



HAL
open science

Modélisation de la croissance pro-pauvre

Ndéné Ka

► **To cite this version:**

Ndéné Ka. Modélisation de la croissance pro-pauvre. Economies et finances. Université Montpellier, 2016. Français. NNT : 2016MONTD038 . tel-01549528

HAL Id: tel-01549528

<https://theses.hal.science/tel-01549528>

Submitted on 28 Jun 2017

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

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

THÈSE

Pour obtenir le grade de
Docteur

Délivré par l'Université de Montpellier

Préparée au sein de l'école doctorale EDEG*
Et de l'unité de recherche LAMETA

Spécialité: **Sciences Économiques**

Présentée par **Ndéné KA**

Modélisation de la Croissance pro-pauvre

Soutenue le 05/12/2016 devant le jury composé de

| | | | |
|----------------------|------------|----------------------|-----------------------|
| Mme.FRANÇOISE SEYTE | MCU HDR | Univ. de Montpellier | Directrice de thèse |
| M. STÉPHANE MUSSARD | Professeur | Univ. de Nîmes | co-Directeur de thèse |
| M. FRANÇOIS BENHMAD | MCU HDR | Univ. de Montpellier | Examineur |
| M. PHILIPPE VAN KERM | DR | LISER, Luxembourg | Rapporteur |
| M. BENOIT MULKAY | Professeur | Univ. de Montpellier | Examineur |
| M. ALEXANDRU MINEA | Professeur | Univ. d'Auvergne | Rapporteur |



L'université n'entend donner aucune approbation ni improbation aux opinions émises dans la thèse ; ces opinions doivent être considérées comme propres à leurs auteurs.

VU et PERMIS D'IMPRIMER



A Montpellier, le

Le Président de l'Université de Montpellier

Philippe Augé

Remerciements

Je tiens à remercier tous ceux sans qui cette thèse n'aurait jamais vu le jour. En tout premier lieu à :

- Madame Françoise Seyte qui fit le pari il y a quelques années de diriger cette thèse. Sa disponibilité, sa rigueur scientifique, ainsi que sa bienveillance discrète ont été déterminantes dans l'accomplissement de ce travail. Je tiens à lui témoigner ici toute ma gratitude.
- Monsieur Mussard Stéphane, qui m'a d'abord fait découvrir la régression Gini. Mais aussi, qui a bien voulu accepter de m'accompagner dans cette entreprise de recherche si exaltante. Par ses conseils avisés et ses remarques pertinentes, il m'a permis d'améliorer la qualité de ce travail. J'ai apprécié la disponibilité, le soutien constant dont il m'a fait preuve et qui m'ont donné la possibilité et la volonté d'avancer. Qu'il trouve ici l'expression de ma gratitude. J'ai beaucoup apprécié le fait de travailler avec lui.
- Mes sincères remerciements vont à Monsieur Philippe Van Kerm et Monsieur Alexandru Minea pour m'avoir fait l'honneur de rapporter cette thèse.
- J'exprime ma gratitude à Monsieur Benoit Mulkay et Monsieur François Benhmad qui m'ont fait l'honneur d'examiner ce travail.

Cette thèse a été réalisée au sein du Laboratoire Montpellierain d'économie théorique et appliquée(LAMETA) où régnait une atmosphère conviviale. Merci à : Ayad Assoil, Olfa Bouallegue, Nour Wehbi, Peguy Ndodjang, Mht Hisseine Saad, Boumediene Souiki, Diogoye Faye, Zoul Moumouni, Abdou salam Diallo, Mamadou Diop, Peguy, Maïté Stéphan ainsi que tous les doctorants du LAMETA pour avoir tous contribué à ma construction.

A mes parents, pour m'avoir éduqué et transmis la meilleure part d'eux même. A ma mère, pour m'avoir soutenu de manière indéfectible durant toutes ces années. A mes frères et sœurs.

A Aldiouma mme KA, pour avoir été mon compagnon de route durant toutes ces années, pour avoir supporté avec patience les externalités négatives de ce travail qui n'aurait certainement pas pu arriver à terme sans son soutien, des remerciements sont très peu de chose.

Il me faut aussi remercier le secrétariat du World income Inequality database de nous avoir fourni plusieurs versions de leur base de données.

Toutes ces personnes, à des degrés variables, ont contribué, directement ou

indirectement, au contenu de ce document, sans que leur responsabilité puisse être engagée en ce qui concerne les erreurs qui pourraient s'y trouver. Je reste, bien évidemment, seul responsable des manquements qui subsisteraient.

SOMMAIRE

Remerciements

Notations

Introduction

Chapitre I : Croissance pro-pauvre : définitions et mesures.

Chapitre II : Méthodologie Gini

Chapitre III : Estimation de la croissance pro-pauvre.

Conclusion Générale

Annexes

Notations

ARCH : AutoRegressif Conditional Heteroscedastixity.

BPC : Biais de pauvreté de la croissance.

CIC : Courbe d'Incidence de la Croissance.

CL : la Courbe de Lorenz.

CNB : Cost of Basic Needs.

CCP : Courbe de croissance de la pauvreté.

Cov représente l'opérateur covariance.

D S : Deininger et Squire

\mathbb{E} représente l'opérateur espérance.

ERS : Economic Research Service.

Eduprim : Taux de scolarisation primaire.

Edusecon : Taux de scolarisation secondaire.

FEI : Food Energy Intake.

FGT : Foster, Greer et Torsbecke.

GMM-AB : L'estimateur d'Arrelano- Bond par la Méthode des Moments Généralisés.

MCO : Les Moindres Carrées Ordinaires.

OCDE : Organisation de coopération et de développement économiques.

PED : Pays en voies de developpement.

PIB : Produit Intérieur Brute.

ICPP : Indice de Croissance Pro-Pauvre.

USDA : United States Department of Agriculture.

R_+ : Ensemble des réels positifs.

TCEP : Taux de Croissance Équivalent à la Pauvreté

Var : représente l'opérateur Variance.

VI : Variable Instrumentale.

WIDER : World Institud for Develpment Research

WIID : World Income Inequality Database.

Table des matières

| | |
|--|-----------|
| Remerciements | 4 |
| Sommaire | 6 |
| Notations | 7 |
| Introduction | 16 |
| 1 La croissance pro-pauvre : une revue de la littérature | 22 |
| 1.1 Pauvreté et croissance économique | 26 |
| 1.1.1 Définitions de la croissance pro-pauvre | 27 |
| 1.1.2 Les mesures agrégées de la croissance pro-pauvre | 30 |
| 1.1.3 Les mesures relatives de la croissance pro-pauvre | 35 |
| 1.1.4 Les mesures absolues de la croissance pro-pauvre | 41 |
| 1.1.5 L'approche économétrique de la croissance pro-pauvre | 43 |
| 1.1.6 Les politiques économiques pro-pauvres | 45 |
| 1.2 Les inégalités de revenu nuisibles ou favorables à la croissance | 49 |
| 1.2.1 Les approches théoriques | 50 |
| 1.2.2 Les approches empiriques | 54 |
| 1.3 Problèmes méthodologiques et qualités des données | 64 |
| 1.3.1 La causalité inverse entre inégalités et croissance | 65 |
| 1.3.2 La qualité des données | 66 |
| 1.3.3 L'hétérogénéité des données sur les inégalités de revenu | 67 |
| 1.4 Conclusion | 70 |
| 2 La méthodologie Gini | 72 |
| 2.1 Généralités sur la méthodologie de Gini | 74 |
| 2.1.1 La différence moyenne de Gini et Co-Gini | 74 |
| 2.1.2 Corrélation au sens de Gini : la G-corrélation | 78 |
| 2.1.3 Décomposition de la différence moyenne du Gini | 80 |

