minimiser la somme du carré des erreurs), nous avons utilisé respectivement la programmation linéaire (**Proc LP**) et la programmation non linéaire²² (**Proc NLP**). Pour lancer notre modèle, nous avons pris un sous-ensemble de données sur une période de cinq ans avec les réserves suivantes :

1. Puisque notre base de données ne porte que sur une courte période, qui est de cinq ans, nous nous permettons d'omettre le changement technologique dans ce secteur.
2. Nous n'avons que 2554 DMU au totale, soit environ 500 DMU par année, et les résultats obtenus par l'approche basée sur la recherche opérationnelle sont très sensibles à la taille de l'échantillon. Plus grande est la taille d'échantillon, plus précis est le résultat qu'on peut obtenir.

²¹ En ce qui concerne la taille optimale de l'entreprise, il faut comparer le prix de marché des inputs quasi-fixes avec leur prix implicite. Si ce dernier est inférieur (supérieur) au prix de marché, alors les DMU sont surcapitalisé (sous-capitalisé) en véhicules ou infrastructures. Mais nous ne disposons pas les informations sur le prix des infrastructures, donc nous ne pouvons pas faire une étude plus précise.

²² Soit la programmation quadratique.

Les intervalles de confiance pour les paramètres et les rendements d'échelles sont calculés. La méthode utilisée est celle du chapitre 4, section 4-4. Les simulations sont réalisées en utilisant un programme **SAS** élaboré par l'auteur à partir de 3000 tirages pour chaque modèle, $B = 3000$ La taille de sous-échantillonnage est de $0,8^{23}$.

8-4-1 Résultats de la fonction de production

8-4-1-1 Coefficients estimés et le résultat de simulation

Les coefficients estimés et le résultat de simulation pour la fonction de production sont présentés aux tableaux 8-2, pour la forme Cobb-Douglas, et 8-3, pour la forme Translog. La première colonne de ces tableaux donne la valeur estimée des paramètres, la seconde le biais estimé, la troisième la valeur estimée du paramètre corrigé pour le biais et enfin les six dernières colonnes des tableaux donnent de l'intervalle de confiance aux niveaux de 90 %, 95 % et 99 %. Dans chaque tableau, PL signifie que la méthode utilisée repose sur un programme linéaire et que le critère est la minimisation de la somme des erreurs; PQ signifie qu'on utilise la programmation quadratique et que le critère utilisé est la minimisation de la somme du carré des erreurs. Primal indique que la fonction de production a été utilisée et Dual est relié à la fonction de coût.

²³ Nous avons essayé d'autre taille (0,9, 0,75), les résultats ne dépendent pas beaucoup de la taille du sous-échantillon.

Tableau 8-2 : Résultats de simulation des coefficients pour la fonction de production sous forme Cobb-Douglas

Primal	Coefficient estimé	Biais d'estimation	Coefficient corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
				Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme									
linéaire									
α_L	0,441	0,016	0,426	0,250	0,520	0,221	0,542	0,141	0,573
α_E	0,231	-0,013	0,245	0,168	0,341	0,161	0,357	0,088	0,382
α_{ENT}	0,052	-0,023	0,075	0,043	0,129	0,032	0,136	0,028	0,140
α_M	0,154	-0,026	0,180	0,101	0,305	0,099	0,308	0,070	0,314
α_V	0,013	0,011	0,002	-0,061	0,025	-0,092	0,027	-0,161	0,029
α_{Inf}	0,669	-0,101	0,771	0,627	1,174	0,617	1,247	0,604	1,339
Programme									
quadratique									
α_L	0,434	0,018	0,416	0,227	0,509	0,201	0,520	0,168	0,546
α_E	0,275	-0,055	0,330	0,243	0,426	0,237	0,437	0,158	0,470
α_{ENT}	0,042	-0,011	0,133	0,037	0,135	0,025	0,193	0,015	0,239
α_M	0,160	-0,029	0,188	0,117	0,305	0,114	0,318	0,093	0,319
α_V	0,012	0,013	-0,001	-0,068	0,024	-0,083	0,028	-0,141	0,031
α_{Inf}	0,639	-0,022	0,661	0,624	0,947	0,578	1,106	0,556	1,253

La première colonne représente les paramètres obtenus avec l'échantillon original ; la deuxième colonne est l'estimateur *bootstrap* du biais ; la troisième colonne correspond à l'estimateur corrigé et finalement $I_{\alpha\%}$ nous donne le seuil de l'intervalle de confiance pour chaque estimateur au niveau α avec $\alpha = 90\%$, 95% , 99% .

Tableau 8-3 : Résultats de simulation des coefficients pour la fonction de production sous forme Translog

Primal	Coefficient estimé	Biais d'estimation	Coefficient corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
				Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme									
linéaire									
α_L	0,903	-0,068	0,970	0,441	1,576	0,354	1,671	0,186	1,758
α_E	0,012	0,059	-0,047	-0,326	0,146	-0,388	0,151	-0,507	0,157
α_{ENT}	-0,038	0,040	-0,077	-0,225	0,014	-0,250	0,030	-0,322	0,140
α_M	-0,042	0,039	-0,081	-0,375	0,056	-0,439	0,083	-0,559	0,237
α_V	-0,021	-0,038	0,017	-0,042	0,203	-0,042	0,256	-0,281	0,393
α_{inf}	2,264	-0,324	2,588	2,010	3,379	1,902	3,466	1,738	3,618
β_{LL}	0,251	-0,026	0,277	0,198	0,419	0,191	0,454	0,167	0,520
β_{EE}	0,063	-0,016	0,079	0,041	0,136	0,033	0,145	0,018	0,163
β_{ENTENT}	0,003	-0,003	0,006	-0,002	0,028	-0,028	0,034	-0,036	0,048
β_{MM}	-0,002	0,000	-0,002	-0,012	0,014	-0,014	0,018	-0,018	0,026
β_{VV}	0,007	0,001	0,005	-0,009	0,014	-0,014	0,016	-0,024	0,018
β_{infInf}	-0,406	0,065	-0,472	-0,626	-0,227	-0,669	-0,202	-0,751	-0,177
β_{L_E}	-0,094	0,013	-0,107	-0,170	-0,067	-0,181	-0,057	-0,191	-0,044
$\beta_{L_{ENT}}$	-0,036	0,001	-0,037	-0,072	-0,008	-0,072	0,019	-0,075	0,038
β_{L_M}	-0,017	0,000	-0,017	-0,035	0,016	-0,040	0,020	-0,058	0,032
β_{L_V}	0,005	0,000	0,005	-0,013	0,018	-0,021	0,026	-0,044	0,036

β_{L_Inf}	-0,086	0,016	-0,102	-0,174	-0,045	-0,201	-0,031	-0,248	0,016
β_{E_ENT}	0,015	0,002	0,012	-0,007	0,029	-0,010	0,030	-0,015	0,030
β_{E_M}	0,006	0,000	0,006	-0,005	0,013	-0,007	0,014	-0,012	0,021
β_{E_V}	-0,009	0,006	-0,015	-0,026	-0,007	-0,029	-0,005	-0,035	-0,001
β_{E_INF}	0,028	-0,022	0,050	0,020	0,103	0,014	0,117	-0,005	0,143
β_{ENT_M}	0,005	-0,001	0,005	-0,004	0,009	-0,007	0,009	-0,013	0,009
β_{ENT_V}	-0,006	-0,002	-0,004	-0,012	0,016	-0,012	0,019	-0,013	0,031
β_{ENT_Inf}	0,028	-0,003	0,031	0,007	0,055	0,003	0,056	-0,013	0,061
β_{M_V}	0,002	-0,001	0,003	0,001	0,018	0,000	0,022	-0,003	0,035
β_{M_Inf}	0,014	-0,001	0,015	-0,017	0,040	-0,023	0,050	-0,044	0,072
β_{V_Inf}	0,007	0,007	0,000	-0,042	0,014	-0,052	0,014	-0,084	0,070

Programme
quadratique

α_L	0,901	0,035	0,866	0,389	1,373	0,333	1,460	0,237	1,594
α_E	0,007	0,003	0,004	-0,399	0,208	-0,495	0,232	-0,607	0,269
α_{ENT}	-0,009	0,027	-0,035	-0,161	0,091	-0,186	0,176	-0,227	0,233
α_M	-0,042	0,058	-0,100	-0,396	0,066	-0,433	0,114	-0,490	0,202
α_V	-0,023	0,041	-0,063	-0,045	0,260	-0,088	0,303	-0,342	0,380
α_{Inf}	2,241	-0,142	2,384	1,807	2,578	1,679	2,709	1,475	2,901
β_{LL}	0,251	-0,072	0,323	0,245	0,446	0,240	0,482	0,218	0,547
β_{EE}	0,063	-0,008	0,071	0,036	0,139	0,032	0,145	0,018	0,160
β_{ENTENT}	0,004	-0,004	0,008	-0,008	0,024	-0,025	0,034	-0,030	0,042

β_{MM}	-0,002	0,000	-0,002	-0,013	0,014	-0,015	0,016	-0,019	0,020
β_{VV}	0,007	0,006	0,001	-0,013	0,014	-0,016	0,014	-0,023	0,014
β_{InfInf}	-0,398	0,013	-0,411	-0,583	-0,209	-0,616	-0,164	-0,668	-0,132
β_{L_E}	-0,094	0,006	-0,100	-0,164	-0,061	-0,176	-0,053	-0,196	-0,033
$\beta_{L_{ENT}}$	-0,036	0,007	-0,043	-0,073	0,002	-0,075	0,016	-0,078	0,034
β_{L_M}	-0,017	0,004	-0,021	-0,043	0,021	-0,048	0,023	-0,055	0,031
β_{L_V}	0,005	0,001	0,004	-0,021	0,023	-0,028	0,027	-0,045	0,035
$\beta_{L_{Inf}}$	-0,085	0,013	-0,098	-0,203	-0,034	-0,234	-0,021	-0,283	0,002
$\beta_{E_{ENT}}$	0,016	0,002	0,014	-0,003	0,031	-0,006	0,033	-0,014	0,034
β_{E_M}	0,006	0,000	0,006	-0,008	0,014	-0,010	0,016	-0,015	0,023
β_{E_V}	-0,010	0,008	-0,017	-0,028	-0,007	-0,031	-0,005	-0,045	0,000
$\beta_{E_{INF}}$	0,029	-0,020	0,049	-0,003	0,090	-0,032	0,119	-0,040	0,122
β_{ENT_M}	0,005	-0,001	0,005	-0,002	0,009	-0,003	0,009	-0,007	0,012
β_{ENT_V}	-0,006	-0,006	0,000	-0,012	0,019	-0,012	0,023	-0,015	0,030
$\beta_{ENT_{Inf}}$	0,021	0,003	0,018	-0,005	0,043	-0,014	0,049	-0,031	0,053
β_{M_V}	0,003	-0,006	0,008	0,001	0,026	0,000	0,028	-0,002	0,035
$\beta_{M_{Inf}}$	0,014	-0,002	0,017	-0,018	0,048	-0,026	0,055	-0,040	0,064
$\beta_{V_{Inf}}$	0,007	0,018	-0,011	-0,059	0,015	-0,066	0,030	-0,086	0,080

La première colonne représente les paramètres obtenus avec l'échantillon original ; la deuxième est l'estimateur *bootstrap* du biais ; la troisième colonne correspond à l'estimateur corrigé et finalement $I_{\alpha\%}$ nous donne l'intervalle de confiance pour chaque estimateur au niveau α avec $\alpha = 90 \%$, 95% , 99% .

Premièrement, pour le modèle Cobb-Douglas, tous les paramètres sont significatifs pour tous les niveaux de confiance à l'exception du paramètre α_V pour lequel nous ne sommes pas en mesure de rejeter que le paramètre est significativement différent de zéro. C'est-à-dire que tous les autres facteurs contribuent à la production mais l'influence des véhicules sur les kilomètres parcourus n'est pas claire. Avec la correction pour le biais, les infrastructures restent le facteur qui contribue le plus à la production, et parmi les inputs variables, c'est la main-d'œuvre qui est le facteur le plus important. Deuxièmement, à partir du tableau 8-3 pour la forme Translog, nous constatons que pour plusieurs paramètres nous ne sommes pas en mesure de dire que les paramètres sont significativement différents de zéro pour tous les niveaux de confiance. Par contre, pour une variable donnée, nous pouvons toujours trouver un terme (ordre un, au carré ou un terme croisé avec une autre variable) significativement différent de zéro, ce qui est assez fréquent avec les estimations Translog. Par exemple, pour les paramètres concernant la variable véhicule : la variable d'ordre un, α_V , d'ordre deux, β_{VV} , et les termes croisés, β_{L_V} , β_{ENT_V} et β_{V_Inf} sont tous significativement différents de zéro; par contre, les termes croisés β_{E_V} , β_{M_V} sont significativement différent de zéro seulement aux seuils de 90 % et de 95 %. Donc nous ne pouvons exclure, à partir des résultats de la forme Translog que les véhicules contribuent à la production des entreprises.

Elasticité des inputs

Pour la fonction de production en forme Cobb-Douglas, les coefficients obtenus représentent l'élasticité de l'output par rapport aux facteurs et nous allons les regarder plus tard. Les termes au carré dans la fonction de production Translog ont une importance particulière. Regardons les formules suivantes :

Par définition d'une élasticité, nous avons :

$$\frac{\partial y}{\partial x_p} = \xi_p \times \frac{y}{x_p} . \quad (8-21)$$

Donc pour connaître l'impact des inputs sur la productivité marginale, nous pouvons chercher la dérivée seconde de l'output par rapport à cet input, soit :

$$\xi_{pp} = \frac{\partial^2 y}{\partial x_p \partial x_p} = \frac{\partial \left(\xi_p \times \frac{y}{x_p} \right)}{\partial x_p} = \frac{\partial \xi_p}{\partial x_p} \times \frac{y}{x_p} - \xi_p \times \frac{y}{x_p^2} = \frac{\beta_{pp}}{x_p} \times \frac{y}{x_p} - \xi_p \times \frac{y}{x_p^2} = (\beta_{pp} - \xi_p) \frac{y}{x_p^2}, \forall p. \quad (8-22)$$

De même, pour les inputs quasi-fixes :

$$\eta_{ss} = \frac{\partial^2 y}{\partial k_s \partial k_s} = \frac{\partial \left(\eta_s \times \frac{y}{k_s} \right)}{\partial k_s} = \frac{\partial \eta_s}{\partial k_s} \times \frac{y}{k_s} - \eta_s \times \frac{y}{k_s^2} = \frac{\beta_{ss}}{k_s} \times \frac{y}{k_s} - \eta_s \times \frac{y}{k_s^2} = (\beta_{ss} - \eta_s) \frac{y}{k_s^2}, \forall s. \quad (8-23)$$

Par exemple, un signe positif pour ξ_{LL} implique que l'ajout d'une unité de main d'œuvre augmentera sa productivité. Puisque y, x_p, k_s sont toujours positifs, le signe de ξ_{pp} et celui de η_{ss} seront entièrement déterminées par le signe de $(\beta_{pp} - \xi_p)$ et $(\beta_{ss} - \eta_s)$. Le tableau 8-4 nous donne l'élasticité de l'output par rapport aux inputs

Tableau 8-4 : Élasticité de l'output par rapport aux inputs

	ξ_L	ξ_E	ξ_{ENT}	ξ_M	η_V	η_{Inf}
Cobb-Douglas PL	0,441	0,231	0,052	0,154	0,013	0,669
Cobb-Douglas PQ	0,434	0,275	0,042	0,160	0,012	0,638
Translog PL	0,613	0,187	0,052	0,024	0,014	0,496
Translog PQ	0,609	0,188	0,057	0,025	0,015	0,442

$\xi_p = \partial \ln y / \partial \ln x_p, p = L, E, ENT, M$ est l'élasticité de l'output par rapport à l'input variable x_p ;
 $\eta_s = \partial \ln y / \partial \ln k_s, s = V, Inf$ est l'élasticité de l'output par rapport à l'input quasi-fixe k_s .

Comme nous le voyons au tableau 8-3, les termes au carré des facteurs sont toujours inférieurs à leur élasticité ($\beta_{pp} < \xi_p, \beta_{ss} < \eta_s$), donc les signes de ξ_{pp} et η_{ss} seront toujours négatifs. Cela correspond au résultat de la forme Cobb-Douglas, avec $\xi_p = \alpha_p (\beta_{pp} - 1) y / x_p^2$ et $\eta_s = \alpha_s (\beta_{ss} - 1) y / k_s^2$, qui sont également négatifs. Cela indique que tous les inputs montrent une productivité marginale décroissante.

Nous constatons que l'influence des différents critères d'optimisation est négligeable. Les coefficients estimés par une même forme fonctionnelle sont très proches. Par contre, entre la forme Cobb-Douglas et la forme Translog, il y a une différence assez claire. Cela signifie que le critère d'optimisation (minimiser la somme des erreurs ou minimiser la somme de carré des erreurs) n'est pas important, mais qu'une forme plus flexible nous donne des résultats différents. Ensuite, nous trouvons que l'élasticité de la main-d'œuvre est toujours plus grande que celles des autres inputs variables. Ceci signifie que parmi les inputs qui peuvent être contrôlés par les firmes, la main d'œuvre reste le facteur le plus important. Une augmentation de 1% de la main d'œuvre permet aux firmes d'augmenter de 0,438% l'output pour la forme Cobb-Douglas, ou 0,611% pour la forme Translog²⁴. Ensuite, l'énergie joue un rôle significatif pour les firmes de transport. Bien que l'élasticité de l'énergie soit moins grande que celle de main d'œuvre, elle reste relativement importante par rapport à celles de l'entretien et des matériaux. La dépense d'énergie est toujours importante dans ce secteur. Par contre, l'élasticité des véhicules est assez basse. Pour la forme Cobb-Douglas, ce paramètre n'est pas significatif ; pour la forme Translog, on trouve qu'une élasticité de 0,0145. Cela signifie que le nombre de véhicules n'influence pas trop l'output des entreprises. Une explication potentielle est que pour les entreprises dans ce secteur, le véhicule est très rarement plein. Puisque ces entreprises ont signé un contrat avec le gouvernement, le but est de satisfaire la demande, donc l'offre est supérieure à la demande, il y a surcapacité. Le fait d'augmenter ou diminuer le nombre de véhicules n'a pas un impact important sur l'output des entreprises. C'est plutôt la capacité de bien utiliser les véhicules existants qui va influencer les performances des entreprises (remplir les véhicules vides, ajouter les chauffeurs pour faire rouler les véhicules sur une plus longue distance, etc.). Pour les infrastructures, le résultat de Cobb-Douglas nous donne une élasticité de 0,654, qui est plus important

²⁴ Ci-après, quand nous parlons du résultat Cobb-Douglas, ou du résultat translog, nous voulons dire qu'il s'agit de la moyenne des résultats obtenus par une même forme fonctionnelle, mais avec différents critères d'optimisation.

que le résultat avec la Translog, avec une élasticité de 0,469. Ceci signifie qu'une augmentation de 1% des infrastructures à long terme peut permettre aux firmes d'avoir des augmentations d'output de 0,654% et 0,469 % avec les formes Cobb-Douglas et Translog respectivement. Selon le résultat avec la forme Cobb-Douglas, la contribution des infrastructures au processus de production est plus importante, ce qui est peu crédible. Par contre, le résultat avec la forme Translog montre que la contribution des infrastructures est moindre que la main-d'œuvre, mais reste beaucoup plus importante que les véhicules. Tout simplement, dans ce secteur, les entreprises ont besoin des infrastructures routières pour aller chercher les clients à leur lieu d'origine et les déplacer vers leur destination. Les entreprises ont également besoin de chauffeurs pour faire rouler des véhicules et des personnels administratifs pour faire fonctionner l'entreprise. Comme nous le disions plus haut, le but des entreprises dans ce secteur est d'assurer le déplacement des individus, leur véhicules sont souvent non remplis, donc une route moins bouchonnée ou plus de chauffeurs permettront aux entreprises d'offrir plus d'output pour une quantité d'inputs donnée, donc ces facteurs contribuent plus que les véhicules dans le processus de production.

Les rendements d'échelle

Les rendements d'échelles sont mesurés par l'équation (8-17) pour la fonction de production et par l'équation (8-20) pour la fonction de coût. Bien entendu, pour la forme Cobb-Douglas, les rendements d'échelle sont les mêmes pour toutes les entreprises, tandis que pour la forme Translog, ils sont spécifiques à chacune des entreprises. Pour la forme Cobb-Douglas, nous présenterons directement les résultats de simulation des rendements d'échelle ; pour la forme Translog, nous projeterons toutes les entreprises sur la frontière estimée et calculerons les rendements d'échelles de chacune d'entre elles. Ces résultats sont présentés via histogramme et une densité estimée des rendements. Nous présenterons également la description statistique de nos résultats comme la médiane, le minimum, le maximum, les quartiles, le pourcentage

des entreprises avec une élasticité supérieure à un et le résultat de simulation de toutes ces statistiques.

Les histogrammes sont présentés aux Figure 8-1 à 8-2. Pour les rendements d'échelle en maintenant fixes les infrastructures, 16,9 % des entreprises présentent des rendements d'échelle supérieurs à 1 pour le programme de minimisation la somme des erreurs (PL) et 16,2 % pour le programme de minimisation la somme du carrés des erreurs (PQ). Ça signifie qu'à court terme, la plupart des entreprises montre des rendements d'échelle décroissants et ce résultat est également justifié par le quantile à 75% des rendements d'échelle de ces deux programmes de minimisations (0,977 pour PL et 0,975 pour PQ). Quand les infrastructures peuvent varier, les entreprises montrent des rendements d'échelles supérieures à 1. Ça signifie qu'à long terme, les entreprises montrent des rendements d'échelle croissants, ce qui est aussi vérifié avec un quantile d'ordre 75% de 1,332 pour PL et 1,337 pour PQ. Les quantiles de 25 % sont de 1,193 pour PL et 1,195 pour PQ.

Les résultats des simulations sont donnés au tableau 8-5. Nous noterons $\eta_{sans}^E = \xi_L + \xi_E + \xi_{ENT} + \xi_M + \eta_V$, les rendements d'échelle pour un niveau d'infrastructures donné, soit les rendements d'échelle sans les infrastructures et $\eta_{avec}^E = \eta_{sans}^E + \eta_{inf}$ correspond aux rendements d'échelle à long terme alors que les infrastructures deviennent variables. Regardons quand les infrastructures sont maintenues constantes. Pour la forme Cobb-Douglas, il n'y a qu'un pourcent de chance que nous ne pouvons rejeter l'hypothèse que $\eta_{sans}^E = 1$. Tous les autres intervalles ne contiennent pas un, la probabilité de taille optimale est assez faible. Donc on rejette plutôt que les rendements sont constants Pour la forme Translog, au contraire du résultat de Cobb-Douglas, même au niveau de confiance 95 %, nous ne pouvons pas refuser l'hypothèse que les entreprises fonctionnent à taille optimale.

Par contre, quand les infrastructures sont considérées variables, c'est-à-dire quand nous incluons l'influence des infrastructures, les rendements d'échelle η_{avec}^E

(identiques aux rendements d'échelles à long terme, soit η_{it}^E) deviennent significativement supérieurs à 1 en toutes circonstances. Cela indique que la taille des entreprises n'est pas optimale, et le gouvernement devrait investir dans les infrastructures et inciter une augmentation de la taille des entreprises afin de réduire le coût moyen. En même temps, le gouvernement doit prendre en considération la capacité du marché à absorber ce genre d'augmentation. Si la demande de marché est capable de satisfaire l'augmentation de l'offre dans ce secteur, nous pouvons bien sûr suivre ce chemin pour augmenter la performance des entreprises. Donc une réallocation des ressources (par exemple, des fusions d'entreprises) pourrait être bénéfique. Cependant, il ne faut pas oublier que les infrastructures ne se manipulent pas comme n'importe quel facteur pour les entreprises. Une fois qu'elles ont choisi leur localisation géographique, il est très difficile de se déplacer d'ailleurs. Il faut que les entreprises choisissent bien leur localisation avant qu'elles ne s'installent quelque part.

Pour conclure, étant donné les réseaux routiers en place, nous ne pouvons pas refuser l'hypothèse que la taille des entreprises est pour une grande majorité optimale à court terme avec la forme fonctionnelle Translog, par contre ce résultat ne semble plus aussi évident avec la forme fonctionnelle Cobb Douglas. Par contre, à long terme, quand les infrastructures ne sont plus fixes, nous voyons clairement que ce sont les infrastructures qui rendent des rendements d'échelle croissants pour l'ensemble du secteur. Une façon d'interpréter ces rendements d'échelle croissants est que les infrastructures ne sont pas suffisantes pour les entreprises et que la taille des entreprises est trop petite, ou il y a de trop nombreuses PME ou encore une combinaison des deux. Pour tirer parti des rendements d'échelle, il faut fusionner les PME ou il faut augmenter la taille des entreprises et le gouvernement doit investir sur les réseaux routiers.

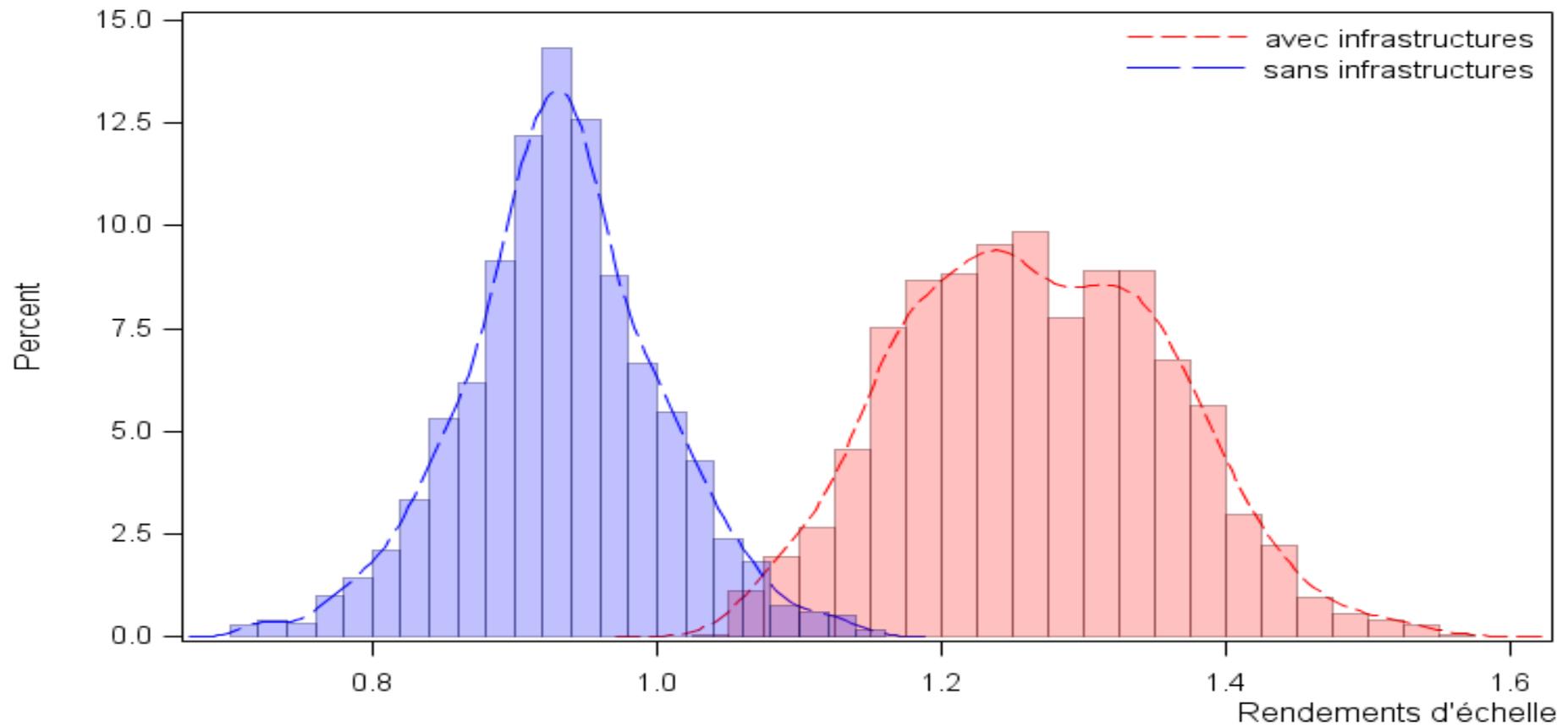


Figure 8-1 : Histogramme des rendements d'échelles avec et sans les infrastructures pour la fonction de production Translog avec le critère de minimisation de la somme des erreurs (cas PL)

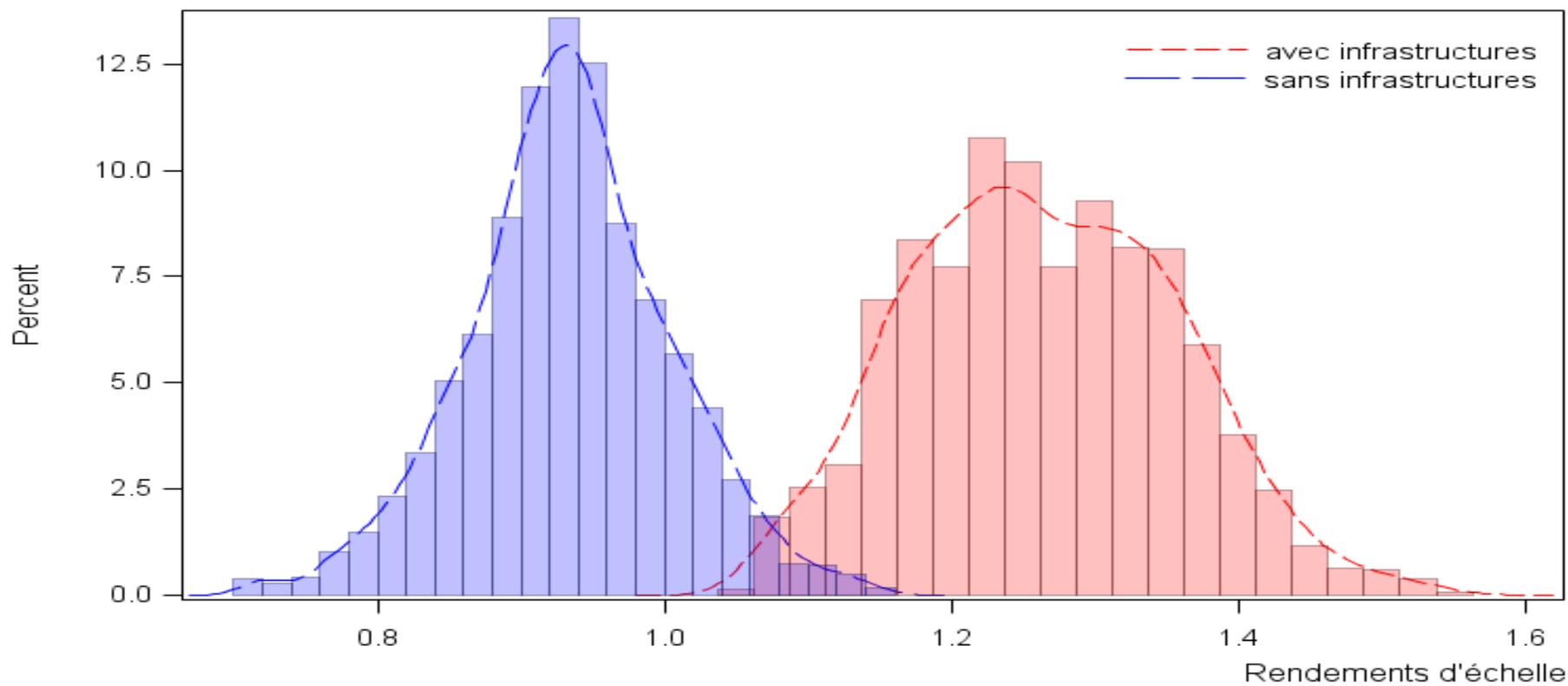


Figure 8-2 : Histogramme des rendements d'échelles avec et sans les infrastructures, pour la fonction de production Translog avec le critère de minimisation de la somme du carré des erreurs (cas PQ)

Tableau 8-5 : Résultats de simulation pour les rendements d'échelles en utilisant une fonction de production

Cobb-Douglas		Coefficient estimé	Biais d'estimateur	Coefficient corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
					Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme linéaire	η_{avec}^E	1,560	-0,188	1,748	1,508	1,884	1,501	1,98	1,491	2,086
	η_{sans}^E	0,891	-0,091	0,982	0,748	0,996	0,739	0,999	0,662	1,085
Programme quadratique	η_{avec}^E	1,561	-0,036	1,597	1,122	1,733	1,094	1,765	1,080	1,839
	η_{sans}^E	0,923	-0,02	0,921	0,799	0,935	0,775	0,963	0,744	1,020

Translog		Coefficient estimé	Biais d'estimateur	Coefficient corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
					Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme linéaire	Médiane	1,258	-0,014	1,272	1,223	1,348	1,212	1,360	1,181	1,385
	Min	1,038	-0,042	1,080	1,026	1,110	1,018	1,179	0,977	1,220
	η_{avec}^E Max	1,562	0,119	1,443	1,234	1,663	1,223	1,700	1,216	1,752
	P25	1,193	-0,021	1,214	1,163	1,286	1,146	1,302	1,120	1,328
	P75	1,332	-0,015	1,347	1,257	1,447	1,239	1,461	1,190	1,479
Programme linéaire	η_{sans}^E Médiane	0,932	0,007	0,925	0,847	0,989	0,843	1,026	0,827	1,080
	Min	0,707	-0,061	0,768	0,598	0,939	0,579	0,947	0,528	0,956
	Max	1,156	0,044	1,112	0,813	1,278	0,801	1,321	0,752	1,351
	P25	0,890	0,002	0,888	0,808	0,955	0,804	0,969	0,792	1,033

	P75	0,977	0,013	0,964	0,830	1,069	0,828	1,090	0,811	1,124
Programme	Médiane	1,261	-0,019	1,280	1,233	1,353	1,228	1,360	1,207	1,380
quadratique	Min	1,033	-0,067	1,100	1,025	1,098	1,016	1,100	0,989	1,101
η_{avec}^E	Max	1,560	0,185	1,375	1,223	1,628	1,217	1,662	1,212	1,702
	P25	1,195	-0,041	1,236	1,178	1,294	1,170	1,302	1,146	1,315
	P75	1,337	-0,009	1,346	1,273	1,446	1,264	1,454	1,233	1,467
Programme	Médiane	0,931	-0,019	0,950	0,888	1,019	0,882	1,028	0,867	1,063
quadratique	Min	0,709	-0,012	0,721	0,559	0,851	0,545	0,873	0,529	0,910
η_{sans}^E	Max	1,150	-0,072	1,202	1,074	1,301	1,035	1,324	0,921	1,351
	P25	0,890	-0,009	0,899	0,822	0,963	0,817	0,977	0,805	1,010
	P75	0,975	-0,031	1,006	0,940	1,077	0,924	1,095	0,887	1,120

La première colonne représente les paramètres obtenus avec l'échantillon original ; $biais_B[\hat{\alpha}_i]$ est l'estimateur *bootstrap* du biais ; la troisième colonne correspond à l'estimateur corrigé et finalement $I_{\alpha\%}$ nous donne le seuil de l'intervalle de confiance pour chaque estimateur au niveau α avec $\alpha = 90\%, 95\%, 99\%$.

Efficacité technique

Le calcul de la fonction de production nous a permis d'obtenir l'efficacité technique mesurée par θ_i pour chacune des DMU i . Le coefficient d'efficacité est obtenu en faisant le rapport entre l'output efficient et l'output observé. Nous présentons au tableau 8-6 les résultats pour les mesures d'efficacité technique. Nous avons calculé l'efficacité moyenne pondérée pour l'ensemble de l'échantillon (équation 8-9). Nous avons sélectionné les plus grandes et les plus petites entreprises et nous avons reporté leur efficacité pour les deux formes fonctionnelles, selon le critère d'optimisation.

Tableau 8-6 : Efficacité technique, modèle semi-paramétrique, primal

	Cobb-Douglas		Translog	
	PL	PQ	PL	PQ
θ	0,831	0,835	0,853	0,859
\dot{N}		92		54
θ 5 % plus petites	0,724	0,727	0,749	0,755
θ 5 % plus grandes	0,865	0,868	0,858	0,864

PL : programme linéaire de minimisation de la somme des erreurs, PQ : programme quadratique de minimisation de la somme du carré des erreurs. θ correspond au coefficient d'efficacité moyenne pondérée de secteur. \dot{N} représente la différence en valeur absolue du classement des DMU selon leur efficacité estimée par PL ou PQ.

Nous remarquons qu'à l'instar des rendements d'échelle, les différents critères d'optimisation nous donnent le même résultat. Par contre, la forme fonctionnelle Translog nous donne une efficacité légèrement plus haute que la forme de Cobb-Douglas. Cela découle du fait que la forme Translog est beaucoup plus souple que la forme Cobb-Douglas et la frontière estimée par une fonction de production de forme Translog enveloppe les données de plus près que la forme Cobb-Douglas. Comme la moyenne d'efficacité technique de ce secteur est d'environ 0,83 avec la forme Cobb-Douglas et d'environ 0,86 avec la forme Translog, cela signifie que nous pouvons diminuer de 14 à 17 % les quantités d'inputs sans affecter la quantité d'output. Ce résultat est assez proche de ce qu'on a obtenu avec la méthode DEA (avec un

coefficient d'efficacité 0,8). Malgré une certaine tendance à la monopolisation locale par des entreprises subventionnées par l'État, le secteur du transport routier est relativement efficace contrairement à l'Allemagne et la Norvège. Hirschhausen et Cullmann (2008) ont obtenu un coefficient d'efficacité 0,42 pour le secteur allemand et Odeck et Alkadi (2001) ont obtenu une efficacité de 72 % pour le secteur norvégien.