| | | |
|----------|--|------------|
| 2.2 | La régression Gini | 84 |
| 2.2.1 | Sensibilité de l'estimateur des MCO aux transformations monotones et aux points aberrants | 85 |
| 2.2.2 | L'approche paramétrique | 87 |
| 2.2.3 | L'approche semi-paramétrique | 89 |
| 2.2.4 | Similarités et différences entre la régression Gini et les MCO | 92 |
| 2.2.5 | Les conditions de "Grenander modifiées" | 94 |
| 2.2.6 | Traitement des ex aequo | 96 |
| 2.2.7 | La régression Gini-VI | 97 |
| 2.2.8 | Inférence et variance des estimateurs | 99 |
| 2.3 | L'estimateur Gini en Données de Panel | 102 |
| 2.3.1 | L'estimateur Gini du modèle à effets fixes | 104 |
| 2.3.2 | Inférence sur l'estimateur intragroupe | 106 |
| 2.3.3 | Simulations | 108 |
| 2.3.4 | Test de l'existence de spécificités individuelles | 111 |
| 2.4 | Les perturbations non sphériques | 115 |
| 2.4.1 | Test de Shelef (2014) | 116 |
| 2.4.2 | Extension du test de Shelef (2014) | 119 |
| 2.4.3 | L'estimateur Gini du modèle à effets aléatoires | 122 |
| 2.5 | Conclusion | 125 |
| 3 | Estimation de la croissance pro-pauvre | 127 |
| 3.1 | Limites de l'approche par les indices de croissance pro-pauvre . . . | 129 |
| 3.1.1 | Croissance pro-pauvre en Thaïlande avec la CIC. | 130 |
| 3.1.2 | Croissance pro-pauvre en Thaïlande avec la CPP. | 132 |
| 3.1.3 | Croissance pro-pauvre en Thaïlande avec le TCEP. | 133 |
| 3.1.4 | Croissance pro-pauvre au Thaïlande avec l'ICPP. | 134 |
| 3.2 | Analyses économétriques de la croissance pro-pauvre | 136 |
| 3.2.1 | Spécification de notre modèle | 137 |
| 3.2.2 | Les données | 138 |
| 3.2.3 | Analyse et interprétation des résultats | 141 |
| 3.2.4 | Estimations de notre modèle sur notre second échantillon . | 147 |
| 3.2.5 | D'autres méthodes d'estimations | 151 |
| 3.3 | Conclusion | 156 |
| | Conclusion Générale | 158 |

| | |
|--|-----|
| Annexes | 173 |
| 3.3.1 Quelques U-statistiques | 173 |
| 3.3.2 Similarités et différences entre la régression Gini et les MCO | 173 |
| 3.3.3 Décomposition de l'Aitken Gini | 175 |
| 3.3.4 Estimation de notre modèle sur les différents quintiles MCO | 178 |

Liste des tableaux

| | | |
|------|---|-----|
| 1.1 | Impact des transferts publics sur la répartition des revenus | 49 |
| 1.2 | les résultats empiriques des modèles d'économie politique | 55 |
| 1.3 | Analyses empiriques des relations inégalité / instabilité et instabilité / croissance | 57 |
| 1.4 | Synthèse de quelques résultats empiriques | 64 |
| 2.1 | Estimateur Gini Global $\hat{\beta}^G$ | 109 |
| 2.2 | Estimateur Gini intergroupe $\hat{\beta}^{BG}$ | 109 |
| 2.3 | Estimateur Gini intragroupe $\hat{\beta}^{WG}$ | 110 |
| 2.4 | Estimateur Gini intragroupe : outliers positifs | 110 |
| 2.5 | Estimateur Gini intragroupe : outliers négatifs | 110 |
| 2.6 | Distribution non normale : sans outlier | 111 |
| 2.7 | Perte de puissance du Test de Shelef | 119 |
| 2.8 | Correction du Test de Shelef | 121 |
| 3.1 | Incidence de la pauvreté en Thaïlande 1990-2000 | 130 |
| 3.2 | Taux de Croissance Équivalent à la Pauvreté en Thaïlande 1990-2000 | 133 |
| 3.3 | Indice de croissance pro-pauvre en Thaïlande 1990-2000 | 134 |
| 3.4 | L'évolution des inégalités en Thaïlande 1990-2000 | 135 |
| 3.5 | Statistiques descriptives | 139 |
| 3.6 | Tests d'outliers | 143 |
| 3.7 | Estimation de notre modèle par les MCO | 145 |
| 3.8 | Estimation de notre modèle par le Gini | 146 |
| 3.9 | Estimation de notre modèle par les MCO | 149 |
| 3.10 | Estimation de notre modèle par le Gini | 150 |
| 3.11 | Estimation de notre modèle sur le premier quintile MCO | 152 |

| | | |
|------|--|-----|
| 3.12 | Estimation de notre modèle sur le premier quintile Gini | 153 |
| 3.13 | Tests d'endogénéité | 154 |
| 3.14 | Quelques U-statistiques | 173 |
| 3.15 | Estimation de notre modèle sur les différents quintiles MCO . . . | 178 |
| 3.16 | Estimation de notre modèle sur les différents quintiles MCO . . . | 179 |
| 3.17 | Estimation de notre modèle sur les différents quintiles Gini . . . | 180 |
| 3.18 | Estimation de notre modèle sur les différents quintiles Gini . . . | 181 |

Université de Montpellier

Faculté d'Économie

Modélisation de la croissance pro-pauvre

Résumé : Cette thèse contribue à l'approche économétrique de la croissance pro-pauvre. Elle présente des apports théoriques et empiriques. En premier lieu, elle présente les différentes définitions, indices et politiques de croissance pro-pauvre proposées dans la littérature théorique. Elle examine également les modèles théoriques et empiriques portant sur les interactions entre distribution du revenu et croissance. Elle montre que les mesures traditionnelles, en plus de leurs caractères partiels, peuvent conduire à des résultats contradictoires. Pour contourner ces limites, cette thèse privilégie l'approche alternative qui consiste à utiliser des modèles économétriques. Cette dernière approche, bien qu'elle présente l'avantage d'inclure l'ensemble des dimensions de la pauvreté, souffre de deux types de biais : le biais de sélection et le biais d'endogénéité. Ces derniers s'expliquent par les limitations inhérentes aux données : erreurs de mesures, points aberrants. En outre, les résultats obtenus avec cette approche sont sensibles aux formes fonctionnelles choisies. Ainsi, il y a de bonnes raisons d'utiliser la régression Gini. Cependant, les régressions de type Gini n'existaient qu'en coupe instantanée et en séries temporelles. Ainsi, dans un second temps, cette thèse propose d'étendre la réflexion sur la régression Gini en panel. Elle introduit les estimateurs intragroupes, intergroupes, le test d'existence de l'effet individuel et l'estimateur Aitken Gini. Enfin, cette thèse présente des applications empiriques qui illustrent de façon concrète la robustesse de nos estimateurs. Elle s'intéresse particulièrement aux conséquences de la méthode d'estimation et à la section de l'échantillon. Elle conclut que le processus de croissance favorise la réduction de la pauvreté à condition que les inégalités de revenu soient maîtrisées. Mais aussi, que l'impact de la croissance agricole sur la réduction de la pauvreté varie en fonction du niveau de développement du pays.

Pro-poor growth modelling

Abstract : This thesis contributes to the econometric approach of pro-poor growth. It presents theoretical and empirical contributions. First, it brings out the different definitions, indices and the policies of pro-poor growth proposed in the theoretical literature. It also examines the theoretical and empirical interactions between income distribution and growth. It shows that the traditional measures, in addition to their partial characteristics, can lead to contradictory results. To avoid these limits this thesis lays the emphasis on the alternative approach by using econometric models. The latter approach, although it has the advantage of including all the dimensions of poverty, suffers from two types of bias : selection bias and bias of endogeneity. These are due to the limitations of the data : measurement errors, outliers. In addition, the results obtained with this approach are sensitive to the selected functional forms. So, there are good reasons to use the Gini regression. However, the Gini regressions existed only for cross sectional data and time series. Thus, in a second time, this thesis proposes to extend the Gini regression for panel data. It introduces within and between-group estimators, the individual effect test and the Gini Aitken estimator. Finally, this thesis presents empirical applications that illustrate the robustness of our estimators. It is particularly interested in the consequences of the estimation method and the sample selection. It concludes that the growth process promotes poverty reduction when income inequalities are overcome. But also, the impact of agricultural growth on poverty reduction varies depending on the country's level of development.

Discipline : Sciences Économiques.

Mots-clés : Croissance pro-pauvre, inégalité de revenu, régression Gini, \mathbb{U} -statistique.

Introduction

Les inégalités de revenu ont pendant longtemps été considérées comme un problème passager qui disparaîtrait tout naturellement avec le processus de croissance économique. Par conséquent, il n'était pas question de mener des politiques sociales coûteuses qui risqueraient d'engendrer des distorsions fiscales ; pour cela, on pouvait compter seulement sur l'accumulation du capital. Kuznets est certainement l'un des premiers auteurs à formaliser cette idée. Déjà, au milieu des années 1950, il affirmait que les inégalités et la croissance entretiennent une relation en forme d'une fonction en « U » inversée. En effet, selon lui, les inégalités augmentent dans un premier temps avec le processus de croissance. Ensuite, elles retombent avec le développement économique. Cependant, loin d'observer ce phénomène, les inégalités continuent de persister malgré des périodes de forte croissance économique. Davantage d'individus se considèrent aujourd'hui marginalisés.

Pourtant, la montée des inégalités a toujours été au cœur des préoccupations des pouvoirs publics. Par exemple, George Bush affirmait en 2007 lors de l'une de ses allocutions que "Nos citoyens déplorent que notre économie dynamique laisse à la traîne certains d'entre eux. [...] l'inégalité des revenus est indéniablement ; elle s'accroît depuis plus de 25 ans". Toutefois, malgré l'importance de cette question et l'intérêt qu'elle a suscité, décideurs et analystes n'ont qu'une connaissance imparfaite du sujet. Pourtant, le véritable enjeu pour mener une bonne politique de développement passe nécessairement par la compréhension des interactions entre distribution du revenu et croissance économique.

Depuis les travaux de Kuznets (1955), de nombreux modèles théoriques ont été développés pour rendre compte de l'impact des inégalités sur le processus de développement. Un premier courant de pensée suggère l'existence d'une relation positive entre les inégalités de revenu et la croissance. Deux arguments principaux fondent ce courant. Le premier s'appuie sur le modèle de croissance de Kaldor selon lequel la propension marginale à épargner (Pms) est plus forte chez les riches que chez les pauvres. Le second s'appuie sur le caractère indivisible de l'investissement. Dans un contexte de marchés financiers imparfaits, une concentration de la richesse serait nécessaire pour financer les coûts fixes liés à la promotion de nouvelles activités industrielles et d'innovations technologiques Galor and Tsiddon (1997). De nombreuses critiques ont été formulées par

la suite à leur rencontre. La vision dominante serait aujourd'hui plutôt celle d'un impact négatif des inégalités sur le taux de croissance. Une première justification est d'ordre socio-politique (*socialpolitical unrest*). De fortes inégalités de revenu peuvent engendrer des mouvements de contestation sociale, des activités criminelles et donc une situation d'instabilité qui inhibe l'investissement. Par ailleurs, dans une société démocratique si le revenu moyen dépasse de loin le revenu médian un système de vote à la majorité simple serait favorable à une redistribution des ressources des riches vers les pauvres (transferts, dépenses publiques, etc.) ce qui pourrait conduire à des distorsions fiscales et détourner ainsi la classe riche de l'investissement. Ces modèles particulièrement connus sous le nom de modèles d'économie politique ont la particularité d'endogénéiser les politiques redistributives.

Les débats entamés depuis le début des années 1990 sur les alternatives possibles à la littérature sur les inégalités ont fait porter un nouveau regard sur le concept de croissance pro-pauvre (ou croissance inclusive). Il est important de noter qu'il est à la fois trompeur et incertain de croire qu'une croissance économique forte conduit forcément à une réduction des inégalités. Cette nouvelle conception, portée par des auteurs comme Anand and Kanbur (1993b), Bourguignon (2003), Kakwani (1993), Klasen (2005), discute des conditions sous lesquelles la croissance profite aux plus pauvres. Et comme le souligne Klasen (2005), cette approche ne tente nullement à réfuter l'idée selon laquelle la croissance réduit, à terme, la pauvreté. Bien au contraire, elle consiste à accroître l'influence de la croissance sur la réduction de la pauvreté. Promouvoir la croissance pro-pauvre revient donc à donner la priorité aux politiques ayant un impact favorable à la fois sur la croissance et sur la réduction des inégalités. Historiquement, on considère deux approches pour aborder ce thème : dans la première, dite relative, la croissance est pro-pauvre lorsqu'elle se manifeste par une réduction des inégalités en faveur des pauvres. La seconde approche, dite absolue, considère que la croissance est pro-pauvre lorsqu'elle s'accompagne d'une réduction du taux de pauvreté en terme absolu. Chacune de ces approches présente des insuffisances majeures et certains chercheurs ont tenté de proposer des approches alternatives.

Aujourd'hui, la question la plus importante et sans nul doute la plus difficile est celle de la mesure de la croissance pro-pauvre. Plus concrètement, comment identifier une croissance inclusive. De combien de point de pourcentage de crois-

sance les pauvres doivent-ils bénéficier afin de la qualifier de pro-pauvre ? Ces dernières années, les praticiens ont développé une multitude de mesures pour déterminer empiriquement l'impact de la croissance sur les plus démunis. Nous pouvons citer en guise d'exemple la courbe d'incidence de la croissance (CIC) de Ravallion and Chen (2003) ; le taux de croissance pro-pauvre de Ravallion and Chen (2003) ; la courbe de croissance de la pauvreté de Son (2004) ; le biais de pauvreté de croissance de McCulloch and Baulch (1999) ; l'indice de la croissance pro-pauvre de Kakwani et al. (2000) ; etc. Toutefois, malgré la grande diversité des indices, d'un point de vue empirique, un travail important est encore nécessaire sur cette question de quantification de la croissance pro-pauvre. En effet, à l'heure actuelle, il n'existe pas de mesure de la croissance pro-pauvre faisant l'unanimité. L'une des principales limites de ces indices est qu'ils se focalisent uniquement sur la dimension monétaire de la pauvreté. Ce qui peut paraître très restrictif bien que nous reconnaissons le lien entre la dimension monétaire et non monétaire de la pauvreté. En outre, en plus de leur caractère partiel, ces indices peuvent conduire à des résultats contradictoires. C'est pourquoi, ces années récentes ont vu naître une approche alternative pour quantifier le caractère pro-pauvre de la croissance. Cette seconde approche utilise généralement des modèles économétriques afin d'établir une causalité entre croissance pro-pauvre et indicateurs de bien-être. **Cette thèse s'inscrit dans la continuité de cette dernière approche. Plus précisément, l'objet de cette thèse est de développer des techniques d'estimation qui nous permettront de conclure avec plus de rigueur quant à la question du signe de cette relation.**

Globalement, bien que l'approche économétrique présente l'avantage d'inclure la dimension non monétaire de la pauvreté, elle est sujette à deux types de biais que sont le biais de sélection et le biais d'endogenéité. Ces derniers s'expliquent principalement par les limites inhérentes aux données. En effet, malgré les améliorations considérables de ces dernières années, les données sur la croissance pro-pauvre continuent de poser de sérieux problèmes aux chercheurs. Ces derniers déplorent le nombre élevé d'observations manquantes et la présence de points aberrants. La grande majorité des données est de piètre qualité et est sujette à d'importantes erreurs de mesure. Malgré les efforts visant à améliorer (ou élargir) les données existantes, elles restent insatisfaisantes. Les données sur la croissance pro-pauvre diffèrent entre les pays en termes de couverture géographique

(couverture nationale, urbaine ou rurale), d'unités statistiques (familles, ménages ou individus) et de définition de revenu (dépenses de consommation, revenu disponible ou revenu brut). Ces erreurs de mesure engendrent une sous-estimation, en valeur absolue, de la vraie valeur du paramètre de la variable sujette à cette erreur de mesure.