Nous avons calculé la moyenne de la différence en valeur absolue du classement des DMU selon leur efficacité estimée par la programmation linéaire et quadratique (Variable \dot{N} dans le tableau 8-6). La différence n'est pas significative (pour 2554 DMU au totale). Encore une fois, le critère d'optimisation n'influence pas beaucoup le classement des entreprises. Cela confirme que la programmation linéaire et la programmation quadratique donnent *grosso modo* les mêmes résultats.

Finalement, nous considérons l'impact de la taille des entreprises sur la performance. Les figures 8-3 à 8-6 représentent la répartition des efficacités techniques des DMU en fonction de leur taille. L'axe horizontal représente le classement des entreprises selon leur chiffre d'affaires (les plus petites à gauche et les plus grandes à droite), l'axe vertical représente le coefficient d'efficacité des firmes.

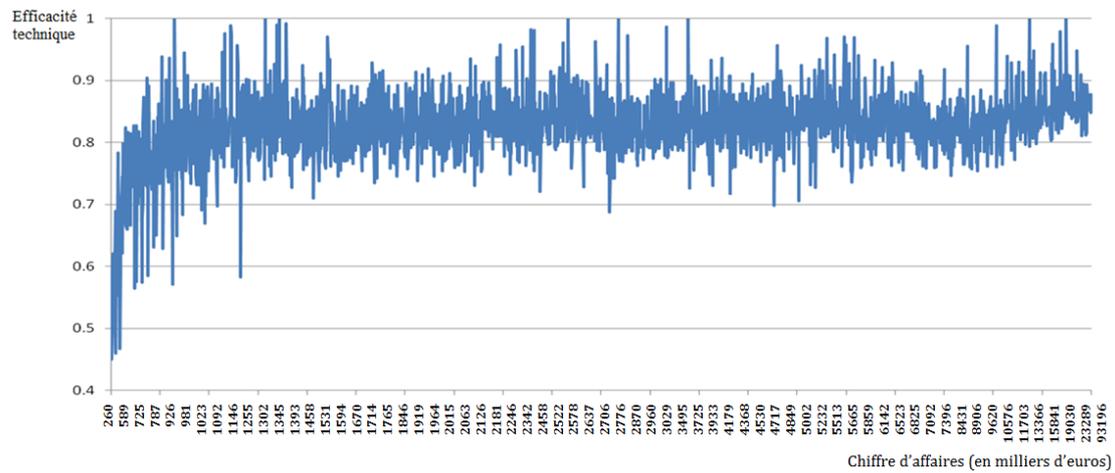


Figure 8-3 : Relation entre l'efficacité et le chiffre d'affaires des firmes en utilisant une fonction de production Cobb-Douglas et en minimisant la somme des erreurs (cas PL)

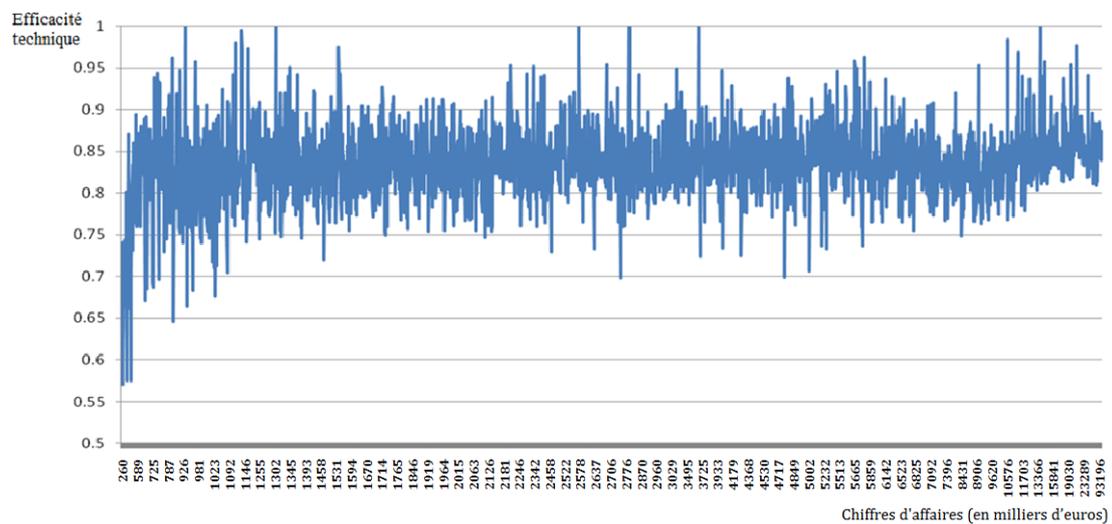


Figure 8-4 : Relation entre l'efficacité et le chiffre d'affaires des firmes en utilisant une fonction de production Cobb-Douglas et en minimisant la somme du carré des erreurs (cas PQ)

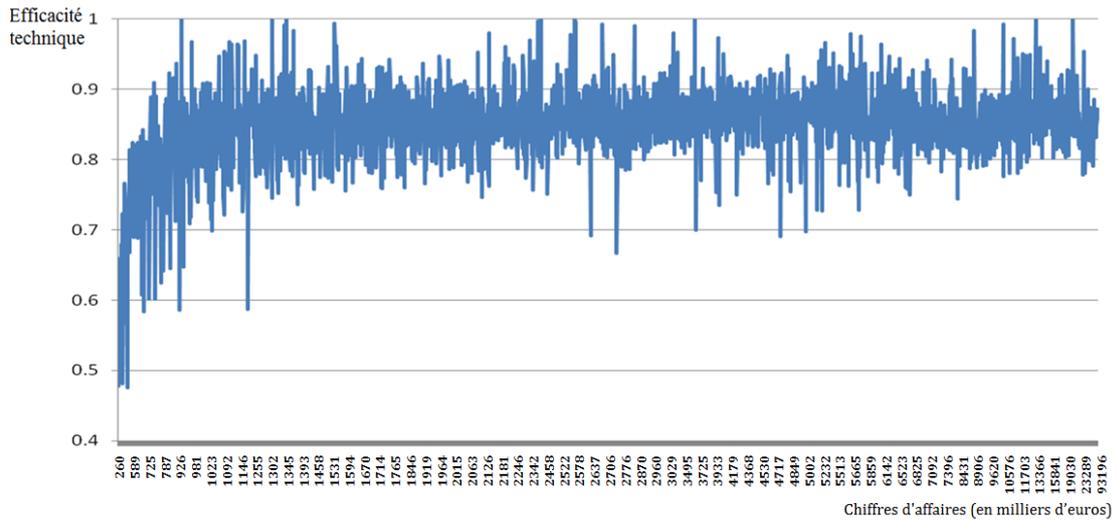


Figure 8-5 : Relation entre l'efficacité et le chiffre d'affaire des firmes en utilisant une fonction de production Translog et en minimisant la somme des erreurs (cas PL)

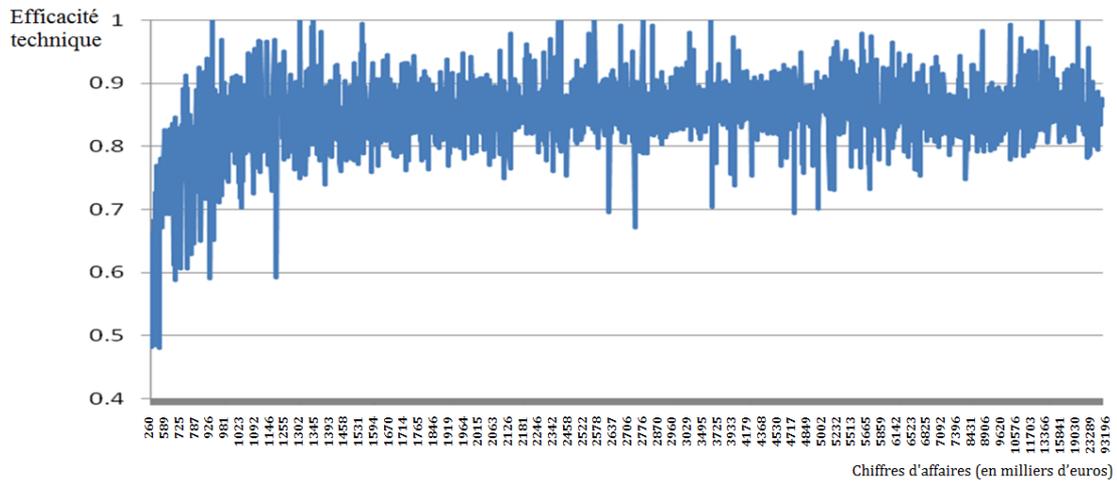


Figure 8-6 : Relation entre l'efficacité et le chiffre d'affaire des firmes en utilisant une fonction de production Translog et en minimisant la somme du carré des erreurs (cas PQ)

Sur l'axe horizontal, la distance d'échelle est 127, qui correspond au 5 % du nombre de DMU. Sur l'axe vertical, nous notons le coefficient d'efficacité des entreprises. Nous pouvons trouver une tendance assez claire : les très petites entreprises (les 5 % plus petites) sont d'une efficacité nettement inférieure à celle des plus grandes. Nous avons calculé la moyenne de θ pour les cinq premiers percentiles (les plus petites entreprises) et les cinq derniers percentiles (les plus grandes entreprises) avec comme critère de taille le chiffre d'affaires (Tableau 8-6). Pour les très petites DMU, la moyenne d'efficacité technique est beaucoup plus basse que la moyenne de toutes les DMU. Cela indique clairement une inefficacité pour les très petites entreprises. Pour les plus grands entreprises, cette différence n'est pas grande (4 % de plus que la moyenne et 2 % de plus que la moyenne respectivement pour PL et PQ), leur efficacité technique est très proche de la moyenne du secteur. Donc à l'exclusion des très petites DMU qui sont défavorisées, la taille ne semble pas avoir une influence positive sur la performance des entreprises. Cela peut provenir du fait que pour les petites, d'une part la forme Translog n'approxime pas correctement la technologie dans cette section du domaine. La forme peut être flexible, mais cette caractéristique est limitée localement puisqu'il s'agit d'une approximation de Taylor. D'autre part, les petites entreprises ne peuvent pas profiter des rendements d'échelle et c'est tout le processus de production qui en est affecté.

Maintenant, nous pouvons regarder la performance des DMU selon différentes régions. Nous calculons la moyenne de coefficient d'efficacité technique pondéré par l'output régional observé :

$$\theta_j = \sum_i \left(\theta_{ij} \times \frac{y_{ij}^{obs}}{\sum_j y_{ij}^{obs}} \right), \quad (8-24)$$

où θ_j représente l'efficacité technique de la région j , et θ_{ij} est l'efficacité technique calculé par notre modèle pour l'entreprise i qui se trouve dans la région j . Le tableau

8-7 suivant présente la moyenne de l'efficacité technique pondérée sur la base de l'output pour la forme Cobb-Douglas et la forme Translog.

Tableau 8-7 : Coefficients d'efficacité par région, modèle semi-paramétrique, primal

Région	Nombre d'observations	Moyenne Cobb-Douglas	Moyenne Translog
Franche Comté	79	0,841	0,862
Picardie	79	0,840	0,858
Alsace	68	0,839	0,863
Languedoc Roussillon	125	0,836	0,855
Haute Normandie	81	0,834	0,854
Rhône Alpes	253	0,834	0,854
Poitou Charentes	69	0,834	0,857
Pays de la Loire	181	0,833	0,850
Auvergne	63	0,832	0,854
Cotes d'Azur	215	0,832	0,853
Limousin	77	0,832	0,850
Basse Normandie	63	0,831	0,848
Lorraine	110	0,831	0,850
Midi Pyrénées	188	0,831	0,852
Centre	94	0,827	0,846
Bretagne	103	0,827	0,847
Bourgogne	71	0,827	0,848
Nord Pas de Calais	172	0,826	0,853
Aquitaine	168	0,825	0,847
Champagne Ardenne	41	0,825	0,844
Ile de France	254	0,815	0,836
Total	2 554	0,832	0,851

Comme nous pouvons le voir, le classement selon la moyenne d'efficacité change avec les modèles d'estimation. Par contre, la région Alsace, Picardie, Franche-Comté et Languedoc Roussillon montrent une efficacité toujours plus haute et la région Ile de France se trouve toujours dernière dans ce classement. La raison est peut-être l'insuffisance des infrastructures dans cette région et une congestion très importante.

8-4-2 Résultat de la fonction de coût

Coefficients estimés

Comme pour les résultats de la fonction de production, nous présentons d'abord les coefficients estimés et le résultat des simulations pour la fonction de coût. Le tableau 8-8 donne les résultats pour la forme Cobb-Douglas et le tableau 8-9 ceux pour la forme Translog. La première observation qui nous frappe est que contrairement au cas primal, le paramètre associé aux véhicules est significatif. Cela découle certainement du processus d'agrégation de certains inputs variables que nous avons imposé à cause du manque de données sur les prix. Donc pour la partie duale, toutes les variables d'inputs sont significatives pour tous les niveaux de confiance. À la section 8-2, nous avons présenté les règles et conditions que la fonction de coût doit respecter pour être compatibles avec la théorie économique. Puisque ces conditions sont imposées dans les programmes mathématiques, linéaire et quadratique, nos résultats respectent automatiquement ces contraintes: les dérivées de la fonction de coût par rapport aux inputs variables et l'output sont positives; les dérivées par rapport aux inputs quasi-fixes sont négatives et la fonction est homogène de degré un dans les prix des inputs variables. Finalement, comme nous le voyons aux tableaux 8-8 et 8-9, il n'y a pas de grandes différences entre les coefficients estimés pour une même forme fonctionnelle en utilisant le programme linéaire ou le programme quadratique. Donc comme dans le cas primal, le critère d'optimisation n'influence pas beaucoup nos résultats.

Tableau 8-8 : Résultats de simulation des coefficients pour la fonction de coût Cobb-Douglas

Dual	Coefficient estimé	Biais d'estimation	Coefficient corrigé	Borne inférieure	$I_{90\%}$ Borne supérieure	Borne inférieure	$I_{95\%}$ Borne supérieure	Borne inférieure	$I_{99\%}$ Borne supérieure
Programme									
linéaire									
φ_{W_L}	0,845	-0,001	0,846	0,731	0,893	0,725	0,941	0,701	0,981
φ_{W_K}	0,155	0,001	0,154	0,107	0,269	0,059	0,275	0,019	0,299
φ_V	0,400	-0,027	0,427	0,382	0,524	0,363	0,555	0,348	0,619
φ_{Inf}	0,381	0,062	0,319	0,310	0,593	0,216	0,603	0,210	0,691
φ_y	1,186	0,064	1,112	0,786	1,219	0,716	1,250	0,617	1,294
Programme									
quadratique									
φ_{W_L}	0,842	0,025	0,817	0,727	0,871	0,716	0,973	0,700	0,979
φ_{W_K}	0,158	-0,025	0,183	0,129	0,273	0,027	0,284	0,021	0,301
φ_V	0,410	-0,026	0,436	0,390	0,531	0,382	0,596	0,348	0,641
φ_{Inf}	0,481	0,114	0,367	0,335	0,623	0,274	0,671	0,208	0,680
φ_y	1,262	0,124	1,138	0,842	1,296	0,678	1,325	0,625	1,387

Tableau 8-9 : Résultats de simulation des coefficients pour la fonction de coût Translog

Dual	Coefficient estimé	Biais d'estimation	Coefficient corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
				Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme linéaire									
φ_{W_L}	0,607	-0,264	0,871	0,483	0,962	0,472	0,980	0,433	1,000
φ_{W_K}	0,393	-0,036	0,429	0,038	0,517	0,020	0,528	0,000	0,567
φ_V	1,732	-0,300	2,032	1,482	2,315	1,463	2,463	1,305	2,582
φ_{Inf}	-1,325	-0,456	-0,869	-1,5192	-0,6288	-1,649	-0,574	-1,762	-0,310
φ_y	-2,217	-0,277	-1,940	-2,4472	-1,819	-2,684	-1,516	-2,910	-1,302
γ_{WL^2}	-0,261	-0,076	-0,185	-0,532	-0,236	-0,665	-0,163	-0,825	-0,032
γ_{WK^2}	-0,261	-0,076	-0,185	-0,532	-0,236	-0,665	-0,163	-0,825	-0,032
γ_{V^2}	0,079	-0,029	0,108	0,020	0,095	0,016	0,158	0,012	0,221
γ_{INF^2}	1,203	-0,192	1,395	1,142	1,518	1,082	1,696	0,811	1,959
γ_{WL_WK}	0,261	0,076	0,185	0,163	0,532	0,136	0,665	0,032	0,825
γ_{WL_V}	-0,002	0,002	-0,004	-0,011	0,005	-0,018	0,008	-0,023	0,019
γ_{WL_Inf}	-0,503	0,171	-0,674	-0,731	-0,475	-0,974	-0,339	-1,208	-0,203
γ_{WK_V}	0,002	-0,002	0,004	-0,05	0,011	-0,08	0,014	-0,19	0,003
γ_{WK_Inf}	0,503	-0,171	0,674	0,475	0,731	0,339	0,974	0,203	1,208

\mathcal{Y}_{V_Inf}	-0,316	0,016	-0,331	-0,505	-0,157	-0,631	-0,131	-0,782	-0,079
\mathcal{Y}_{y^2}	0,205	-0,194	0,399	0,137	0,281	0,125	0,302	0,119	0,373
\mathcal{Y}_{WL_y}	0,094	0,000	0,094	0,009	0,217	0,008	0,260	0,007	0,304
\mathcal{Y}_{V_y}	-0,083	-0,007	-0,076	-0,125	-0,064	-0,166	-0,046	-0,206	-0,028
\mathcal{Y}_{Inf_Y}	-0,134	0,131	-0,265	-0,178	-0,075	-0,182	-0,063	-0,221	-0,038

Programme

quadratique

φ_{WL}	0,453	-0,309	0,762	0,147	0,692	0,113	0,961	0,089	1,00
φ_{WK}	0,547	-0,112	0,659	0,308	0,853	0,039	0,887	0	0,911
φ_V	1,257	-0,415	1,672	1,088	1,440	1,014	1,839	0,855	2,012
φ_{INF}	-1,625	-0,156	-1,468	-1,714	-1,530	-1,822	-1,166	-1,969	-0,913
φ_y	-1,750	-0,418	-1,332	-1,867	-1,196	-1,884	-0,978	-1,955	-0,636
\mathcal{Y}_{WL^2}	0,119	-0,185	0,304	0,077	0,400	0,059	0,511	0,052	0,631
\mathcal{Y}_{WK^2}	0,119	-0,185	0,304	0,077	0,400	0,059	0,511	0,052	0,631
\mathcal{Y}_{V^2}	0,082	0,094	-0,012	-0,063	0,103	-0,074	0,142	-0,096	0,195
\mathcal{Y}_{Inf^2}	0,939	-0,228	1,168	0,819	0,989	0,756	1,173	0,632	1,288
\mathcal{Y}_{WL_WK}	-0,119	0,185	-0,304	-0,400	-0,077	-0,511	-0,059	-0,631	-0,052
\mathcal{Y}_{WL_V}	0,008	0,043	-0,035	-0,181	0,163	-0,212	0,207	-0,301	0,239
\mathcal{Y}_{WL_Inf}	-0,249	-0,040	-0,209	-0,736	-0,146	-0,987	-0,101	-1,402	-0,079

γ_{wK_v}	-0,008	-0,043	0,035	-0,163	0,181	-0,207	0,212	-0,239	0,301
γ_{wK_Inf}	0,249	0,040	0,209	0,146	0,736	0,101	0,987	0,079	1,402
γ_{V_Inf}	-0,045	-0,012	-0,033	-0,111	-0,007	-0,130	-0,006	-0,163	-0,004
γ_{y^2}	0,208	0,100	0,108	0,046	0,236	0,037	0,270	0,030	0,330
γ_{wL_y}	0,039	0,042	-0,003	0,026	0,145	0,015	0,208	0,009	0,296
γ_{V_y}	-0,103	0,083	-0,111	-0,233	-0,097	-0,253	-0,089	-0,288	-0,057
γ_{Inf_y}	-0,063	0,043	-0,106	-0,166	-0,058	-0,233	-0,046	-0,306	-0,028

Elasticités

Au tableau 8-10, nous présentons l'élasticité de coût par rapport au travail, aux matériaux, à l'output, aux véhicules, aux infrastructures et l'élasticité de l'output par rapport aux infrastructures obtenu par la fonction de coût.