Par ailleurs, dans la plupart des travaux empiriques les échantillons sont composés d'une soixantaine de pays parmi lesquels les pays développés sont surreprésentés et les pays pauvres surtout ceux d'Afrique sub-saharienne quasiment inexistantes, ceci conduit à des échantillons très hétérogènes et à une sélection drastique. A cela s'ajoute, le fait que la dimension temporelle des données est extrêmement limitée. En conséquence, la variance inter-individuelle est largement supérieure à la variance inter-temporelle (Cogneau et al., 2002; White and Anderson, 2001; Ghura et al., 2002, entre autres). En outre, il serait nécessaire de tester la robustesse des conclusions en analysant les canaux de transmission et de voir aussi si les résultats restent valables dans d'autres échantillons et également avec des combinaisons différentes de variables explicatives. D'ailleurs, certains auteurs, (Barro, 2000; Ghura et al., 2002, entre autres), montrent la sensibilité des résultats au type d'échantillon et aux formes fonctionnelles retenues. L'une des particularités de notre travail est justement de proposer un modèle de correction de ces trois limites. A cet effet, nous utiliserons la régression Gini. Cette dernière introduit par Olkin and Yitzhaki (1992) présente plusieurs avantages : elle est moins sensible que les méthodes traditionnelles aux erreurs de mesures et aux points aberrants (outliers). Mais aussi, elle est plus robuste car les estimateurs obtenus par cette approche sont insensibles aux formes fonctionnelles retenues et au biais de sélection.

Formellement, la régression Gini consiste à utiliser la différence moyenne du Gini (GMD) comme mesure de dispersion. Cette dernière a été introduite en 1912 par Corrado Gini, et depuis, plusieurs opérateurs ou coefficients ont été déduits comme la covariance au sens de Gini (co-Gini), la corrélation au sens de Gini (G-corrélation), l'analyse du Gini (ANOGI). Comme nous le verrons plus loin la co-Gini n'est pas symétrique. Cela explique d'ailleurs le fait qu'il existe deux types de régression Gini. La première consiste à minimiser la GMD des erreurs. Les estimateurs obtenus par cette approche ont la propriété d'optimalité. La seconde approche consiste à construire l'estimateur en question de sorte qu'il soit moins

sensible aux outliers. Cette seconde approche, puisqu'elle ne découle pas d'un programme de minimisation et par conséquent ne requiert pas de définir une forme fonctionnelle quelconque, est qualifiée de régression semi-paramétrique. L'estimateur obtenu par cette dernière approche est similaire à celui obtenu par les MCO ; excepté que chaque variance est remplacée par la GMD et chaque covariance par la co-Gini appropriée voir Ka and Mussard (2015).

D'un point de vue théorique et empirique, un travail important est encore nécessaire sur la régression Gini notamment en panel. En effet, introduite par Olkin and Yitzhaki (1992), la réflexion sur la régression Gini a été étendue aux variables instrumentales par Yitzhaki and Schechtman (2004) et aux séries temporelles par (Shelef and Schechtman, 2011; Shelef, 2014; Carcea and Serfling, 2015, entre autres). *Nous proposons de l'étendre aux données de panel. C'est dans ce cadre que nous avons développé l'estimateur par effet fixe en utilisant la GMD comme mesure de dispersion Ka and Mussard (2016). Nous proposons également le test de l'existence de spécificités individuelles et l'estimateur du modèle à effets aléatoires.* L'intérêt essentiel de développer la régression Gini en panel réside dans la volonté de pouvoir exploiter au mieux les avantages spécifiques de ces données. En particulier, la possibilité de prendre en compte les outliers. Puisqu'un échantillon de données de panel contient un très grand nombre d'observations, on peut penser a priori que cet ensemble d'observations est capable de "neutraliser" l'influence de quelques points aberrants. En réalité, ce n'est pas vraiment le cas. Huber (1980) a montré que seulement 3% de valeurs atypiques dans un échantillon suffisent à perturber significativement la qualité des estimations. Plus généralement, dans le cadre de l'économétrie sur données de panel, et comme nous le verrons d'ailleurs dans le chapitre suivant, quelques points atypiques peuvent suffire à modifier considérablement les résultats d'une estimation.

Cette thèse est composée de trois chapitres structurés de la manière suivante.

Le premier chapitre est une revue de la littérature de la croissance pro-pauvre. Il est composé de deux sections. La première section s'intéresse principalement aux définitions. Aujourd'hui, il existe une multitude de définitions et de mesures de la croissance pro-pauvre. Et le lecteur non averti peut se perdre facilement dans cette vaste littérature. C'est pour cette raison que nous avons envisagé de rédiger cette section, dans l'objectif de faire une synthèse de toutes ces mesures et définitions. La seconde section examine les modèles théoriques et empiriques

portant sur les interactions entre distribution du revenu et croissance. Ainsi, il vise essentiellement à rendre accessible au lecteur cette vaste littérature. Il s'agit principalement d'un travail de présentation, d'éclaircissement et de critique des principaux courants impliqués dans ce débat.

Le second chapitre s'intéresse à la régression Gini. La première section de ce chapitre revient sur les généralités de la méthode Gini. Nous allons y présenter entre autres la différence moyenne du Gini, le co-Gini, la corrélation au sens de Gini et les différentes méthodes de décomposition de l'indice de Gini. La seconde section s'intéresse aux régressions Gini, à leurs propriétés et aux hypothèses sur la matrice des moments. Une remarque additionnelle sera également portée au traitement des ex-æquo. Dans la troisième section nous développons l'estimateur intragroupe et intergroupe au sens de Gini du modèle à effets fixes et le test d'existence de l'effet individuel et enfin la dernière section constitue nos derniers développements théoriques.

Le dernier chapitre constitue une application empirique. Il s'intéresse particulièrement aux limites des indices de la croissance pro-pauvre mais également aux conséquences de la méthode d'estimation et de la sélection de l'échantillon. Dans la première section de ce chapitre 3 nous proposons une étude comparative de quelques indices de la croissance inclusive. Plus précisément, nous allons étudier le caractère inclusif de la croissance Thaïlandaise en utilisant plusieurs indices. Ceci nous permettra de constater que ces derniers peuvent conduire à des résultats contradictoires. La seconde section s'intéresse aux conséquences de la limitation des données, des choix de spécification et à la méthode d'estimation. La sensibilité des estimateurs par effets fixes standards aux outliers, aux erreurs de mesures et aux formes fonctionnelles fait que nous privilégierons l'utilisation de notre estimateur intra-individuel (ou intragroupe) de la régression Gini. Cette dernière est plus robuste que l'estimateur intra basé sur les MCO Ka and Mussard (2016).

Nous allons maintenant aborder le premier chapitre de cette thèse sur la croissance pro-pauvre.

Chapitre 1

La croissance pro-pauvre : une revue de la littérature

Au début des années 1950, la littérature économique considérait que la croissance n'était favorable qu'aux riches Kakwani et al. (2000). En effet, ne disposant pas de capital humain ni financier, les pauvres ne pouvaient recevoir qu'une faible partie des bénéfices de la croissance (grâce à la redistribution notamment) : c'est la théorie du *trickle down*. Dans un tel contexte, nous pouvons bien comprendre qu'une croissance économique forte ne conduit pas nécessairement à une rapide baisse de la pauvreté. D'ailleurs, il est tout à fait envisageable qu'elle engendre une augmentation de la pauvreté : croissance appauvrissante.

Le début des années 1990 a vu naître un nouveau concept, celui de la croissance pro-pauvre. Cette dernière est apparue comme une alternative aux modèles de redistribution qui conduisent à une très faible réduction de la pauvreté. L'idée est ici, contrairement à la théorie du *trickle down* notons-le, de faire émerger la croissance à partir de la base (les pauvres), c'est-à-dire de mettre les pauvres au cœur du processus de création de richesse. En effet, comme le souligne Dollar and Kraay (2002), la croissance à elle seule est insuffisante pour engendrer une réduction significative de la pauvreté. Cette dernière passe par des stratégies de croissance¹ et des politiques distributives dont la combinaison pourrait être propre à chaque pays Bourguignon (2004). Par exemple, une croissance moyenne de 1% du revenu par habitant peut entraîner une réduction de la proportion des gens vivant dans une pauvreté extrême allant jusqu' à 4%, mais pouvant aussi être in-

1. volontairement orientées vers les pauvres

férieure à 1%, selon le pays et la période Ravallion (2004). Ceci nous amène à nous demander pourquoi la croissance ne profite pas toujours aux plus pauvres. Dans une économie décentralisée, les réponses à cette question résident dans le degré d'accès des pauvres aux marchés et plus généralement aux inégalités de revenu et sociales. Par conséquent, la prise en compte du lien entre la distribution des revenus et la croissance est cruciale pour toute politique tendant à assurer une croissance pro-pauvre.

Manifestement, le véritable enjeu pour mener une bonne politique de développement passe nécessairement par la compréhension des interactions entre distribution du revenu et croissance économique. En effet, comme le souligne Bourguignon (2004) et comme la réalité tend à le montrer, la réduction de la pauvreté réside beaucoup plus dans les interactions entre inégalités de revenu et croissance que dans les relations entre, d'une part, pauvreté et croissance et, d'autre part, pauvreté et inégalités, qui restent essentiellement arithmétiques. Une forte croissance économique, par exemple, a été le principal levier de réduction de la pauvreté en Extrême-Orient et en Asie du Sud-Est où la proportion des gens vivant dans la pauvreté est passée respectivement de 33% en 1990 à 9,9% en 2004 et de 41% à 29,5% (durant la même période). Par contre, en Afrique sub-saharienne où la croissance économique a été relativement faible, l'incidence de la pauvreté n'a diminué que légèrement, de 47% en 1990 à 41% en 2004 Wiggins and Higgins (2008). De même, plusieurs travaux tendent à montrer qu'une détérioration de la distribution des revenus favorise l'augmentation de la pauvreté. Toutefois, une question intéressante est de savoir si la distribution du revenu et la croissance sont étroitement liées ou si, au contraire, elles sont indépendantes? Depuis les travaux de Kuznets dans les années 1950, le débat sur cette question ne cesse d'être alimenté par de nouveaux arguments. Cette hypothèse de Kuznets, selon laquelle les inégalités de revenu augmentent dans les premières phases du processus de croissance puis retombent avec le développement économique, largement corroborée par les travaux empiriques antérieurs (Adelman and Morris, 1973; Ahluwalia, 1976; Papanek and Kyn, 1986) a très vite laissé place à la controverse. Tout d'abord, Anand and Kanbur (1993a) montrent que sur le plan empirique l'apparition de la courbe de Kuznets dépendrait en grande partie de la forme fonctionnelle retenue. Ensuite et surtout peut être, dans les années 1990, un nouveau fait stylisé issu de l'économétrie révèle une relation de causalité in-

verse entre les inégalités de revenu et la croissance. Ces travaux sont l'œuvre de Persson and Tabellini (1994). De nombreux modèles théoriques ont été élaborés pour rendre compte de cette régularité. Nous pouvons citer déjà à titre d'exemple les modèles d'économie politique Persson and Tabellini (1994) et Alesina and Rodrik (1994), les modèles d'imperfection du marché du crédit (Banerjee and Newman, 1993; Galor and Zeira, 1993), les modèles portant sur l'instabilité sociopolitique que pourraient engendrer de fortes inégalités de revenu Alesina and Perotti (1996) etc. Malheureusement, cette nouvelle relation, tout comme la courbe de Kuznets d'ailleurs, a très vite montré sa fragilité. En effet, ces premiers travaux ont été largement remis en cause par les analyses sur données de panel. Ces dernières ont trouvé une relation positive, non linéaire, voire nulle entre les inégalités de revenu et la croissance (Forbes, 2000; Barro, 2000; Banerjee and Duflo, 2003; Deininger and Squire, 1998). D'ailleurs c'est la raison pour laquelle cette nouvelle conception portée par des auteurs comme Anand and Kanbur (1993a), (Kakwani et al., 2000; Bourguignon, 2004; Klasen, 2004) discute des conditions sous lesquelles la croissance profite aux plus pauvres. Cependant, tout comme la pauvreté, il est difficile de définir ce nouveau concept. Schématiquement, il s'agit d'une croissance qui engendre une diminution significative de la pauvreté OCDE (2001)². Toute la difficulté tient au mot "significatif" ce qui conduit ainsi à plusieurs approches définitionnelles.

La croissance est manifestement favorable aux pauvres lorsqu'elle entraîne l'augmentation de leurs revenus. Toutefois, comme le souligne Lopez (2004), la question la plus importante et sans nul doute la plus difficile est celle de savoir de combien les pauvres doivent bénéficier de la croissance pour qu'elle soit qualifiée de pro-pauvre. Cette question s'intéresse donc à la mesure de la croissance pro-pauvre. Heureusement, les chercheurs ont proposé une multitude de mesures pour déterminer empiriquement l'impact de la croissance sur les plus démunis : la courbe d'incidence de la croissance (CIC) de Ravallion and Chen (2003) ; le taux de croissance pro-pauvre de Ravallion and Chen (2003) ; la courbe de croissance de la pauvreté de (Son, 2004) ; le biais de pauvreté de croissance de McCulloch and Baulch (1999) ; l'indice de la croissance pro-pauvre de Kakwani et al. (2000) ; etc. Ces indices étant largement connus, quelques brefs rappels sont seulement proposés dans ce chapitre. Toutefois, malgré la grande diversité des indices, d'un

2. (Pro-pauvre growth may be referred as growth that benefits the poor and provides them with opportunities to improve their economic situation OCDE (2001)).

point de vue empirique, un travail important est encore nécessaire sur cette question de quantification de la croissance pro-pauvre. En effet, à l'heure actuelle, il n'existe pas de mesure de la croissance pro-pauvre faisant l'unanimité. L'une des principales limites de ces indices est qu'ils se focalisent uniquement sur la dimension monétaire de la pauvreté, ce qui peut paraître très restrictif bien que nous reconnaissons le lien entre la dimension monétaire et non monétaire de la pauvreté. En outre, comme nous allons le voir plus loin dans cette thèse (chapitre 3), une réflexion en profondeur nous permet de constater que ces derniers peuvent aboutir à des conclusions contradictoires. C'est pourquoi, ces années récentes ont vu naître une approche alternative pour quantifier le caractère pro-pauvre de la croissance. Cette seconde approche utilise généralement des modèles économétriques afin d'établir une causalité entre croissance pro-pauvre et indicateurs de bien-être.

Dans l'objectif d'introduire cette réflexion sur la problématique de la croissance pro-pauvre et les limites des approches traditionnelles, ce chapitre propose une synthèse de la littérature sur les concepts de base. Ainsi, nous tenterons de recenser les définitions et principales mesures de la croissance pro-pauvre. Il comprend deux grandes sections.

La première section est consacrée aux concepts de base : définitions, mesures et politiques pro-pauvres. Plus précisément, dans cette première section nous présenterons les différentes définitions, indices et politiques de croissance pro-pauvre. Dans la seconde section, nous faisons une revue sélective des avancés théoriques et empiriques sur la réflexion dans la relation entre inégalités de revenu et croissance.