Tableau 8-10 : Elasticité des inputs pour la fonction de coût

	S_L	S_K	ξ_{Cy}	ξ_{CV}	ξ_{CInf}	ξ_{yInf}
Cobb-Douglas						
Programme linéaire	0,845	0,155	1,186	-0,400	-0,381	0,321
Programme Quadratique	0,842	0,158	1,262	-0,410	-0,481	0,381
Translog						
Programme linéaire	0,777	0,223	1,188	-0,330	-0,552	0,465
Programme Quadratique	0,795	0,205	1,288	-0,329	-0,537	0,417

$s_p = \partial \ln C / \partial \ln w_p$ est l'élasticité du coût variable par rapport à l'input variable x_p ;

$\xi_{Cy} = \partial \ln C / \partial \ln y$ est l'élasticité du coût variable par rapport à l'output ;

$\xi_{Ck_s} = \partial \ln C / \partial \ln k_s, k_s = V, Inf$ est l'élasticité du coût variable par rapport à l'input quasi-fixe k_s ,

$\xi_{yInf} = \partial \ln y / \partial \ln Inf = -\xi_{CInf} / \xi_{Cy}$ est l'élasticité de l'output par rapport aux infrastructures calculée par la fonction de coût.

Comme dans le cas de la fonction de production, nous constatons que l'influence des différents critères d'optimisation est négligeable. Les coefficients estimés par une même forme fonctionnelle sont très proches. Par contre, encore une fois, entre la forme Cobb-Douglas et la forme Translog, il y a une différence assez claire. Cela signifie que le critère d'optimisation (minimiser la somme des erreurs ou minimiser la somme de carré des erreurs) n'est pas important, mais qu'une forme plus flexible nous donne des résultats différents. Ensuite, nous trouvons que l'élasticité de la main-d'œuvre est toujours plus grande que celles des autres inputs variables. Ceci signifie que parmi les inputs qui peuvent être contrôlés par les firmes, la main d'œuvre reste le facteur le plus important avec une élasticité d'environ 0,84 pour la forme Cobb-

Douglas et 0,78 pour la forme Translog, signifiant que la part du coût dépensé sur la main d'œuvre est la plus importante dans le coût variable total. Par contre, la part des matériaux dans le coût variable est beaucoup moins importante, avec environ 0,16 pour la Cobb-Douglas et entre 0,20 et 0,22 pour la forme Translog. Les valeurs de ξ_{Cy} indiquent qu'une augmentation de 1% de l'output se traduit par une augmentation de coût variable de 1,186% à 1,288%. ξ_{CV} et ξ_{CInf} nous donne l'ampleur de la variation du coût variable suit une variation des inputs quasi-fixes, une augmentation de 1% des véhicules fera diminuer d'environ 0,4% le coût variable pour la forme Cobb-Douglas et 0,33% pour la forme Translog. Une augmentation de 1% des infrastructures entraînera une baisse du coût variable de 0,38% ou encore de 0,48% pour la forme Cobb-Douglas selon le critère retenu, et d'environ 0,54% pour la forme Translog. L'effet des infrastructures sur la réduction de coût variable est donc plus important que celui des véhicules.

L'équation (8-13) montre comment mesurer l'effet des infrastructures sur l'output avec la fonction de coût. Nous obtenons une élasticité de 0,351 avec la Cobb-Douglas et 0,441 avec la Translog. Par rapport au cas primal, nous trouvons que l'effet des infrastructures sur l'output est plus important pour la fonction de production Cobb-Douglas, qui a une élasticité de $\eta_{inf} = 0,654$; par contre, cet effet est très proche des résultats obtenus par la fonction de production Translog, qui a une élasticité de $\eta_{inf} = 0,469$. Mais ces deux résultats nous confirment que l'influence des infrastructures est importante dans ce secteur. Finalement, comme pour les résultats que nous venons de présenter, nous pouvons étudier de quelle façon la demande d'inputs variables peut être affectée ou ajustée en raison d'une mise à disposition additionnelle d'infrastructures publiques pour les entreprises de transport. Ces ajustements de la demande de facteurs ou les effets du biais d'utilisation des facteurs variables causés par les infrastructures peuvent être estimés grâce à l'équation (8-12) alors que l'effet total sur la demande de facteurs variables est mesuré par l'équation

(8-19). Nous dirons que les infrastructures et l'input variable x_p sont substituables, indépendants, ou complémentaires si ξ_{p_Inf} est respectivement inférieure, égale ou supérieure à zéro. Nous présentons les valeurs calculées au tableau 8-11. Ces résultats sont intéressants car ils indiquent que les infrastructures favorisent le processus de production pour les entreprises avec une forte intensité sur les matériaux et sur les véhicules, mais ont un effet opposé pour les entreprises qui ont une forte intensité en main d'œuvre. Les infrastructures et la main d'œuvre sont substituables alors que les infrastructures et le capital sont complémentaires. Mais la relation entre les infrastructures et les véhicules change selon le critère d'optimisation. Une explication pour cette variabilité est que ces mesures dépendent de termes de deuxième ordre qui sont moins bien estimés que les termes de premier ordre.

Tableau 8-11 : Effet externe des infrastructures sur les autres inputs

Fonction de coût Translog	Programme linéaire	Programme quadratique
τ_{L_Inf}	-0,647	-0,313
τ_{cap_Inf}	2,256	1,215
ξ_{L_Inf}	-1,199	-0,850
ξ_{cap_Inf}	1,704	0,678
ξ_{V_Inf}	0,406	-0,400

$\tau_{p_Inf} = \partial \ln S_p / \partial \ln Inf$, $p = L, Cap$ est l'élasticité de la part des autres inputs dans le coût variable par rapport aux infrastructures;

$\xi_{p_Inf} = \partial \ln x_p / \partial \ln Inf$, $p = L, Cap$ est l'élasticité des inputs variables par rapport aux infrastructures et

$\xi_{V_Inf} = \partial \ln V / \partial \ln Inf$ est l'élasticité des véhicules par rapport aux infrastructures.

Rendements d'échelle

Comme pour le cas primal, nous présentons d'abord l'histogramme des rendements d'échelle aux figures 8-7 et 8-8. Pour les rendements d'échelle à court terme, il n'y a que 18,3% des entreprises avec des rendements d'échelle supérieurs à 1 pour le programme de minimisation de la somme des erreurs (PL) et 17,4 % pour le programme de minimisation de la somme du carré des erreurs (PQ). Ça signifie qu'à

court terme, la plupart des entreprises montre des rendements d'échelle décroissants, et ce résultat est également vérifié pour les quantiles de 75 % des rendements d'échelle de ces deux programmes (0,953 pour PL et 0,899 pour PQ). À long terme, le pourcentage des entreprises qui sont caractérisées par des rendements d'échelles supérieures à 1 est très important (95,5 % pour PL et 99,6 % pour PQ). Ça signifie qu'à long terme, les entreprises sont en situation de rendements d'échelle croissants. Ce résultat est vérifié également pour les quantiles de 75 % de 1,789 pour PL et 1,655 pour PQ. Les quantiles de 25 % est de 1,262 pour PL et 1,294 pour PQ.

Les résultats des simulations pour les rendements d'échelles sont présentés au tableau 8-12. En maintenant les infrastructures fixes, avec la forme Cobb-Douglas, pour tous les niveaux de confiance, les entreprises sont inefficaces à l'échelle. Pour la forme Translog, le résultat est moins clair. Pour le programme linéaire, la possibilité pour laquelle nous ne pouvons pas refuser l'hypothèse que les entreprises fonctionnent efficacement à l'échelle est très faible (1 %). Pour le programme quadratique, la possibilité montre à 5 %.

En intégrant les indicateurs d'accessibilité, le secteur du transport routier de voyageur en France se caractérise par des rendements d'échelle significativement croissants pour les deux formes fonctionnelles et tous les niveaux de confiance. Cela signifie qu'il est plus coûteux lorsque l'offre de transport est assurée par plusieurs unités plutôt que par une seule. Le fait qu'une DMU bénéficie d'économies d'échelle au cours du processus de production justifie sa position de monopole sur le marché. Donc du point de vue de la société, il faut des entreprises plus grandes et plus d'infrastructures à la condition que le marché soit capable d'absorber cette hausse de production.

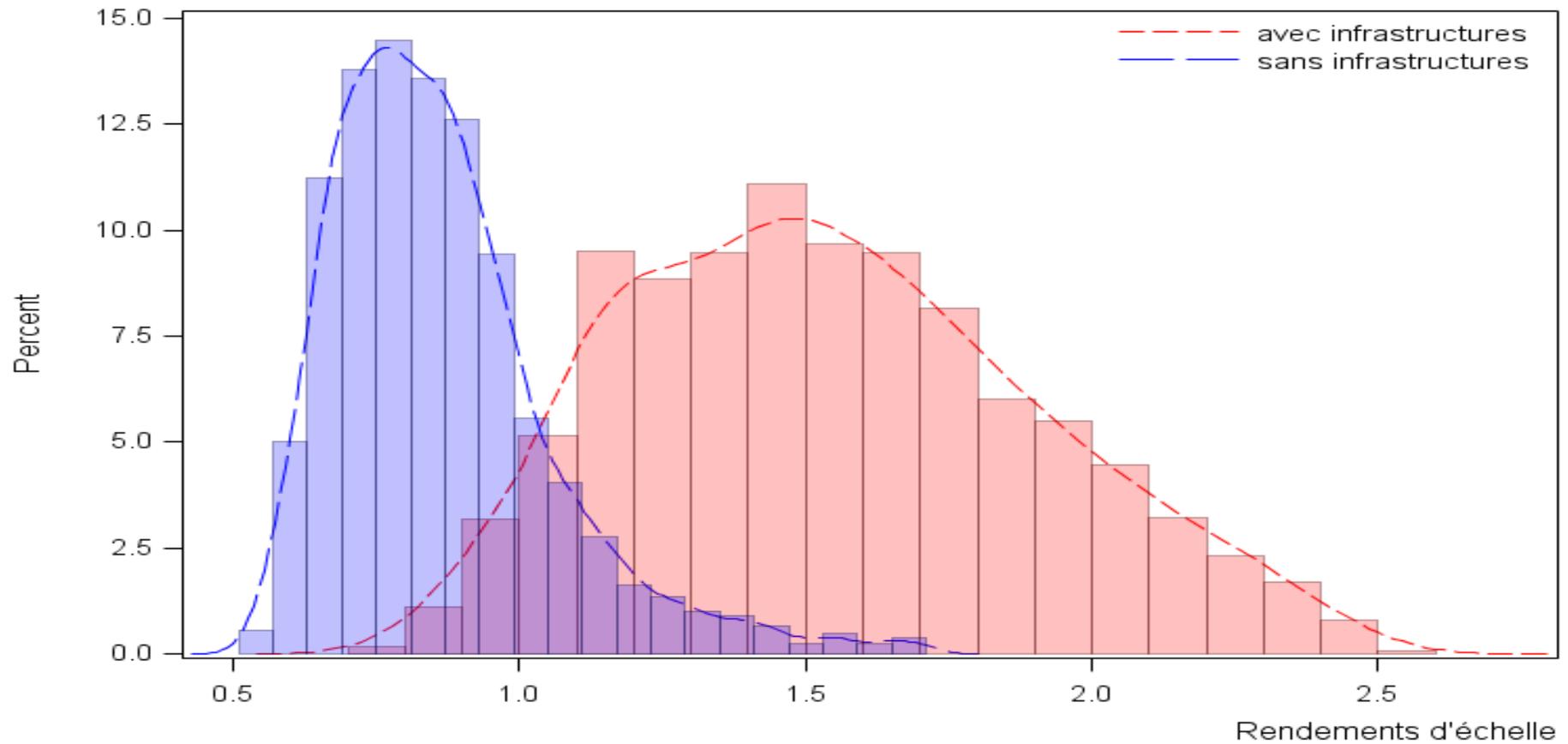


Figure 8-7 : Histogramme des rendements d'échelles avec et sans infrastructures, pour la fonction de coût Translog en minimisant la somme des erreurs (cas PL)

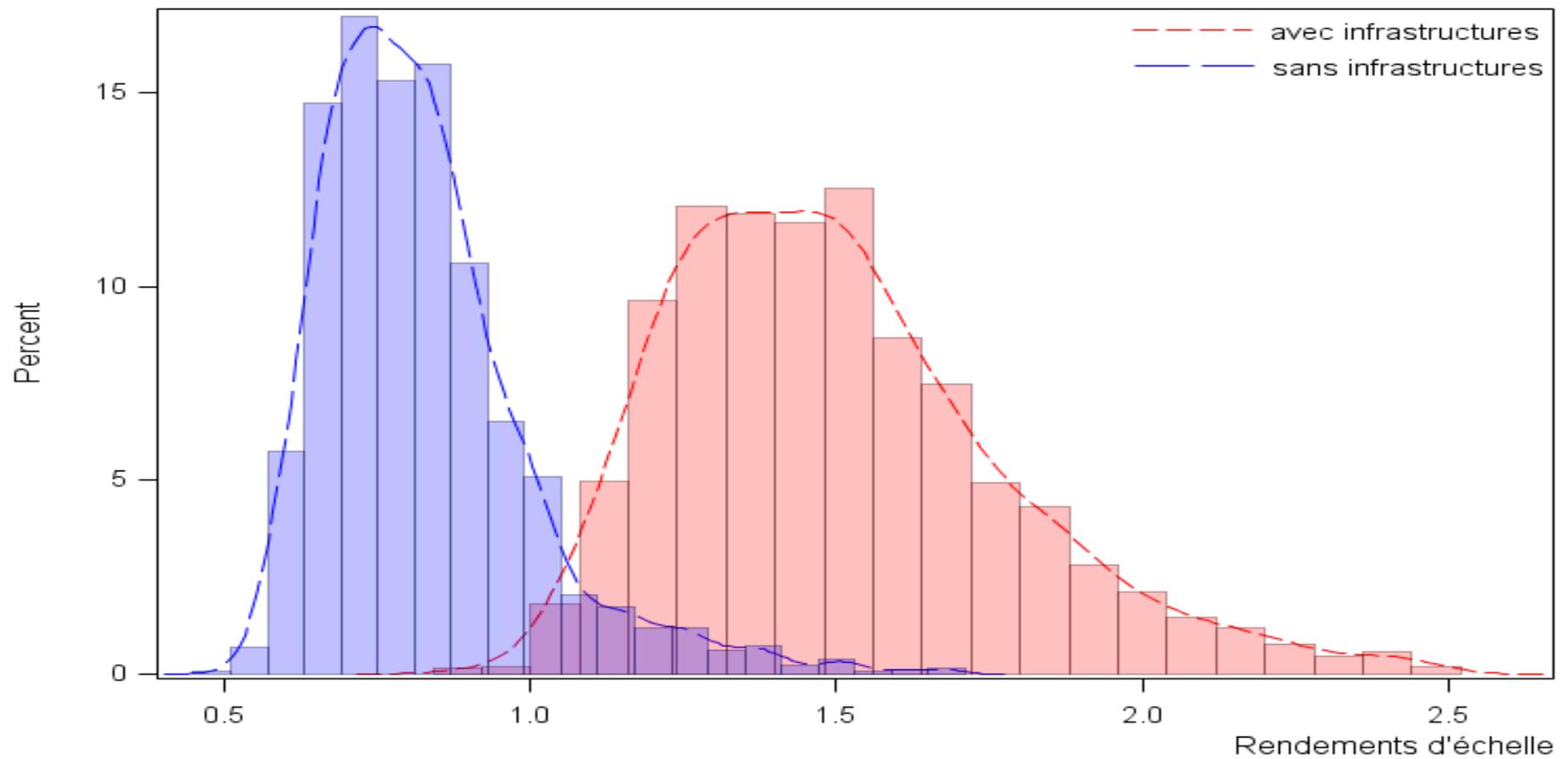


Figure 8-8 : Histogramme des rendements d'échelles avec et sans infrastructures, pour la fonction de coût Translog en minimisant la somme du carré des erreurs (cas PQ)

Tableau 8-12 Résultats de simulation des rendements d'échelles pour la fonction de coût

Cobb-Douglas		Coefficient Estimé	Biais d'estimateur	Coefficient Corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
					Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme linéaire	η_{sans}^E	0,843	-0,026	0,869	0,609	0,902	0,578	0,937	0,528	0,960
	η_{avec}^E	1,487	-0,015	1,502	1,213	1,573	1,194	1,645	1,030	1,720
Programme quadratique	η_{sans}^E	0,792	-0,017	0,809	0,701	0,905	0,641	0,946	0,608	0,975
	η_{avec}^E	1,555	-0,011	1,566	1,204	1,673	1,182	1,695	1,129	1,734

Translog		Coefficient Estimé	Biais d'estimateur	Coefficient Corrigé	$I_{90\%}$		$I_{95\%}$		$I_{99\%}$	
					Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Programme linéaire	Médiane	0,833	0,012	0,821	0,760	0,900	0,747	0,912	0,713	1,093
	Min	0,527	0,054	0,473	0,331	0,607	0,303	0,634	0,287	0,653
	η_{sans}^E Max	1,700	0,047	1,653	1,718	1,789	1,550	1,808	1,474	1,866
	P25	0,728	0,075	0,653	0,583	0,789	0,578	0,826	0,559	0,857
	P75	0,953	-0,11	1,063	0,900	1,071	0,893	1,128	0,863	1,176
Programme linéaire	η_{avec}^E Médiane	1,514	-0,021	1,535	1,491	1,579	1,463	1,641	1,336	1,681
	Min	0,765	0,031	0,734	0,671	0,929	0,660	0,971	0,569	1,015
	Max	2,57	0,134	2,436	2,469	2,592	2,258	2,631	2,142	2,712
	P25	1,262	0,145	1,117	1,055	1,297	1,036	1,323	0,989	1,351

	P75	1,789	-0,209	1,998	1,704	1,793	1,690	1,868	1,634	1,947
Programme	Médiane	0,796	0,016	0,780	0,727	0,920	0,677	1,096	0,642	1,154
linéaire	Min	0,485	0,044	0,441	0,426	0,838	0,403	0,872	0,374	0,924
η_{sans}^E	Max	1,692	0,047	1,645	1,461	1,750	1,439	1,955	1,407	1,967
	P25	0,706	0,082	0,624	0,610	0,730	0,600	0,816	0,587	0,822
	P75	0,899	-0,152	1,051	0,776	0,930	0,704	1,139	0,677	1,185
Programme	Médiane	1,465	-0,018	1,483	1,267	1,548	1,257	1,631	1,215	1,701
linéaire	Min	0,869	0,025	0,844	0,600	1,508	0,331	1,542	0,107	1,676
η_{avec}^E	Max	2,501	0,11	2,391	2,159	2,587	2,127	2,890	2,079	2,908
	P25	1,294	0,165	1,129	1,117	1,339	1,100	1,495	1,076	1,505
	P75	1,655	-0,200	1,855	1,429	1,712	1,407	1,912	1,376	1,924

Efficacité

Comme pour la fonction de production, nous pouvons calculer l'efficacité technique à partir de la fonction de coût. Par définition, nous aurons une efficacité supérieure à un. Car dans le cas de la fonction de coût, nous enveloppons les données de dessous, nous avons une enveloppe inférieure et les coûts inefficaces sont au-dessus de l'enveloppe.

La formule est donnée par l'équation (8-11) et correspond au rapport entre le coût observé et le coût efficient. Nous présentons au tableau 8-13 les coefficients d'efficacité. Peu importe la forme retenue, les entreprises dans ce secteur peuvent diminuer de 11,9% leur coût variable sans diminuer l'output. Encore une fois, ce résultat n'est pas très loin du résultat obtenu avec la méthode DEA. Les entreprises dans ce secteur sont toujours relativement efficaces.

Tableau 8-13 : Efficacité des entreprises par la fonction de coût

		Cobb-Douglas		Translog	
		PL	PQ	PL	PQ
θ		1,130(0,885)	1,129(0,886)	1,122(0,891)	1,120(0,893)
\dot{N}		42		49	
θ	5 % plus petites	1,362	1,336	1,236	1,233
	5 % plus grandes	1,093	1,085	1,056	1,060

PL : programme linéaire de minimisation de la somme des erreurs, PQ : programme quadratique de minimisation de la somme du carré des erreurs. θ correspond au coefficient d'efficacité moyenne pondérée de secteur. \dot{N} représente la différence en valeur absolue du classement des DMU selon leur efficacité estimée par PL ou PQ.

Pour les figures ci-dessous nous adoptons la même présentation que celle de la figure 8-1. Sur l'axe des abscisses, chaque groupe de 127 DMU représente 5 % de l'échantillon. Sur l'axe vertical, nous notons le coefficient d'efficacité en coût. Nous trouvons une tendance claire : les très petites entreprises (les 5 % les plus petites) sont nettement moins efficaces. Elles sont environ 20 % moins efficaces que les 5 % les plus grandes. Nous avons calculé la moyenne de θ pour le premier 5 % et pour le dernier 5 % (par chiffre d'affaires). Et comme le résultat dans la fonction de

production, les plus petites montre une inefficacité plus importante, les plus grandes sont relativement plus efficaces.

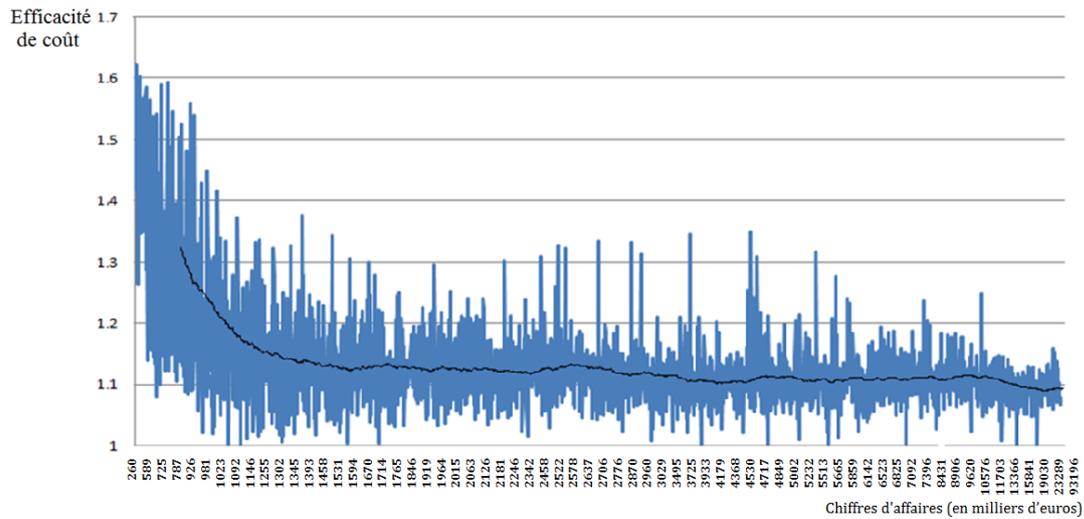


Figure 8-9 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Cobb-Douglas en minimisant la somme des erreurs

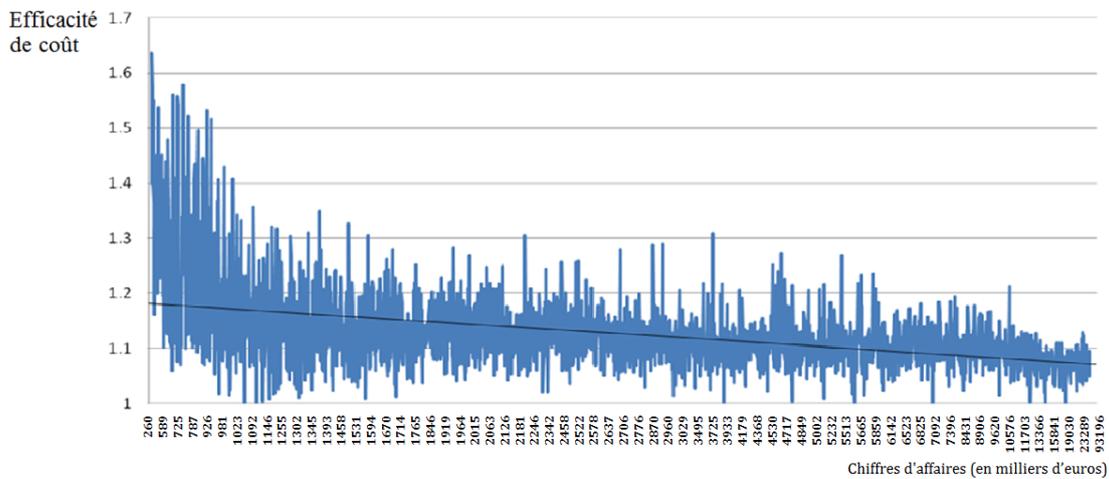


Figure 8-10 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Cobb-Douglas en minimisant la somme du carré des erreurs

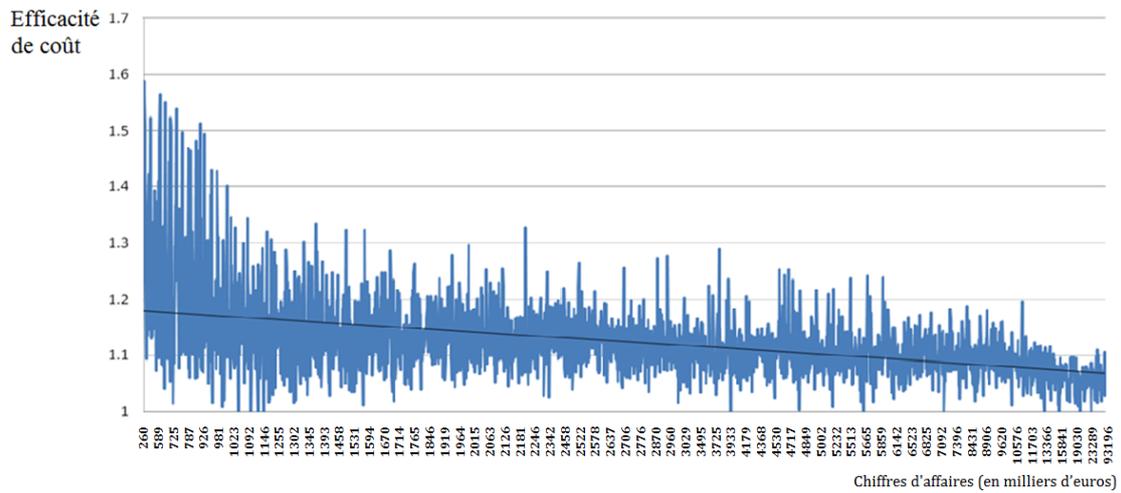


Figure 8-11 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Translog en minimisant la somme des erreurs

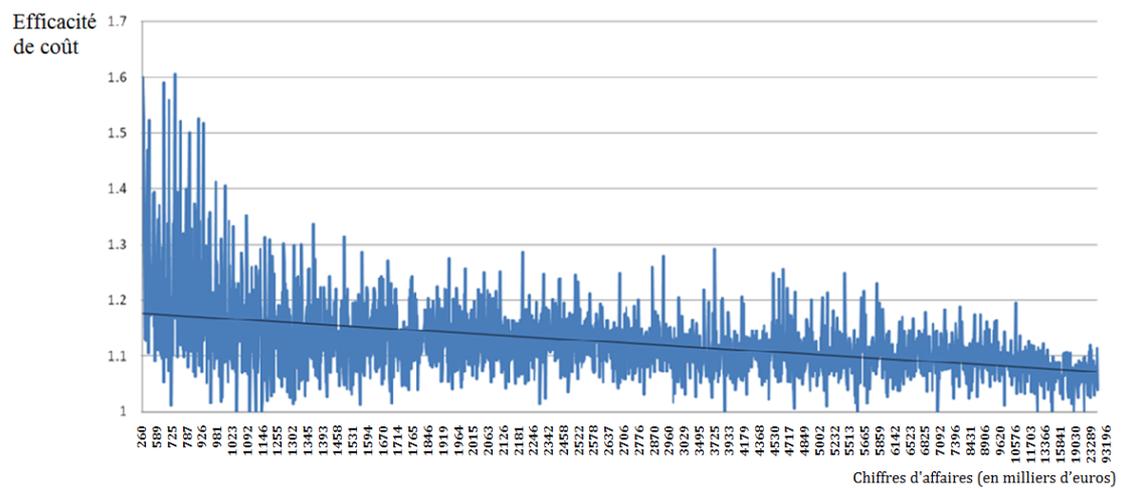


Figure 8-12 : Relation entre la performance et le chiffre d'affaire avec la fonction de coût Translog en minimisant la somme du carré des erreurs

Maintenant, nous nous tournons vers la distribution des coefficients d'efficacité en coût pondéré (par le coût), par région française. Encore une fois, le classement régional pour l'efficacité de coût ne change pas beaucoup : Alsace, Franche Comté, Rhône Alpes et Languedoc Roussillon sont les régions les plus efficaces en coût, et la région Ile de France est encore une fois la moins efficace. Nous avons trouvé presque

le même classement des régions avec la fonction de production, cela signifie que la méthode duale donne des résultats conformes à ceux de la méthode primale.

Tableau 8-14 : Coefficients de l'efficacité-coût par région, modèle semi-paramétrique, forme Translog

Région	Nombre d'observations	Moyenne
Alsace	68	1,101
Languedoc Roussillon	125	1,104
Franche Comté	79	1,104
Rhône Alpes	253	1,105
Bourgogne	71	1,110
Picardie	79	1,102
Bretagne	103	1,122
Centre	94	1,115
Pays de la Loire	181	1,124
Côtes d'Azur	215	1,115
Lorraine	110	1,123
Poitou Charentes	69	1,117
Haute Normandie	81	1,130
Limousin	77	1,131
Basse Normandie	63	1,130
Nord-Pas de Calais	172	1,127
Champagne Ardenne	41	1,133
Aquitaine	168	1,139
Midi Pyrénées	188	1,131
Auvergne	63	1,147
Ile de France	254	1,149
Total	2554	1,122

8-4-3 Comparaison entre les modèles primal et dual de Aigner et Chu et le modèle DEA

Pour compléter nos études, nous avons également estimé la fonction de production avec les inputs agrégés comme ce que nous avons fait pour la fonction de coût. Les résultats sont présentés dans les tableaux de l'Annexe III.

Nous constatons d'abord que nous avons eu presque les mêmes résultats pour l'élasticité de l'output par rapport à la main-d'œuvre, aux véhicules et aux infrastructures. L'efficacité en moyenne n'est pas trop changée. Cela semble indiquer que l'agrégation des inputs est acceptable et que les résultats sont robustes.

Maintenant, nous pouvons comparer les résultats entre la fonction de production avec inputs agrégés et la fonction de coût avec inputs agrégé. Qualitativement, nous obtenons les mêmes conclusions : lorsque l'on enlève les infrastructures, nous ne pouvons pas refuser l'hypothèse que les entreprises montrent qu'une taille optimale (pour le niveau de confiance à 95 %) ; quand nous prenons en compte des infrastructures, les entreprises produisent avec des rendements d'échelles croissants. De plus, pour la forme Translog, nous constatons que c'est toujours la main-d'œuvre qui possède l'élasticité la plus importante. Quantitativement, quand nous analysons la contribution des infrastructures sur la production (η_{inf} pour la fonction de production au tableau III-5 et ξ_{yinf} au tableau 8-10), nous trouvons qu'il n'y a pas de grandes différences dans les résultats obtenus pour la forme Translog entre le modèle primal et le modèle dual. Par contre, pour la forme Cobb-Douglas, les résultats sont assez différents ; cela vient du fait que la forme Cobb-Douglas est très restrictive. Donc, pour conclure, la fonction de production et la fonction de coût nous donne toujours les mêmes résultats et l'agrégation de certains inputs semble justifiée dans notre cas d'étude. Le modèle primal et le modèle dual analyse différents aspects des caractéristiques technologiques des entreprises, mais dans les deux modèles les

infrastructures influencent positivement la performance des entreprises : soit par une augmentation de la production soit par une diminution des coûts de production.

Comparons maintenant la méthode semi-paramétrique et la méthode non-paramétrique. Les deux méthodes nous montrent que sans considérer les infrastructures pour mesurer la performance des entreprises, nos résultats seront biaisés mais la méthode semi-paramétrique nous retourne une efficacité des entreprises plus grande que la méthode non-paramétrique. Par contre, car la frontière estimée par la méthode paramétrique enveloppe des données de plus près, nous devons d'avoir une efficacité mesurée plus importante. Avec les résultats au contraire (l'efficacité mesuré par la méthode semi-paramétrique est plus importante), cela indique que la question des rendements d'échelle est importante quand nous mesurons l'impact des infrastructures sur la performance des entreprises. La méthode semi-paramétrique les capte mieux et peu mieux coller aux données. La méthode semi-paramétrique nous permet d'une part, d'obtenir les coefficients techniques des entreprises comme la méthode non-paramétrique ; elle nous permet également de mesurer quantitativement les caractéristiques techniques des entreprises (comme les élasticités, les rendements d'échelles, etc.).

8-4-4 Résultats sans indicateur d'infrastructures

Dans cette partie, pour démontrer la robustesse de nos résultats, nous présentons une comparaison des paramètres obtenus avec et sans l'indicateur d'infrastructure.

Tableau 8-15 : Coefficients estimés avec la forme Cobb-Douglas

	Programme linéaire avec infrastructures	Programme linéaire sans infrastructures	Programme quadratique avec infrastructures	Programme quadratique sans infrastructures
Primal	Fonction de production Cobb-Douglas			
α_L	0,441	0,448	0,434	0,441
α_E	0,231	0,245	0,275	0,292
α_{ENT}	0,052	0,079	0,042	0,047
α_M	0,154	0,208	0,160	0,206
α_V	0,013	0,032	0,012	0,029
α_{inf}	0,669		0,638	
Dual	Fonction de coût Cobb-Douglas			
φ_{WL}	0,845	0,769	0,842	0,766
φ_{WCap}	0,155	0,231	0,158	0,234
φ_V	0,400	0,312	0,410	0,326
φ_{inf}	0,381		0,481	
φ_y	1,186	1,300	1,262	1,319

Note : « avec » indique le cas où les infrastructures sont incluses dans le traitement ; « sans » indique le cas sans infrastructures. « PL » indique le programme linéaire ; « PQ » indique le programme quadratique.