1.1 Pauvreté et croissance économique

La croissance pro-pauvre est au centre du débat économique ce qui témoigne de sa popularité ces dernières années même si sa définition reste difficile. Schématiquement, la croissance est qualifiée de pro-pauvre lorsqu'elle s'accompagne d'une réduction significative de la pauvreté OCDE (2001). Toutefois, cette définition est vaste et nous fournit peu d'informations. En effet, nous ne savons pas de combien les pauvres doivent bénéficier de la croissance pour que celui-ci soit qualifiée de pro-pauvre. Que signifie une diminution de la pauvreté ? Toutes ces interrogations font l'objet de divergence dans les définitions et conduisent à plusieurs approches quant à sa mesure. Certains favorisent une approche relative pour définir la croissance pro-pauvre, en considérant que la croissance est pro-pauvre lorsqu'elle s'accompagne d'une réduction des inégalités de revenu (White and Anderson, 2001; Klasen, 2004).³ D'autres préfèrent l'approche absolue qui définit la croissance pro-pauvre comme étant une croissance qui réduit le taux de pauvreté (Kakwani et al., 2000, 2002). Comme nous allons le voir chacune de ces approches présente des limites et certains chercheurs ont tenté de proposer des approches alternatives (exemple Osmani et al. (2005)). Toutefois, la question la plus importante est celle de savoir de combien l'incidence de la pauvreté doit diminuer pour que la croissance soit considérée comme étant pro-pauvre Lopez (2004). Cette question s'intéresse donc à la mesure de la croissance pro-pauvre. Deux grandes approches se sont dégagées pour mesurer une telle croissance. Une première se fonde sur des indices.⁴ Et la seconde utilise des modèles économétriques afin d'établir une causalité entre croissance pro-pauvre et indicateurs de bien-être. Ainsi, cette section établit un rapprochement de ces deux méthodes afin de pouvoir dégager leurs limites. Mais, nous exposerons dans un premier temps les différentes approches retenues dans la littérature en termes de définition. Ceci semble être indispensable car quelle que soit la méthode utilisée, elle reste largement tributaire de la définition considérée. Nous nous intéresserons également aux politiques pro-pauvres.

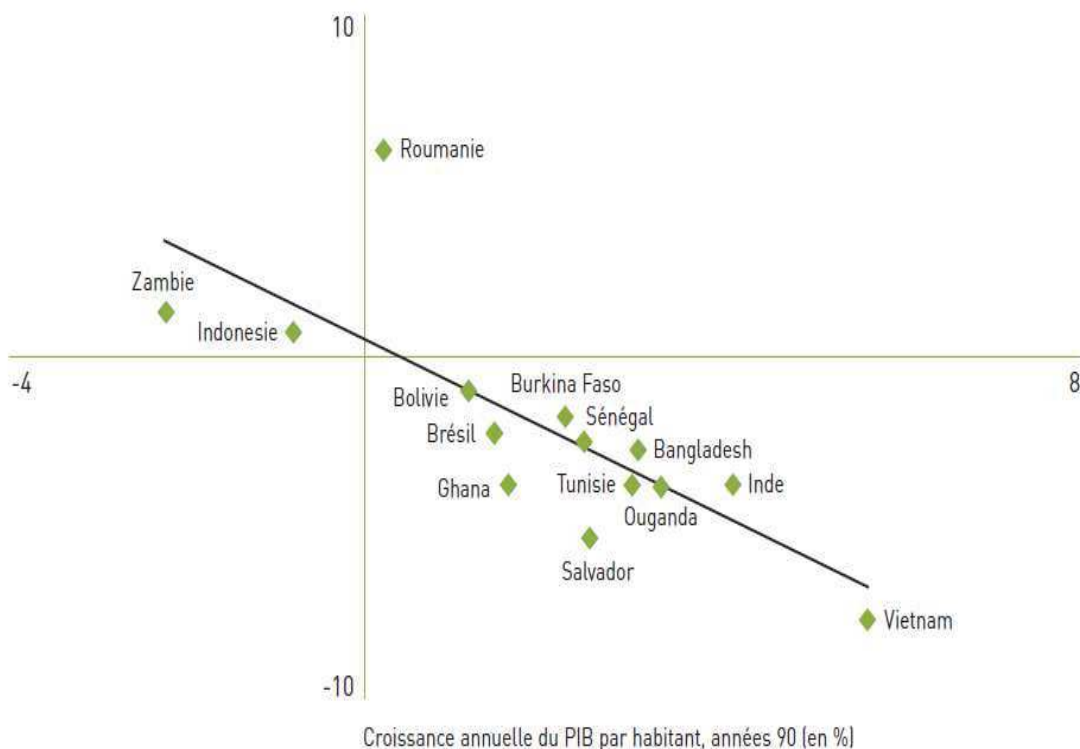
3. Intuitivement, dans le contexte d'une croissance économique, une baisse des inégalités de revenu signifie que les pauvres ont profité plus de la croissance que les autres.

4. Ces indices étant largement connus, la formalisation sera volontairement réduite au minimum nécessaire.

1.1.1 Définitions de la croissance pro-pauvre

Au cours des années 90, la littérature sur la croissance pro-pauvre a fait l'objet d'évolutions profondes. En effet, elle a montré que les stratégies fondées sur la seule promotion de la croissance sont insuffisantes pour engendrer une réduction significative de la pauvreté. L'idée de base est relativement simple : la croissance est un vecteur indéniable de réduction de la pauvreté ; toutefois à croissance égale, certains pays parviennent mieux que d'autres à réduire la pauvreté. Par exemple, comme le montre le graphique ci-dessous, bien qu'en général la croissance s'accompagne d'une réduction de la pauvreté, on constate de fortes divergences entre les pays : le Salvador a vu durant les années 90 une réduction de la pauvreté plus importante que l'Inde et l'Ouganda malgré une croissance plus lente. La croissance n'explique pas à elle seule la totalité de l'évolution de la pauvreté.

FIGURE 1.1 – Croissance et réduction de la pauvreté



Source : Cord (2005)

A l'évidence, l'efficacité de la croissance comme vecteur de réduction de la pauvreté dépend en partie des inégalités de revenu. C'est pour cette raison que

deux grandes approches se sont développées pour définir la croissance pro-pauvre. Certains comme (White and Anderson, 2001; Klasen, 2004) favorisent l'approche relative et d'autres, (Ravallion and Chen, 2003; Kraay, 2006), préfèrent l'approche absolue.

Selon l'approche relative la croissance est dite pro-pauvre lorsque le taux de croissance du revenu des individus pauvres est plus important que celui des individus non pauvres (White and Anderson, 2001; Klasen, 2004). Ainsi, dans le cadre d'une politique économique pro-pauvre, la réduction de la pauvreté sera plus forte comparée à une politique de croissance pour laquelle les inégalités de revenu restent inchangées pour tous, (McCulloch and Baulch, 1999; Kakwani and Son, 2002). Cette définition a trait donc à l'évolution de la distribution des revenus. Autrement dit, elle s'intéresse à la réduction des inégalités de revenu en faveur des pauvres suite à une période de croissance économique. C'est principalement pour cette raison qu'on parle de définition relative de la croissance pro-pauvre. Malheureusement, en se concentrant sur la réduction des inégalités de revenu, cette définition ouvre la voie à plusieurs critiques. Tout d'abord, une politique économique axée sur la réduction des inégalités risque d'être sous-optimale, tant pour les ménages pauvres que pour les autres. Ainsi, selon cette approche un décideur doit préférer une politique de croissance économique qui accroît plus le revenu des pauvres que la croissance moyenne des revenus de l'économie. Ensuite, sous cette définition, une récession économique pourrait être qualifiée de pro-pauvre. En effet, supposons que l'économie soit en récession et que cette dernière affecte plus les non pauvres. Dans ce cadre, nous aurons une réduction des inégalités de revenu. Et par conséquent la récession sera pro-pauvre. En outre, d'après cette définition, la croissance chinoise de ces dernières années n'est pas pro-pauvre, au contraire, elle a été défavorable aux pauvres. Pourtant, il est difficile d'admettre qu'une croissance qui s'est soldée par le recul de la pauvreté qui fut incontestablement la plus importante de ces dernières décennies ne soit pas qualifiée de pro-pauvre Ravallion (2004). Enfin, cette définition pose un problème lié au rôle de l'État quant à sa manière d'intervenir. Elle pourrait légitimer les interventions de l'état visant à réduire les inégalités indépendamment des effets sur la croissance économique Boccanfuso and Ménard (2009).