Tableau 8-16 : Coefficients estimés de la fonction de production Translog avec et sans infrastructures

	Fonction de production Translog			
	Programme linéaire avec infrastructures	Programme linéaire sans infrastructures	Programme quadratique avec infrastructures	Programme quadratique sans infrastructures
α_L	0,903	0,501	0,901	0,571
α_E	0,012	0,055	0,007	0,195
α_{ENT}	-0,038	0,092	-0,009	0,138
α_M	-0,042	0,053	-0,042	0,040
α_V	-0,021	0,014	-0,023	0,021
α_{inf}	2,264		2,241	
β_{L^2}	0,251	0,244	0,251	0,246
β_{E^2}	0,063	0,070	0,063	0,063
β_{ENT^2}	0,003	0,003	0,004	0,006
β_{M^2}	-0,002	-0,001	-0,002	-0,001
β_{V^2}	0,007	0,011	0,007	0,014
β_{inf^2}	-0,406		-0,398	
$\beta_{L,E}$	-0,094	-0,077	-0,094	-0,092
$\beta_{L,ENT}$	-0,036	-0,039	-0,036	-0,052
$\beta_{L,M}$	0,017	0,038	0,017	0,027
$\beta_{L,V}$	0,005	0,012	0,005	0,010
$\beta_{L,inf}$	-0,086		-0,085	
$\beta_{E,ENT}$	0,015	0,022	0,016	0,024
$\beta_{E,M}$	0,006	0,011	0,006	0,008
$\beta_{E,V}$	-0,009	-0,019	-0,010	-0,022
$\beta_{E,inf}$	0,028		0,029	
$\beta_{ENT,M}$	0,005	0,009	0,005	0,006
$\beta_{ENT,V}$	-0,006	-0,010	-0,006	-0,008
$\beta_{ENT,inf}$	0,028		0,021	
$\beta_{M,V}$	0,002	0,007	0,003	0,005
$\beta_{M,inf}$	0,014		0,014	
$\beta_{V,inf}$	0,007		0,007	

Tableau 8-17 : Coefficients estimés de la fonction de coût Translog avec et sans infrastructures

	Fonction de Coût Translog			
	PL avec	PL sans	PQ avec	PQ sans
φ_{WL}	0,607	-0,254	0,453	-0,404
φ_{WK}	0,393	0,749	0,547	0,596
φ_V	1,732	0,458	1,257	0,378
φ_{inf}	-1,325		-1,625	
γ_{WL^2}	-0,261	-0,298	0,119	0,156
γ_{WK^2}	-0,261	-0,298	0,119	0,155
γ_{v^2}	0,079	0,079	0,082	0,087
γ_{inf^2}	1,203		0,939	
$\gamma_{WL_{WK}}$	0,261	0,298	-0,119	-0,155
γ_{WL_V}	-0,002	-0,196	0,008	-0,134
$\gamma_{WL_{inf}}$	-0,503		-0,249	
γ_{WK_v}	0,002	0,196	-0,008	0,134
$\gamma_{WK_{inf}}$	0,503		0,249	
$\gamma_{v_{inf}}$	-0,316		-0,045	
γ_Y	-2,217	-0,013	-1,750	-0,166
γ_{Y^2}	0,205	0,001	0,208	0,028
γ_{WL_Y}	0,094	0,245	0,039	0,191
γ_{v_Y}	-0,083	-0,008	-0,103	-0,019
γ_{inf_Y}	-0,134		-0,063	

Tableau 8-18 : Efficacité technique avec et sans infrastructure, modèle semi-paramétrique de Aigner et Chu

	Primal		Dual	
	avec	sans	avec	Sans
	infrastructures	infrastructures	infrastructures	infrastructures
Cobb-Douglas				
Programme linéaire	0,831	0,827	1,130	1,133
Programme quadratique	0,835	0,830	1,129	1,132
Translog				
Programme linéaire	0,853	0,849	1,122	1,129
Programme quadratique	0,859	0,853	1,12	1,127

Tableau 8-19 : Élasticité des inputs et rendements d'échelle par la fonction de production

	ξ_L	ξ_E	ξ_{ENT}	ξ_M	η_V	η_{Inf}	η_{sans}^E	η_{avec}^E
Cobb-Douglas								
PL avec infrastructures	0,441	0,231	0,052	0,154	0,013	0,669	0,891	1,560
PL sans infrastructures	0,448	0,245	0,079	0,208	0,032		1,012	1,012
PQ avec infrastructures	0,434	0,275	0,042	0,160	0,012	0,638	0,923	1,561
PQ sans infrastructures	0,441	0,292	0,047	0,206	0,029		1,015	1,015
Translog								
PL avec infrastructures	0,613	0,187	0,052	0,024	0,014	0,496	0,845	1,258
PL sans infrastructures	0,618	0,197	0,067	0,062	0,023		0,932	0,932

PQ avec infrastructures	0,609	0,188	0,057	0,025	0,015	0,442	0,881	1,261
PQ sans infrastructures	0,612	0,200	0,093	0,046	0,035		0,931	0,931

PL : Programme linéaire; PQ : Programme quadratique

Tableau 8-20 : Élasticité des inputs et rendements d'échelle avec la fonction de coût

Fonction de coût	S_L	S_K	ξ_{Cy}	ξ_{CV}	ξ_{CInf}	ξ_{yInf}	η_{sans}^E	η_{avec}^E
Cobb-Douglas PL avec	0,845	0,155	1,186	-0,400	-0,381	0,321	0,843	1,487
Cobb-Douglas PL sans	0,769	0,231	1,300	-0,312			1,009	
Cobb-Douglas PQ avec	0,842	0,158	1,262	-0,410	-0,481	0,381	0,792	1,555
Cobb-Douglas PQ sans	0,766	0,234	1,319	-0,326			1,005	
Translog PL avec	0,777	0,223	1,188	-0,330	-0,552	0,465	0,833	1,514
Translog PL sans	0,739	0,261	1,309	-0,297			0,991	
Translog PQ avec	0,795	0,205	1,288	-0,329	-0,537	0,417	0,796	1,465
Translog PQ sans	0,749	0,251	1,323	-0,283			0,970	

Note : « avec » indique le cas où les infrastructures sont incluses dans le traitement ; « sans » indique le cas sans infrastructures. « PL » indique le programme linéaire ; « PQ » indique le programme quadratique.

Pour bien capter l'influence des infrastructures sur la performance des entreprises, nous avons également estimé nos modèles empiriques en excluant les infrastructures. Nous présentons aux tableaux 8-15, 8-16, 8-17 les coefficients estimés de la fonction de production et de la fonction de coût avec et sans l'indicateur d'infrastructures. Nous présentons au tableau 8-18 l'efficacité des entreprises avec et sans infrastructures. Clairement, sans intégrer les infrastructures, le coefficient d'efficacité des entreprises va diminuer pour l'espace primal et l'espace dual. C'est-à-dire qu'en excluant les infrastructures nous sous-estimons l'inefficacité. Ensuite, les tableaux 8-19 et 8-20 nous donnent les élasticités de l'output et les élasticités de coût par rapport aux inputs et les rendements d'échelle (médiane). Pour l'espace primal, si nous

enlevons l'indicateur d'infrastructures, nous pouvons constater une augmentation des élasticités par rapport à l'output. Dans ces conditions l'effet des infrastructures est capté (partagé) par les autres inputs. À court terme (lorsque le nombre de véhicules n'est pas ajustable), les entreprises sont en rendements d'échelle décroissants, cependant qu'à long terme (lorsque les entreprises peuvent ajuster le nombre de véhicules) et si nous intégrons les infrastructures, les entreprises sont en rendements d'échelle croissants. En n'incluant pas les infrastructures, nous trouvons des rendements d'échelles presque constants, c'est-à-dire que les entreprises sont presque à la taille optimale. Ceci signifie qu'en excluant les infrastructures, il y a une tendance à sous-estimer les rendements d'échelle des entreprises et il y a un biais dans les résultats. Finalement, dans l'espace dual et sans les infrastructures, nous trouvons que la part de la main d'œuvre dans le coût variable va diminuer et la part de matériaux dans le coût variable va diminuer. Sans prendre les infrastructures, la part implicite des véhicules dans le coût variable va également diminuer. Pour les rendements d'échelles, nous trouvons les mêmes résultats que pour la fonction de production, soit une tendance à sous-estimer les rendements d'échelle à long terme.

Et finalement, nous présentons au tableau 8-21 une comparaison des résultats avec et sans les infrastructures entre le modèle non-paramétrique avec rendements d'échelles variables et le modèle semi-paramétrique avec fonction de production sous forme Translog en minimisant la somme des erreurs. Les deux méthodes montrent que les petites entreprises sont moins efficaces. Par contre, avec la méthode DEA nous trouvons que les grandes entreprises sont plus efficaces que les entreprises de taille moyenne alors que cette différence n'est pas clairement apparente avec la méthode semi-paramétrique. Les deux méthodes concordent pour trouver qu'en considérant la possibilité de changement des infrastructures, les firmes opèrent avec des rendements d'échelles croissants, tandis que l'exclusion des infrastructures nous amène à conclure que les entreprises opèrent à rendements d'échelles décroissants. Dans les deux cas,

sans les infrastructures, les résultats indiquent une sous-estimation de la performance des entreprises.

Tableau 8-21 : Comparaison des mesures d'efficacité et des rendements d'échelle selon la méthode, la taille des entreprises et la présence ou non des infrastructures

		Avec infrastructures		Sans infrastructures		Taille des entreprises mesurée par l'output (km)
		Efficacité moyenne	Rendements d'échelle médiane	Efficacité moyenne	Rendements d'échelle médiane	
DEA (BCC)	Toutes tailles confondues	0,812	1,118	0,801	0,886	2 047 491,542
	Petites entreprises	0,711	1,167	0,701	0,916	513 201,332
	Entreprises de taille moyenne	0,786	1,133	0,775	0,882	1 343 365,296
	Grandes entreprises	0,817	1,060	0,810	0,846	4 993 854,098
Méthode semi-paramétrique primal	Toutes tailles confondues	0,856	1,260	0,846	0,954	2 047 491,542
	Petites entreprises	0,808	1,347	0,795	0,992	513 201,332
	Entreprises de taille moyenne	0,852	1,258	0,849	0,963	1 343 365,296
	Grandes entreprises	0,860	1,174	0,853	0,935	4 993 854,098

Note : Les petites entreprises sont les entreprises du premier quartile sur la base du chiffre d'affaires et les grandes correspondent au dernier quartile et les entreprise de taille moyenne sont celles appartenant aux deux quartiles de part et d'autre de la médiane. Les résultats du modèle semi-paramétrique est pour la fonction de production Translog.

Conclusion

Dans ce chapitre, nous utilisons d'abord le modèle semi-paramétrique pour analyser l'efficacité des entreprises de transport routier de voyageurs pour la France métropolitaine au cours de la période 2000-2004. Nous avons estimé une fonction de production Cobb-Douglas et une fonction Translog. Nous avons également testé notre modèle dans l'espace dual de la technologie en estimant une fonction de coût. Nos résultats indiquent que le secteur du transport de voyageurs est relativement efficace, mais la plupart des petites entreprises souffrent d'inefficacité d'échelle. En plus, les très petites entreprises montrent une efficacité beaucoup moins importante que les autres entreprises. Les infrastructures sont caractérisées par une complémentarité avec les matériaux et une substituabilité avec la main d'œuvre et les véhicules. Les infrastructures ont un impact positif sur la performance des entreprises et leur influence est plus importante pour les entreprises avec une grande intensité sur les matériaux et sur les véhicules. Nous remarquons que les résultats changent selon la méthode utilisée. Nous avons constaté que le modèle de Aigner et Chu nous donne une efficacité légèrement plus grande que celle obtenue avec le modèle de DEA. Tant dans le cas dual que primal nous avons trouvé une influence positive des infrastructures sur la performance des entreprises. Pour la fonction de production, en intégrant les infrastructures dans le processus de production, les entreprises se retrouvent avec des rendements d'échelle croissants à long terme. Pour la fonction de coût, un prix implicite positif implique que les infrastructures peuvent aider les entreprises à diminuer le coût de production.

Conclusion Générale

Notre thèse avait comme deux buts. Premièrement, nous voulions proposer une façon d'inclure un traitement statistique des estimateurs semi-paramétriques de Aigner et Chu afin de mesurer l'efficacité et les caractéristiques technologiques d'une entreprise. Cet objectif a fait l'objet du chapitre 4 de notre thèse. En utilisant un théorème de Wald (1949), nous avons pu contourner certains problèmes qui invalidaient les méthodes traditionnelles de preuve de convergence des estimateurs et en recourant à des techniques de ré-échantillonnage, nous avons montré comment déduire les distributions empiriques des estimateurs.

Deuxièmement, nous voulions analyser le secteur du transport de voyageurs en France à l'aide de la méthodologie proposée. En plus de l'utilisation de notre méthode d'estimation, un aspect important de notre étude empirique consiste en l'adoption d'un indicateur d'infrastructures permettant d'incorporer la contribution de l'État qui fournit un input essentiel à ce secteur. Nos résultats permettent non seulement de caractériser et de tester la technologie de ce secteur, mais en plus de quantifier l'impact de la contribution de l'État.

Nous avons trouvé qu'avec l'introduction des infrastructures, les entreprises dans ce secteur se caractérisent par des rendements d'échelle croissants alors qu'en les excluant, elles sont en rendements décroissants pour la plupart. Ceci signifie qu'en excluant les infrastructures, il y a une tendance à sous-estimer les rendements d'échelle des entreprises. L'exclusion des infrastructures a aussi pour effet de sous-estimer l'inefficacité des firmes. Donc, nos résultats montrent que les infrastructures jouent un rôle important pour les entreprises de ce secteur et que le *design* optimal de ce secteur doit prendre en considération le rôle de cet input essentiel sans lequel l'existence même du secteur est remise en question.

BIBLIOGRAPHIE

- Afriat, S.N. (1972). "Efficiency Estimation of Production Functions". *International Economic Review*, 13(3), 568-98.
- Aigner, D.J. et Chu, S. F. (1968). "On Estimating the Industry Production Function". *The American Economic Review*, 58(4), 826-839.
- Aigner, D.J., Lovell C.A.K. et Schmidt, P. (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models". *Journal of Econometrics*, 6, 21-37.
- Albriksen, R. et Førsund, F.R. (1990), "A Productivity Study of the Norwegian Building Industry". *Journal of Productivity Analysis*, 2(1), 53-66.
- Andrews, D.W.K. (2000). "Inconsistency of the Bootstrap When a Parameter Is on the Boundary of the Parameter Space". *Econometrica*, 68, 399-405.
- Arrow, K. J., Chenery, H. B., Minhas, B. S. et Solow, R. M. (1961). "Capital-Labor Substitution and Economic Efficiency". *Review of Economics and Statistics*, 43(3), 225-250.
- Aschauer, D. (1989a). "Is Public Expenditure Productive?". *Journal of Monetary Economics*, 23, 177-200.
- Aschauer, D. (1989b). "Does Public Capital Crowd Out Private Capital?". *Journal of Monetary Economics*, 24, 171-188.
- Aschauer, D. (1989b). "Why is Infrastructure Important?" In A. Munnell (Ed.) *Is There a Shortfall in Public Capital Investment?*. Federal Reserve Bank of Boston. Conference series, 34, 21-50.
- Attaran, K. et Auclair, P. (1990). "Highway Stock and Private-Sector Productivity". *Transportation Research Record*, 1274, 119-124.
- Banister, D. et Berechman, Y. (2001). "Transport Investment and the Promotion of Economic Growth". *Journal of Transport Geography*, 9 (3), 209-218.
- Banker, R.D. et Morey, R.C. (1984). "Efficiency Analysis for Exogenously Fixed Inputs and Outputs". *Operations Research Archive*, 34, 515-518.
- Banker, R.D., Chang, H. et Cooper, W.W. (1996). "Equivalence and Implementation of Alternative Methods for Determining Returns to Scale in Data Envelopment Analysis". *European Journal of Operational Research*, 89, 473-481.

- Banker, R.D., Charnes, A. et Cooper, W.W. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis". *Management Science*, 30(9), 1078–1092.
- Banker, R.D. et Morey, R.C. (1986). "The Use of Categorical Variables in Data Envelopment Analysis". *Management Science*, 32(12), 1613-1627.
- Banque Mondiale. (1994). "*Rapport sur le développement dans le monde : une infrastructure pour le développement*", Washington DC.
- Barnett, W.A. (1983a). "New Indices of Money Supply and the Flexible Laurent Demand System". *Journal of Business and Economic Statistics*, 1, 7-23.
- Barro, R. (1990). "Government Spending in a Simple Model of Endogenous Growth". *Journal of Political Economy*, 98, 103-26.
- Barro, R. (1991). "Economic Growth in a Cross Section of Countries". *The Quarterly Journal of Economics*, 106, 407-43.
- Battese, G.E et Corra, G.S. (1977). "Estimation of a Production Frontier Model with Application to the Pastoral Zone of Eastern Australia". *Australian Journal of Agricultural Economics*, 21, 169-179.
- Berndt, E.R. et Hansson, B. (1991). "Measuring the Contribution of Public Infrastructure Capital in Sweden". NBER *Working Papers*, 3842, National Bureau of Economic Research.
- Bhatta, S.D. et Drennan, M. (2003). "The Economic Benefits of Public Investment in Transportation: A Review of Recent Literature". *Journal of Planning Education and Research*, 200, 22(3), 288-296.
- Bickel, P. et Freedman, D. (1981). "Some Asymptotic Theory for the Bootstrap". *The Annals of Statistics*, 9(6), 1196-1217.
- Bickel, P., Götze, F. et van Zwet, W. (1997). "Resampling Fewer than n Observations: Gains, Losses and Remedies for Losses". *Statistica Sinica*, 7, 1-31.
- Biggar, D. (2002). "Concurrence dans le transport routier", *Revue de l'OCDE sur le Droit et la Politique de la Concurrence*, 4, 39-119.
- Boarnet, M.G. (1998). "Spillovers and the Locational Effects of Public Infrastructure". *Journal of Regional Science*, 38, 381-400.
- Boopen, S. (2006). "Transport Infrastructure and Economic Growth: Evidence from Africa Using Dynamic Panel Estimates". *Empirical Economic Letters*, 5(1), 37.

- Bretagnolle, J. (1983). "Lois limites du bootstrap de certaines fonctionnelles". *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 19(3), 281-296.
- Bronzini, R. et Piselli, P. (2009). "Determinants of Long-Run Regional Productivity with Geographical Spillovers: The Role of R&D, Human Capital and Public Infrastructure". *Regional Science and Urban Economics*, Elsevier, 39(2), 187-199.
- Button, K. (1998). "Infrastructure Investment, Endogenous Growth and Economic Convergence". *Annals of Regional Science*, 32, 145-162.
- Buurman, J. et Rietveld, P. (1999). "Transport Infrastructure and Industrial Location: the Case of Thailand": *Journal of the Applied Regional Science Conference*, 11(1), 45-62.
- Calderón, C., Easterly, W. et Servén, L. (2002). "How Did Latin America's Infrastructure Fare in the Era of Macroeconomic Crises". The World Bank, Mimeo.
- Calderón, C., Easterly, W. et Servén, L. (2011). "Is Infrastructure Capital Productive ? A Dynamic Heterogeneous Approach". Policy Research Working Paper Series, 5682, The World Bank.
- Canning, D. et Bennathan, E. (2000). "The Social Rate of Return on Infrastructure Investments". *World Bank research project*, 680-89, Washington, D.C.
- Cazals, C., Florens, J.P. et Simar, L. (2002). "Nonparametric Frontier Estimation: A Robust Approach". *Journal of Econometrics*, 106, 1-25.
- Chandra, A. et Thompson, E. (2000). "Does Public Infrastructure Affect Economic Activity? Evidence from the Rural Interstate Highway System". *Regional Science and Urban Economics*, Elsevier, 30(4), 457-490.
- Charlot, S. et Schmitt, B. (1999). "Public Infrastructure and Economic Growth in France's Regions". *ERSA conference papers, European Regional Science Association*.
- Charnes, A., Cooper, W.W. et Rhodes E. (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2, 429-444.
- Charnes, A., Cooper, W.W., Harnes, A., Lewin, A.Y. et Seiford, L.M. (1995). "Data Envelopment Analysis : Theory, Methodology, and Applications". *Kluwer Academic Publishers, Boston*.
- Chin, M.S. et Chou, Y.K. (2004). "Modelling Social Infrastructure and Economic Growth". *Australian Economic Papers*, 43(2), 136-157.