La seconde définition de la croissance pro-pauvre permet de surmonter ces

limites et se concentre essentiellement sur le lien entre pauvreté et croissance et non sur la distribution du revenu. Sous l'angle absolu, la croissance est dite pro-pauvre lorsqu'elle se traduit par une réduction du taux de pauvreté en terme absolu. Cette seconde approche est beaucoup moins contraignante que la précédente dans la mesure où elle se focalise sur les variations de l'indice de mesure de la pauvreté suite à un épisode de croissance. D'ailleurs, elle est partagée par Ravallion and Chen (2003) et Kraay (2006) qui ont fondé leurs mesures de croissance pro-pauvre sur cette définition. Il serait également intéressant d'insister sur le fait que selon cette approche la croissance sera dite pro-pauvre lorsqu'elle se traduit par une variation positive des revenus des plus pauvres, peu importe la proportion. Cependant, tout comme la première, cette approche n'est pas exempte de critiques. En effet, un problème lié à cette seconde définition est qu'elle peut considérer une croissance économique accompagnée d'une forte augmentation des inégalités de revenu comme étant pro-pauvre.

Osmani et al. (2005) propose une version particulière en agrégeant les deux approches précédentes. Ainsi, selon lui, la croissance sera pro-pauvre lorsqu'elle réduit à la fois la pauvreté et les inégalités. Cette approche a le mérite d'insister sur les interactions possibles entre croissance, inégalité et pauvreté.

Rappelons que, trois grandes étapes sont généralement nécessaires pour effectuer une analyse de la croissance pro-pauvre d'un pays : choisir un indicateur de bien-être, définir un seuil de pauvreté et enfin utiliser une mesure de la croissance pro-pauvre. Ce dernier point fait l'objet de la section suivante.

De façon générale, on utilise les revenus ou les dépenses de consommation comme indicateur de bien-être. Cependant, la littérature privilégie les dépenses pour deux raisons principales. D'abord, elles reflètent assez bien le niveau de vie des ménages puisqu'elles constituent une bonne approximation des revenus. Ensuite, les dépenses sont plus faciles à mesurer que les revenus surtout dans les pays en voie développement où ces derniers proviennent en grande partie du secteur informel. Ces approches permettent de déterminer un seuil de pauvreté (ligne de pauvreté) nécessaire à la mesure agrégée de la pauvreté.⁵

5. On peut se référer aux méthodes basées soit sur l'énergie calorifique (Food Energy Intake (FEI)) soit sur le coût des besoins de base (Cost of Basic Needs, CBN).

1.1.2 Les mesures agrégées de la croissance pro-pauvre

Les mesures traditionnelles de la pauvreté

Les mesures de la croissance pro-pauvre se basent sur des indices de pauvreté agrégés, le plus connu étant l'indice de Foster, Foster et al. (1984) généralisé noté FGT_α :

$$FGT_\alpha = \frac{1}{n} \sum_{i=1}^n \left(\frac{z - y_i}{z} \right)^\alpha, \forall y_i < z, \alpha \in \mathbf{R}_+. \quad (1.1)$$

Avec :

n le nombre d'individus dans la population ;

z le seuil de pauvreté ;

y les dépenses de consommation annuelles par équivalent adulte ;

α un degré d'aversion à l'inégalité. Plus le coefficient α est élevé plus l'indice accorde de l'importance à la situation des pauvres. Un indice moins utilisé, mais qui respecte les mêmes propriétés que l'indice FGT_α , est l'indice de Watts (1968) :

$$W(z) = \int_0^1 (\ln z - \ln y_i) dp. \quad (1.2)$$

Ces indices satisfont notamment les axiomes suivants. Soit P une mesure de pauvreté.

Définition 1.1.1 – Axiome focus – Une mesure de pauvreté doit être invariante aux changements chez les riches. Pour tout vecteur de revenus y obtenu à partir de x avec une augmentation du revenu d'au moins une personne riche alors :

$$P(x, z) = P(y, z). \quad (1.3)$$

Définition 1.1.2 – Axiome de Monotonie – Toute perte (gain) de revenu d'un individu pauvre accroît (réduit) la pauvreté P .

Définition 1.1.3 – Axiome de transfert – La mesure P diminue si un pauvre reçoit un transfert de revenu positif d'un autre individu pauvre plus riche que lui.

Ces indices sont aussi connus pour leurs propriétés liées à l'approche par dominance stochastique.⁶ Cependant, quantifier l'impact de la croissance sur la réduction de la pauvreté nécessite d'aller plus loin dans l'utilisation des indices

6. Il s'agit d'une méthode (comparaisons de courbes) dont l'objectif est d'obtenir des résultats

agrégés. A croissance égale, certains pays parviennent mieux que d'autre à réduire la pauvreté. C'est pour cette raison que ces dernières années plusieurs chercheurs ont tenté de proposer des mesures basés à la fois sur la croissance économique et les indicateurs agrégés de pauvreté. Les premières tentatives d'évaluation de la croissance pro-pauvre se basent sur l'idée selon laquelle la réduction de la pauvreté peut être décomposée en une composante "croissance" et une composante "redistribution" Kakwani (2000) à laquelle pourrait s'ajouter éventuellement un résidu Datt and Ravallion (1992).

La décomposition de Datt and Ravallion (1992)

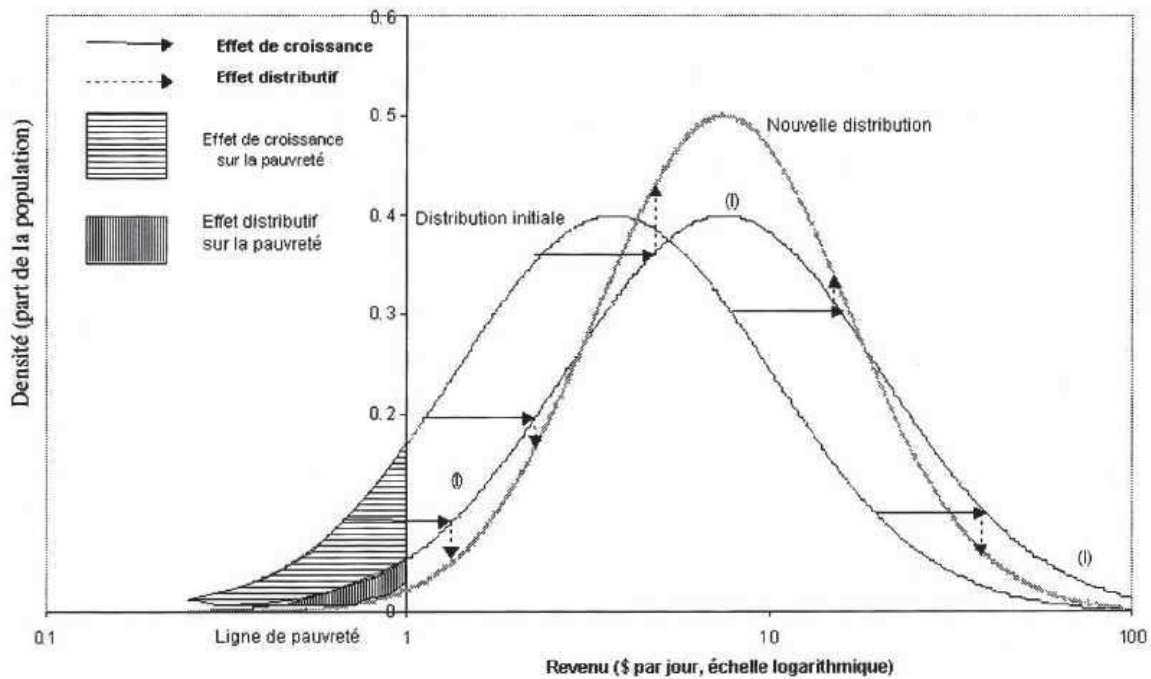
Les institutions internationales comme la Banque mondiale définissent la croissance pro-pauvre comme une croissance qui réduit la pauvreté (même faiblement). Sous cette définition, les pauvres profitent proportionnellement moins que les riches des retombées de la croissance⁷. Cette définition de la banque mondiale est très générale et considère, globalement, la croissance comme étant toujours pro-pauvre. Formellement, la réduction de la pauvreté dépend de deux facteurs : la croissance et comment elle est distribuée entre les pauvres et les riches. Donc toute chose étant égale par ailleurs, la réduction de la pauvreté peut être obtenue par la croissance, la distribution ou par une interaction entre les deux.