- Christensen, L.R., Jorgenson, D.W. et Lau, L.J. (1971). "Conjugate Duality and the Transcendental Logarithmic Production Function". *Econometrica*, 39, 255-256.
- Christensen, L.R., Jorgenson, D.W. et Lau, L.J. (1973). "Transcendental Logarithmic Production Frontiers". *The Review of Economics and Statistics*, February, 28-45.
- Chu, X., Fielding, G.J. et Lamar, B.W. (1992). "Measuring Transit Performance Using Data Envelopment Analysis". *Transportation Research*, 26A(3), 223-30.
- Chung, K.L. (1974). "*A Course in Probability Theory*". San Diego, Academic Press.
- Coelli, T.J. et Battese, G.E. (1998). "*An Introduction to Efficiency and Productivity Analysis*". Boston, Kluwer Academic Publisher.
- Coelli, T.J., Prasada, R., Donnell, C.J. et Battese, G.E. (2005). "*An Introduction to Efficiency and Productivity Analysis, Second Edition*". New York, Springer.
- Cohen, J.P. et Morrison Paul, C.J. (2004). "Public Infrastructure Investment, Interstate Spatial Spillovers, and Manufacturing Costs". *The Review of Economics and Statistics*, 86(2), 551-560.
- Cohen, J.P. et Monaco, K. (2008). "Ports and Highways Infrastructure: An Analysis of Intra- and Interstate Spillovers". *International Regional Science Review*, 31(3), 257-274.
- Cooper, W.W., Seiford, M. et Tone, K. (2006). "*A Comprehensive Text with Models, Applications, References and DEA-Solver Software, second edition*". Springer Science Business Media, LLC.
- Cootner, P.H. (1964). "The Random Character of Stock Market Prices". *Operations Research*, 14(5), 962-965.
- Cowie, J. et Asenova, D. (1999). "Organisational Form, Scale Effects and Efficiency in the British Bus Industry". *Transportation*, 26(3), 231-48.
- Cullinane, K., Wang, T.F., Song, D.W., et Ji, P. (2006). "The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis". *Transportation Research Part A: Policy and Practice*, Elsevier, 40(4), 354-374.
- Davison, A.C. et Hinkley, D.V. (1997). "*Bootstrap Methods and Their Application*". Cambridge University Press.
- Dean, J., (1951). "*Managerial Economics*". Prentice Hall, Englewood Cliffs, New Jersey.

- Debreu, G. (1951). "The Coefficient of Ressource Utilization". *Econometrica*, 19, 273-292.
- Del Bo C. et Florio, M. (2008). "Infrastructure and Growth in the European Union: an Empirical Analysis at the Regional Level in a Spatial Framework". Departmental *Working Papers 37*, Department of Economics, Management and Quantitative Methods at Università degli Studi di Milano.
- Demetriades, P. et Mamuneas T.P. (2000). "Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies". *The Economic Journal*, 110, 687-712.
- Diewert, E.W. (1974). "Applications of Duality Theory". In Intrilligator, M.D. et Kendrick, D.A. (Eds.), "*Frontiers of Quantitative Economics*", Volume II, Amsterdam, North-Holland.
- Duggal, V.G., Saltzman, C. et Klein, L. (1999). "Infrastructure and Productivity: A Nonlinear Approach". *Journal of Econometrics*, 92, 47-74.
- Duggal, V. G., Saltzman, C. et Klein, L. (2007). "Infrastructure and Productivity: An Extension to Private Infrastructure and Its Productivity". *Journal of Econometrics*, Elsevier, 140(2), 485-502.
- Dujardin, C. (2001). "Effets de frontière et interaction spatiale : Les migrations alternantes et la frontière linguistique en Belgique". *L'Espace géographique*, 30(4), 307-320.
- Efron, B. (1979). "Computer and the Theory of Statistics: Thinking the Unthinkable". *Society for Industrial and Applied Mathematics Review*, 21(4), 460-480.
- Efron, B. (1983). "Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation". *Journal of American Statistical Association*, 78(382), 316-331.
- Efron, B. (1986). "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy". *Statistical Science*, 1(1), 54-75.
- Efron, B. et Tibshirani, R. (1993). "*An Introduction to Bootstrap*". New York, Chapman & Hall.
- Essid, H., Ouellette, P. et Vigeant, S. (2010). "Measuring efficiency of Tunisian schools in the presence of quasi-fixed inputs: A bootstrap data envelopment analysis approach ". *Economics of Education Review*, 29(4), 589-596.
- Everaert, G. et F. Heylen. (2004). "Public Capital and Long-Term Labour Market Performance in Belgium". *Journal of Policy Modeling*, 26, 95-112.

- Färe, R., Grosskopf, S. et Logan, J. (1983). "The Relative Efficiency of Illinois Electric Utilities". *Resources and Energy*, 5(4), 349-367.
- Färe, R., Grosskopf, S., Lindgren, B. et Roos, P. (1994) "Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach". In Charnes, A., Cooper, W., Lewin, A.Y. and Seiford, L.M. (Eds.) *Data Envelopment Analysis: Theory, Methodology and Applications*. Boston: Kluwer, 253-272.
- Farrell, M.J. (1957). "The Measurement of Productive Efficiency". *Journal of Royal Statistical Society, Series A, General*, 120, 253-81.
- Fernandez, C., Koop, G. et Steel, M. (2000). "A Bayesian Analysis of Multiple Output Production Frontiers". *Journal of Econometrics*, 98, 47-79.
- Førsund, F.R. (1992). "A Comparison of Parametric and Nonparametric Efficiency Measures: The Case of Norwegian Ferries". *Journal of Productivity Analysis*, 3, 25-44.
- Førsund, F.R. (1996). "On the Calculation of Scale Elasticities in DEA Models". *Journal of Productivity Analysis*, 7, 283-302.
- Førsund, F.R. et Jansen, E.S. (1977). "On Estimating Average and Best Practice Homothetic Production Functions via Cost Functions". *International Economic Review*, 18(2), 463-76.
- Førsund, F.R. et Hjalmarsson, L. (1979). "Frontier Production Functions and Technical Progress: A Study of General Milk Processing in Swedish Dairy Plants". *Econometrica, Econometric Society*, 47(4), 883-900.
- Fujita, M., Krugman, P. et Venables, A.J. (1999). *The Spatial Economy*. MIT Press.
- Gagné, R. et Ouellette, P. (1998). "On the Choice of Functional Forms: Summary of a Monte Carlo Experiment". *Journal of Business and Economic Statistics*, 16(1), 118-124.
- Gagné, R. et Ouellette, P. (2002). "The Effect of Technological Change and Technical Inefficiencies on the Performance of Functional Forms". *Journal of Productivity Analysis*, 17(3), 233-247.
- Gallant, A.R. (1981). "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form". *Journal of Econometrics*, 1(5), 211-245.
- Gannon, C.A. et Liu, Z. (1997). "Poverty and Transport". *Transportation, Water and Urban Development Papers, WDU-30*. The World Bank, Washington, DC.

- Garcia-Mila, T. et McGuire, T.J. (1992). "The Contribution of Publicly Provided Inputs to States Economies". *Regional Science and Urban Economics*, 22, 229-24.
- Geurts, K.T., Ritsema van Eck J.T. (2001). "Accessibility Measures: review and applications". RIVM Report 408505 006, Utrecht University, 147.
- Gramlich, E.M. (1994). "Infrastructure Investment: A Review Essay". *Journal of Economic Literature*, 32(3), 1176-1196.
- Greene, W.H. (1980a). "Maximum Likelihood Estimation of Econometric Frontier Functions". *Journal of Econometrics*, 13, 27-56
- Greene, W. H. (1980b). "On the Estimation of A Flexible Frontier Production Model". *Journal of Econometrics*, 13(1), 101-115.
- Greene, W. H. (2000). "*Econometric Analysis. (4th ed.)*". London, Prentice Hall.
- Greene, W.H. (2005). "Fixed and Random Effects in Stochastic Frontier Models". *Journal of Productivity Analysis*, 23, 7-32.
- Greene, W. H. (2008). "The Econometric Approach to Efficiency Analysis". Chapter 2 in Fried, H.O., Lovell, C.A.K et Schmidt, S. (Eds.) "*The Measurement of Productive Efficiency: Technics and Applications*". New York, Oxford University Press.
- Haan, J., Sturm, J. et Sikken, B. (1996). "Government Capital Formation: Explaining the Decline". *Review of World Economics (Weltwirtschaftliches Archiv)*, Springer, 132(1), 55-74.
- Hall, P. (1992a). "*The Bootstrap and Edgeworth Expansion*". New York: Springer-Verlag.
- Hall, R. et Jones, C. (1999). "Why Do Some Countries Produce So Much More Output per Worker Than Others?". National Bureau of Economic Research *Working Papers*, 65-64.
- Hansen, W.G. (1959). "How Accessibility Shapes Land Use". *Journal of American Institute of Planners*, 25(1), 73-76.
- Haworth, J M. et Vincent, P. J. (1979). "The Stochastic Disturbance Specification and its Implications for Log-Linear Regression". *Environment and Planning A*, 11(7), 781-790.
- Hildebrand, G.H. et Liu, T.C. (1965). "*Manufacturing Production Functions in the United States, 1957 An Interindustry and Interstate Comparison of Productivity*".

Ithaca, New York. New York State School of Industrial and Labor Relations, Cornell University.

Hirschman, A.O. (1958). *"The Strategy of Economic Development"*. New Haven, Conn.: Yale University Press.

Hirschhausen, C. et Cullmann, A. (2008). "Next Stop: Restructuring?: A Nonparametric Efficiency Analysis of German Public Transport Companies". *Transportation Research Part E: Logistics and Transportation Review*, 46(3), 436-445.

Holl, A. (2004). "Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain". *Regional Science and Urban Economics*, 34(3), 341-363.

Holtz-Eakin, D. (1994). "Public Sector Capital and Productivity Puzzle". *Review of Economics and Statistics*, 76, 12-21.

Hotelling, H. (1932). "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions". *Journal of Political Economy*, 40(5), 577-616.

Hulten, C.R. et Schwab, R. (1997). "A Fiscal Federalism Approach to Infrastructure Policy". *Regional Science and Urban Economics*, 27, 139-159.

Hulten, C.R., Bennathan, E. et Srinivansan, S. (2006). "Infrastructure, Externalities and Economic Development: A Study of Indian Manufacturing Industry". *World Bank Review*, 20(2), 291-308.

Ingram, D.R. (1971). "The Concept of Accessibility: A Search for Operational Form". *Regional Studies*, 5, 101-107.

Johansen, S. (1995). *"Likelihood-Based Inference in Cointegrated Vector Autoregressive Models"*. Oxford University Press, New York.

Johnston, J. (1959). *"Statistical Cost Analysis"*. McGraw-Hill, New York.

Jorgensen, F., Pedersen, P.A. et Volden, R. (1997). "Estimating the Inefficiency of the Norwegian Bus Industry from Stochastic Cost Frontier Models". *Transportation*, 24, 421-433.

Juan, Z., Wu, J. et McDonald, M. (2003). "The Socio-economic Impacts Assessment of Advanced Convoy Driving on motorway". *Transportation Research Part A: Policy and Practice*, 37(9), 731-747.

Kamps, C. (2004). "New Estimates of Government Net Capital Stocks for 22 OECD Countries 1960-2001". *IMF Working Papers 04/67*, International Monetary Fund.

- Karlaftis, M.G. et McCarthy, P.S. (1997). "Subsidy and Public Transit Performance: A Factor Analytic Approach". *Transportation*, 24, 253-70.
- Karlaftis, M.G. (2003). "A DEA Approach for Evaluating the Efficiency and Effectiveness of Urban Transit Systems". *European Journal of Operational Research*, 152, 354-64.
- Kerstens, K. (1996). "Technical Efficiency Measurement and Explanation of French Urban Transit Companies". *Transportation Research Part A: Policy and Practice*, 30(6), 431-452.
- Kilkenny, M. (1998). "Transport Costs and Rural Development". *Journal of Regional Science*, 38, 293-312.
- Kneip, A., Park, B. et Simar, L. (1998) "A Note on the Convergence of Nonparametric DEA Efficiency Measures". *Econometric Theory*, 14, 783- 793.
- Koopmans, T.C. (1951). "Analysis of Production as an Efficient Combination of Activities". In T.C. Koopmans (Ed.) *Activity Analysis of Production and Allocation*, Wiley, New York London, 33-97.
- Kumbhakar, S.C. et Lovell, C.A.K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press, U.K.
- Lasserre, P. et Ouellette, P. (1994). "Factor Demands, Cost Functions, and Technology Measurements of Regulated Firms". *Canadian Journal of Economics*, 27, 218-242.
- Lau, L.J. (1986). "Functional Forms in Econometric Model Building", *Handbook of Econometrics*. 3, North-Holland Publishing Company, Amsterdam, 1515-1566.
- Lau, S.H.P. et Sin, C.Y. (1997). "Public Infrastructure and Economic Growth: Time Series Properties and Evidence". *Economic Record*, 73(221), 125-135.
- Linneker, B. et Spence, N. (1992). "Road Transport Infrastructure and Regional Economic Development: The Regional Development Effects of the M25 London Orbital Motorway". *Journal of Transport Geography*, 4(2), 77-92.
- Lovell, C.A.K. (1993). "Production Frontier and Productive Efficiency". In Lovell and Schmidt (Eds.) *The Measurement of Productive Efficiency* Edited by Harold, O., Fried, Lovell, C.A.K. et Schmidt, S.S., Oxford University Press, 3-67.
- Lovell, C.A.K. (2003). "The Decomposition of Malmquist Productivity Indexes". *Journal of Productivity Analysis*, 20, 437-458.

- Lucas, J. Robert, E. (1988). "On the Mechanics of Economic Development". *Journal of Monetary Economies*, 22, 3-42.
- Malmquist, S. (1953). "Index Numbers and Indifference Curves". *Trabajos de Estadística*, 4, 209-242.
- Mamatzakis, E.C. (1999a). "Public Infrastructure, Private Input Demand, and Economic Performance of the Greek Industry". *Working Papers* 406, Queen Mary, University of London, School of Economics and Finance.
- Marazzo, M., Scherre, R. et Fernandes, E. (2010) "Air transport demand and economic growth in Brazil: A time series analysis". *Transport Research*, 46(2), 261–269.
- Martin, J.C., Gutierrez, J. et Roman, C. (2004). "Data Envelopment Analysis (DEA) Index to Measure the Accessibility Impacts of New Infrastructure Investments: The Case of the High- speed Train Corridor Madrid-Barcelona-French Border". *Regional Studies*, 38(6), 697-712.
- Meeusen, W. et Van Den Broeck. (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Errors". *International Economic Review*, 18, 435-444.
- Morris, J.M., Dumble, P.L. et Wigan, M.R. (1978). "Accessibility Indicators for Transport Planning". *Transportation Research A*, 13, 91-109.
- Morrison, C.J. et Schwartz, A.E. (1996). "State Infrastructure and Productive Performance". *The American Economic Review*, 86, 1095-1111.
- Munnell, A. (1990a). "Why has productivity growth declined? Productivity and Public Investment". *New England Economic Review, Federal Reserve Bank of Boston*, 01, 3-22.
- Munnell, A. (1990b). "How Does Public Infrastructure Affect Regional Economic Performance?". *Conference Series, Federal Reserve Bank of Boston*, 69-112.
- Munnell, A. (1992). "Policy Watch: Infrastructure Investment and Economic Growth". *Journal of Economic Perspectives*, 6, 189-98.
- Munnell, A. (1993). "Les investissements d'infrastructure: évaluation de leurs tendances Actuelles et de leurs effets économiques". In *OCDE, Politiques d'Infrastructures pour les Années*, OCDE, Paris, 23-60.
- Nadiri, M.I. et Mamouneas, T.P. (1994). "The Effects of Public Infrastructure and R&D Capital on the Cost Structure and Performance of US Manufacturing Industries". *Review of Economics and Statistics*, 76(1), 22–37.

- Nerlove, M. (1963). "Returns to Scale in Electricity Supply". In Christ, C. (Ed.) *"Measurement in Economics"*. Stanford, Stanford University Press, 167-200.
- Nishimizu, M. et Page, J. M. (1982). "Total Factor Productivity Growth, Technological Progress and Technical Efficiency Change: Dimensions of productivity Change in Yugoslavia, 1965-1978". *Economic Journal*, 92(368), 920-936.
- Odeck, J. (2008). "The Effect of Mergers on Efficiency and Productivity of Public Transport Services". *Transportation Research Part A*, 42, 696–708.
- Odeck, J. et Alkadi, A. (2001). "Evaluating Efficiency in the Norwegian Bus Industry using Data Envelopment Analysis". *Transportation*, 28, 211-32
- Ouellette, P. et Petit, P. (2010) *"Mesure de l'Efficienc e des Établissements de Santé : Revue et Synthèse Méthodologique"* Centre sur la productivité et la prospérité, HEC Montréal
- Ouellette, P. et Vierstraete, V. (2002). "Technological Change and Efficiency in the Presence of Quasi-Fixed Inputs: A DEA Application to the Hospital Sector". *European Journal of Operational Research*, 154(3), 755-763.
- Ouellette, P. et Vigeant, S. (2001). "On the Existence of a Regulated Production Function". *Journal of Economics*, 73, 193-200.
- Ouellette, P. et Vigeant, S. (2003). "Technological Choices and Regulation: The Case of the Canadian Manufacturing Sectors". *Canadian Journal of Economics*, 36(1), 88-125.
- Ouellette, P., Quesnel, J.P et Vigeant, S. (2012). "Measuring Returns to Scale in DEA Models When Firm Is Regulated", *European Journal of Operational Research*, 220(2), 571-576.
- Pacheco, R.R. et Fernandes, E. (2003). "Managerial Efficiency of Brazilian Airports". *Transportation Research Part A: Policy and Practice*, 37, 667-280.
- Pacione, M. (1989). "Access to Urban Service-the Case of Secondary Schools in Glasgow". *Scottish Geographical Services*, 105, 12-18.
- Patton, T.A. et Clark, N. (1970). "Towards an Accessibility Model for Residential Development". *Tewkesbury Symposium*, University of Melbourne, 266-2103.
- Pereira, A. et Andraz, J. (2012). "On the Economic and Budgetary Effects of Investments in SCUTS: the Portuguese Toll-Free Highways". *The Annals of Regional Science*, Springer, 48(1), 321-338.

- Politis, D. N. et Romano, J. P. (1992). "A Circular Block-Resampling Procedure for Stationary Data". In LePage, R., Billard, L. (Eds.) *"Exploring the Limits of Bootstrap"*. New York, John Wiley, 263-270.
- Politis, D. N. et Romano, J. P. (1994). "The Stationary Bootstrap". *Journal of American Statistical Association*, 89, 1303-1313.
- Politis, D. N., Romano, J. P. et Wolf, M. (1999). *"Subsampling"*. New York: Springer.
- Porter, M.E. (1979). "The Structure within Industries and Companies' Performance". *The Review of Economics and Statistics*, 61, 214-27.
- Prebish, R. (1950). *"The Economic Development of Latin American and Some of its Principal Problems"* Lake Success, New York: United Nations.
- Prud'homme, R. (2004). "Infrastructure and Development". In Bourguignon, F. & B. Pleskovic (Eds.) *"Lessons of Experience (Proceedings of the Annual World Bank Conference on Development Economics)"*, Washington : World Bank and Oxford University Press, 153-180.
- Richmond, J. (1974). "Estimating the Efficiency of Production". *International Economic Review*, 15, 515-521.
- Romer, P.M. (1986). "Increasing Returns and Long-run Growth". *Journal of Political Economy*, 94, 1002-37.
- Rosenstein-Rodan, P. N. (1943). "Problems of Industrialisation of Eastern and South-Eastern Europe". *The Economic Journal*, 53(210/211), 202-21.
- Rostow, W. W. (1956). "The Take-Off Into Self-Sustained Growth". *The Economic Journal*, 66(261), 25-48.
- Royden, H. L. (1988). *"Real Analysis. 3rd Ed"*. Prentice Hall, Englewood Cliffs, New Jersey.
- SAS Institute Inc. (1989). *"SAS/OR 9.1 User's Guide: Mathematical Programming"*. SAS Institute Inc, SAS Publishing.
- Schmidt, P. (1975). "On the Statistical Estimation of Parametric Frontier Production Functions", *Review of Economics and Statistics*, 58, 238-239.
- Seitz, H. (1993). "The Economic Benefits of the Public Road Network". *The Annals of Regional Science*, 27, 223-239.
- Seitz, H. (2001). "Infrastructure Investment and Economic Activity: Theoretical Issues and International Evidence". In Deutsche Bundesbank (Ed.), *"Inventing"*

- Today for the World of Tomorrow: Studies on the Investment Process in Europe*", Springer Verlag, 85-124.
- Sengupta, J.T. (1995). "Estimating efficiency by cost frontiers: a comparison of parametric and nonparametric methods". *Applied Economics Letters*, 2(4), 86-90.
- Shao, J. et Tu, D. (1995). *The Jackknife and Bootstrap*. New York: Springer-Verlag.
- Shephard, R.W. (1953). *Cost and Production Functions*. Princeton University Press.
- Shephard, R.W. (1970). *Theory of Cost and Production Functions*. Princeton. Princeton University Press.
- Shephard, R.W. et Färe, R. (1974). "The law of diminishing returns". *Zeitschrift für Nationalökonomie*, 34, 69-90.
- Simar, L. et Wilson, P.W. (1998). "A General Methodology for Bootstrapping in Nonparametric Frontier Models". *Papers 9811*, Catholique de Louvain - Institut de Statistique.
- Simar, L. et Wilson, P.W. (1999). "Some Problems with the Ferrier/Hirschberg Bootstrap Idea". *Journal of productivity Analysis*, 11, 67-80.
- Simar, L. et Wilson, P.W. (2011). "Inference by the m out of n bootstrap in nonparametric frontier models". *Journal of Productivity Analysis*, 36, 33-53.
- Sims, C. A., Stock, J. H., et Watson, M. W., (1990). "Inference in Linear Time Series Models with some Unit Roots". *Econometrica* 58, 113-144.
- Small, K. A. (1992). *Urban Transportation Economics*. Harwood Academic Press, Chur, Switzerland.
- Solow, R. (1957). "Technical Change and the Aggregate Production Function". *Review of Economics and Statistics*, 39, 312-20.
- Stathopoulos, A. et Karlaftis, M. (2003). "A Multivariate State-Space Approach for Urban Traffic Flow Modeling and Prediction". *Transportation Research, Part C*, 11(2), 121-135.
- Stevenson, R. (1980). "Likelihood Functions for Generalized Stochastic Frontier Estimation". *Journal of Econometrics*, 13, 50-66.
- Stuart, A., Ord, J. K. et Arnold, S. (1999). *Kendall's Advanced Theory of Statistics*. Vol. 2A. Oxford University Press, New York

- Sun, Q., Tang, W. et Yu, Q. (2002). "Determinants of Foreign Direct Investment across China". *Journal of International Money and Finance*, 21(1), 79-113.
- Tinbergen, J. (1942). "On the Theory of Long-Term Economic Growth". *Weltwortschaftliches Archiv*, 55, 511-549.
- Timmer, C. P. (1971). "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency". *Journal of Political Economy*, University of Chicago Press, 79(4), 776-94.
- Viton, P.A. (1997). "Technical Efficiency in Multi-mode Bus Transit: A Production Frontier Analysis". *Transportation Research Part B*, 31 (1), 23–39.
- Viscusi, W.K., Vernon, J.M. et Harrington, J.E. Jr (2000). "*Economics of Regulation and Antitrust*". 3rd ed., The MIT Press, Cambridge.
- Wald, A. (1949). "Note on the consistency of the maximum likelihood estimate". *The Annals of Mathematical Statistics*, 20, 595–601.
- Yilmaz, S., Haynes, K. et Dinc, M. (2001). "The Impact of Telecommunications Infrastructure Investment on Sectoral Growth". *Australasian Journal of Regional Studies*, 7 (3), 383-397.
- Youngson, A. J. (1967). "*Britain's economic growth, 1920-1966*". New York, A.M. Kelley.
- Zellner, A. et Revankar, N.S. (1969). "Generalized Production Functions". *The Review of Economic Studies*, 36(2), 241-250.

Annexes

Annexe A. Les hypothèses de Wald (1949)

Soit X_1, X_2, \dots , etc., des vecteurs des variables aléatoires, indépendantes et identiquement distribuées. Le cas le plus souvent considérée dans la littérature est celui où la distribution commune est connue, sauf pour les valeurs d'un nombre fini des paramètres, $\theta^1, \theta^2, \dots, \theta^k$. Ici nous allons traiter le cas paramétrique. Pour un point quelconque, $\theta = (\theta^1, \theta^2, \dots, \theta^k)$, prenons $F(X, \theta)$ comme la fonction de répartition cumulative de X_i , soit $F(X, \theta) = \text{prob}\{X_i < \mathbf{X}\}$. L'ensemble Ω est nommé l'espace des paramètres. Donc Ω est un sous-ensemble de l'espace cartésien de dimension k .

Hypothèse 1. $F(X, \theta)$ est absolument continue pour tous les θ .

Quelques notations, Soit, pour une fonction de densité $f(y, \theta)$ et un vecteur de variables explicatives X_i , pour un θ et un ρ donnés : $f(y, \theta | \rho) := \sup_{\theta'} f(y, \theta')$ avec $\|\theta - \theta'\| \leq \rho$. Pour un θ et un r donnés : $\varphi(y, r) := \sup_{\theta} f(y, \theta)$. avec $\|\theta\| \geq r$. En plus, notons :

$$f^*(y, \theta | \rho) = \begin{cases} f(y, \theta | \rho) & \text{si } f(y, \theta | \rho) > 1 \\ 1 & \text{si } f(y, \theta | \rho) \leq 1 \end{cases} \quad \text{et} \quad \varphi^*(y, r) = \begin{cases} \varphi(y, r) & \text{si } \varphi(y, r) > 1 \\ 1 & \text{si } \varphi(y, r) \leq 1 \end{cases}$$

Hypothèse 2. Pour ρ suffisamment petit et r suffisamment grand, les espérances de

$$\int_{-\infty}^{+\infty} \ln f^*(X, \theta | \rho) dF(X, \theta_0) \quad \text{et} \quad \int_{-\infty}^{\infty} \log \varphi^*(X, \gamma) dF(X, \theta_0)$$

sont finies, où θ_0 représente la vraie valeur du paramètre.

Hypothèse 3. Si $\lim_{i \rightarrow \infty} \theta_i = \theta$, alors $\lim_{i \rightarrow \infty} f(X, \theta_i) = f(X, \theta)$ pour tous les x .

Hypothèse 4. Si θ_1 , un paramètre différent du vrai point paramétrique θ_0 , alors on a $F(X, \theta_1) \neq F(X, \theta_0)$ pour au minimum une valeur de X .

Hypothèse 5. Si $\lim_{i \rightarrow \infty} \|\theta_i\| = \infty$, alors, $\lim_{i \rightarrow \infty} f(X, \theta_i) = 0$ pour tous les X .

Hypothèse 6. Pour le vrai paramètre θ_0 , nous avons :

$$\int_{-\infty}^{\infty} |\ln f(X, \theta_0)| dF(X, \theta_0) \leq M < \infty$$

Hypothèse 7. L'espace des paramètres Ω est un sous-ensemble fermé de l'espace cartésien de dimension k .

Hypothèse 8. $f(X, \theta | \rho)$ est une fonction mesurable de X pour toutes les valeurs de θ et ρ .

Théorème de Convergence de Wald (1949). Soit $\bar{\theta}(x_1, \dots, x_n)$, une fonction des observations, x_1, \dots, x_n , telle que :

$$\frac{f(x_1, \bar{\theta}_n) \cdots f(x_n, \bar{\theta}_n)}{f(x_1, \theta_0) \cdots f(x_n, \theta_0)} \geq c > 0 \quad (\text{A-1})$$

pour tous les n et pour tous les x_1, \dots, x_n . Si les hypothèses 1 à 8 tiennent, alors :

$$\text{prob} \left\{ \lim_{n \rightarrow \infty} \bar{\theta}_n = \theta_0 \right\} = 1 \quad (\text{A-2})$$

Dans la mesure où l'équation (A-1) a une solution (l'estimateur existe) pour $c = 1$, alors $\hat{\theta}(x_1, \dots, x_n)$ est un estimateur du maximum de vraisemblance. Donc dans ces conditions, l'estimateur du maximum de vraisemblance est convergent en probabilité.

Annexe B. Les indices de prix utilisé pour calculer l'indice de Laspeyres

Tableau Annexe B-1 : Tableau entrées-sorties pour le transport routier de voyageur

BRANCHES	Valeur
FABRICATION DE PRODUITS CERAMIQUES ET DE MATERIAUX DE	6
INDUSTRIE DU CAOUTCHOUC	65
SERVICES INDUSTRIELS DU TRAVAIL DES METAUX	9
FABRICATION DE PRODUITS METALLIQUES	6
COMMERCE ET REPARATION AUTOMOBILE	30
COMMERCE DE GROS INTERMEDIAIRES	18
MANUTENTION ENTREPOSAGE ET GESTION D'INFRASTRUCTURE	296
AGENCES DE VOYAGE	57
ORGANISATION DU TRANSPORT DE FRET ET TRANSPORT SPATIAL	28
INTERMEDIATION FINANCIERE	106
ASSURANCES	69
LOCATION IMMOBILIERE	104
TELECOMMUNICATION, POSTE ET COURRIER	46
ACTIVITES INFORMATIQUES	76
SERVICES PROFESSIONNELS	87
ADMINISTRATION D'ENTREPRISES	98
LOCATION SANS OPERATEUR	226
SELECTION ET FOURNITURE DE PERSONNEL	132
SECURITE NETTOYAGE ET SERVICES DIVERS AUX ENTREPRISES	227
HOTELS ET RESTAURANTS	87
EDUCATION MARCHANDE	137
Total des produits	1,910

Tableau Annexe B-2 : Indices de prix utilisés

Indice de prix de la production	Décembre	Décembre	Décembre	Décembre	Décembre
	2000	2001	2002	2003	2004
Produits céramiques et matériaux de construction	100	103,8	106,4	108,1	111,3
Produits en caoutchouc	100	101,114	101,695	100,769	100,660
Produits du travail des métaux	100	101,9	102,7	102,9	103,5
Produits métallurgiques	100	99,355	96,811	96,628	106,705
Commerce et réparation automobile	100	100,9	103,8	108,5	112,3
Commerce de gros, intermédiaires	100	104,3	104,8	108,4	109,7
Services auxiliaires de transports	100	101,5	103,1	103,3	105,1
Intermédiation financière	100	96,5	99,8	97,9	96,6
Assurances	100	101,603	104,309	107,715	111,824
Location immobilière	100	101,5	104,4	107,2	110,6
Postes et télécommunications	100	96,5	95,1	95,6	96,3
Activités Informatiques	100	97,223	94,897	90,909	88,251
Services Professionnel	100	100,459	103,904	107,463	110,677
Administration d'entreprise	100	102,1	106,0	108,7	111,0
Location Sans Operateur	100	104,360	106,671	106,353	105,222
Sélection et Fourniture de Personnel	100	102,6	105,0	107,3	109,5
Sélection et Fourniture de Personnel	100	102,6	105,0	107,3	109,5
Hôtels et Restaurants	100	102,6	107,2	109,9	112,5
Éducation	100	103,2	108,2	110,9	111,4

Annexe C. Les résultats pour la fonction production avec des inputs agrégés

Tableau Annexe 31: Coefficients estimés de la fonction de production avec inputs agrégés sous forme Cobb-Douglas

Primal	Coefficient estimé	Biais d'estimation	Coefficient corrigé	Limites de confiance pour un niveau de 95 %	
				Borne supérieure	Borne inférieure
Programme linéaire					
α_L	0,477	0,007	0,470	0,286	0,564
α_K	0,338	-0,029	0,367	0,263	0,591
α_V	0,023	0,016	0,008	-0,067	0,047
α_{Inf}	0,615	-0,158	0,774	0,601	1,196
Programme quadratique					
α_L	0,469	0,002	0,467	0,309	0,520
α_K	0,339	0,000	0,339	0,297	0,544
α_V	0,035	-0,001	0,037	-0,027	0,071
α_{Inf}	0,674	-0,007	0,680	0,555	0,951

Tableau Annexe 42: Coefficients estimés de la fonction de production avec inputs agrégés sous forme Translog

Primal	Coefficient estimé	Biais d'estimation	Coefficient corrigé	Limites de confiance pour un niveau de 95 %	
Programme linéaire					
α_L	0,430	0,175	0,255	-0,817	0,731
α_K	-0,196	-0,143	-0,053	-0,641	0,556
α_V	0,058	0,052	0,006	-0,587	0,381
α_{Inf}	1,421	-0,186	1,608	0,857	2,707
β_L^2	0,187	-0,073	0,260	0,111	0,549
β_K^2	0,027	0,024	0,003	-0,128	0,123
β_V^2	0,017	0,005	0,011	-0,041	0,054
β_{INF^2}	-0,313	0,065	-0,378	-0,678	-0,210
β_{L-K}	-0,036	0,003	-0,039	-0,223	-0,024
β_{L-V}	0,030	0,005	0,025	-0,018	0,061
β_{L-Inf}	-0,111	-0,008	-0,103	-0,206	-0,029
β_{K-V}	-0,019	-0,002	-0,017	-0,038	0,047
β_{K-Inf}	0,133	0,004	0,128	0,053	0,245

β_{V_Inf}	-0,010	0,000	-0,010	-0,048	0,094
<hr/>					
Programme quadratique					
α_L	0,836	0,581	0,255	-0,735	0,877
α_K	-0,469	-0,416	-0,053	-0,577	0,667
α_V	0,131	0,125	0,006	-0,528	0,457
α_{Inf}	0,716	-0,892	1,608	0,557	1,760
β_{L^2}	0,160	-0,100	0,260	0,099	0,494
β_{K^2}	0,106	0,103	0,003	-0,115	0,135
β_{V^2}	0,028	0,017	0,011	-0,037	0,059
β_{Inf^2}	-0,100	0,078	-0,178	-0,271	-0,051
β_{L_K}	-0,039	0,000	-0,039	-0,089	-0,006
β_{L_V}	0,034	0,009	0,025	-0,020	0,055
β_{L_Inf}	-0,186	-0,083	-0,103	-2,245	-0,032
β_{K_V}	-0,044	-0,027	-0,017	-0,071	0,053
β_{K_Inf}	0,119	-0,009	0,128	0,060	0,232
β_{V_Inf}	-0,002	0,008	-0,010	-0,091	0,080
<hr/>					

Tableau Annexe C-3 : Efficacité technique de la fonction de production avec inputs agrégés

	Fonction de production Cobb-Douglas		Fonction de production Translog	
	Programme linéaire	Programme quadratique	Programme linéaire	Programme quadratique
θ	0,834	0,835	0,842	0,847
\dot{N}		76		45
θ'	5 % plus petites	0,785	0,788	0,789
	5 % plus grandes	0,842	0,844	0,853

Tableau Annexe C-5 : Élasticités de l'output par rapport aux inputs et les rendements d'échelles de la fonction de production avec inputs agrégés

		ξ_L	ξ_K	η_V	η_{INF}	η_{sans}^E	η_{avec}^E
Cobb-Douglas	Programmation linéaire	0,477	0,338	0,023	0,615	0,838	1,453
	Programmation quadratique	0,469	0,339	0,035	0,674	0,843	1,517
Translog	Programmation linéaire	0,583	0,266	0,086	0,403	0,935	1,338
	Programmation quadratique	0,592	0,275	0,098	0,441	0,965	1,406

Tableau Annexe C-6 : Intervalle de confiance pour les rendements d'échelles de la fonction de production avec inputs agrégés

Cobb-Douglas		Coefficient estimé	Biais d'estimateur	Coefficient corrigé	$I_{95\%}$	
					Borne inférieure	Borne supérieure
Programme linéaire	η_{avec}^E	1,453	-0,016	1,469	1,400	1,571
	η_{sans}^E	0,838	0,002	0,836	0,757	0,912
Programme quadratique	η_{avec}^E	1,517	0,116	1,401	1,188	1,651
	η_{sans}^E	0,843	0,006	0,837	0,762	0,928

Translog		Coefficient estimé	Biais d'estimateur	Coefficient corrigé	$I_{95\%}$	
					Borne inférieure	Borne supérieure
Programme linéaire	η_{avec}^E	1,338	0,018	1,320	1,134	1,493
	η_{sans}^E	0,935	-0,003	0,965	0,904	1,062
Programme quadratique	η_{avec}^E	1,406	-0,025	1,431	1,351	1,534
	η_{sans}^E	0,965	-0,039	1,004	0,946	1,032