Plus précisément, et comme le montre le graphique ci-dessous, la variation du taux de pauvreté peut être décomposée en deux composantes : une composante "croissance", représentée par une translation de la distribution 1 vers la distribution 2, qui correspond au changement du revenu moyen (indépendamment de la distribution du revenu national) et une composante « redistribution », représentée par une translation de la distribution 2 vers la distribution 3, qui résulte d'un changement dans la répartition du revenu national (indépendamment du revenu moyen).

robustes à tout changement d'indice de pauvreté ou de seuil de pauvreté. Un des avantages de cette méthode est de permettre d'ordonner des distributions de revenus sans avoir à fixer un unique seuil de pauvreté *a priori*. Ceci se révèle ainsi particulièrement pertinent lors de l'analyse de la croissance pro-pauvre d'un pays donné puisque définir la ligne de pauvreté est un exercice délicat Duclos et al. (2014). Pour une présentation détaillée le lecteur pourra se référer à l'ouvrage de Duclos and Araar (2006).

7. Cette situation correspond au *trickle down*

FIGURE 1.2 – Décomposition de la pauvreté.



Source : Bourguignon (2003)

Ce type d'approche met en évidence le rôle majeur que jouent la croissance et les inégalités dans toute stratégie de lutte contre la pauvreté et permet de mesurer l'impact de chacune des deux variables sur la variation du taux de pauvreté. Ainsi, cette décomposition a le mérite d'expliquer pourquoi, dans certains cas, une croissance économique forte ne conduit pas forcément à une réduction de la pauvreté, dès lors que la hausse des inégalités vient annihiler le bénéfice d'une croissance positive. Toutefois, l'application empirique de cette méthodologie a mis en évidence des résultats très différents entre les pays en termes d'impact des deux effets. Il semble en effet, que la contribution de la croissance économique à la réduction de la pauvreté soit corrélée positivement au niveau initial de développement et négativement au niveau d'inégalités prévalant dans une économie donnée.

A partir d'une mesure agrégée de pauvreté, les interactions entre croissance, redistribution et réduction de la pauvreté peuvent être captées. Soit P_t une mesure de la pauvreté à la date t :

$$P_t = P(z/\mu_t; CL_t), \quad (1.4)$$

où μ_t est le revenu moyen à la période t et CL_t correspond à la courbe de Lorenz à la période t . Ainsi, une variation de P_t peut provenir d'une variation de μ_t par rapport à z ou d'une variation des inégalités de revenu. Par conséquent, un changement de la pauvreté entre deux date t et $t+n$ peut être décomposé comme suit :

$$\Delta P_{t,t+1} = P_{t+n} - P_t = \int_t^{t+n} \frac{\partial P}{\partial s} ds = \int_t^{t+n} \left(\frac{\partial P}{\partial \mu} \frac{\partial \mu}{\partial s} + \frac{\partial P}{\partial CL} \frac{\partial CL}{\partial s} \right) ds. \quad (1.5)$$

Puisque les valeurs initiales et finales de μ et de CL ne sont pas connues, il est impossible de déterminer ces deux intégrales. Une première option pour surmonter cette difficulté est d'imposer la situation initiale comme référence. Cette possibilité suggérée par Datt and Ravallion (1992) conduit à la décomposition suivante :

$$\Delta P_{t,t+1} = \underbrace{[P(\mu_{t+1}, CL_t) - P(\mu_t, CL_t)]}_{\text{effet croissance}} + \underbrace{[P(\mu_t, CL_{t+1}) - P(\mu_t, CL_t)]}_{\text{effet redistribution}} + \underbrace{R_{t+1, t}}_{\text{résidu}}. \quad (1.6)$$

Les effets obtenus sont :

$$\text{l'effet croissance } G(t; t+n; r) = P(z/\mu_{t+n}; CL_t) - P(z/\mu_t; CL_t);$$

$$\text{l'effet redistributif } D(t; t+n; r) = P(z/\mu_t; CL_{t+n}) - P(z/\mu_t; CL_t);$$

r étant la date de référence à laquelle le changement observé dans la pauvreté est décomposé. Selon Datt and Ravallion (1992), pour $r = t$ le terme résiduel noté $R(t; t+n; r)$ devient :

$$R(t; t+n; r) = G(t; t+n; t+n) - G(t; t+n; t) = D(t; t+n; t+n) - D(t; t+n; t). \quad (1.7)$$

Ainsi, le résidu correspond à la différence entre la composante croissance (respectivement redistribution) évaluée aux niveaux terminal et initial de la courbe de Lorenz.

La principale limite de cette approche concerne le résidu. En effet, ce dernier peut être extrêmement important, et même plus important que l'effet redistribution dans certains cas. C'est pour cette raison que Kakwani (2000) a proposé une va-

riante de cette décomposition de la pauvreté sans le terme résiduel de Datt and Ravallion (1992). Selon Kakwani (2000) la principale explication d'une variation globale de la pauvreté réside dans l'effet croissance et l'effet redistribution.

La courbe d'incidence de la croissance (CIC) – Ravallion and Chen (2003)

D'autres analyses mettent davantage l'accent sur les mesures de la croissance. La courbe d'incidence de la croissance, introduit par Ravallion and Chen (2003), permet d'analyser l'impact de la croissance économique sur les différents centiles de la distribution de revenu (ou des dépenses de consommation) de la population en utilisant le taux de croissance au p ième centile. Globalement, elle représente les centiles de la population ordonnée par les dépenses de consommation (ou du revenu) sur l'axe des abscisses et le taux de croissance annuel des dépenses de consommation du centile correspondant en ordonnée Günther et al. (2006).

Soit $F_t(y)$ la fonction de répartition des revenus, c'est à dire la proportion de la population ayant un revenu inférieur à y à la période t . Ainsi l'inverse de $F_t(y)$ au p ième centile donne le revenu de ce centile et est défini par :

$$y_t(p) = F_t^{-1}(p) = CL'_t(p)\mu_t. \quad (1.8)$$

Avec :

$$y'_t(p) > 0;$$

$CL'_t(p)$ la pente de la courbe de Lorenz ;

μ_t le revenu moyen.

Par conséquent, entre deux périodes $t - 1$ et t , le taux de croissance du p ième centile donne un point de la courbe d'incidence de la croissance (CIC) :

$$G_{t(p)} = \frac{y_{t(p)} - y_{t-1(p)}}{y_{t-1(p)}} = \frac{y_{t(p)}}{y_{t-1(p)}} - 1, \quad (1.9)$$

avec $y_{t(p)}$ représente les revenu ou les dépenses du p ième centile, et $G_{t(p)}$ la taux de croissance de $y_{t(p)}$. De manière équivalente (1.9) la CIC peut être définie par :

$$G_{t(p)} = \frac{CL'_{t(p)}}{CL'_{t-1(p)}}(\gamma + 1) - 1, \quad \forall p \in [0, 1], \quad (1.10)$$

où $\gamma = (\mu_t/\mu_{t-1}) - 1$ représente le taux de croissance du revenu moyen. L'intérêt de cette dernière présentation est qu'elle permet de mieux comprendre la

décomposition de la pauvreté précédente de Datt and Ravallion (1992). En effet, lorsque la courbe de Lorenz CL reste inchangée entre $t - 1$ et t , c'est à dire lorsque l'effet redistributif est nul ($CL'_{t-1(p)} = CL'_{t(p)}$) alors $G_{t(p)} = \gamma$: il s'agit de l'effet croissance Ravallion and Chen (2003).

Selon cette approche la croissance est dite pro-pauvre lorsqu'il y'a dominance stochastique de premier ordre entre les distributions relatives aux deux périodes. Autrement dit, lorsque les taux de croissance $G_{t(p)}$ sont tous positifs, pour tout p , jusqu'à un certain point p^* critique, la croissance est pro-pauvre en terme absolue. Par contre, lorsque la dominance stochastique de premier ordre n'est pas satisfaite, c'est-à-dire lorsque les courbes se croisent (changement de signe pour $G_{t(p)}$), il devient impossible de conclure sur la seule base de cet indicateur.

En conclusion, nous pouvons constater que la CIC présente deux insuffisances majeures. En premier lieu, cette mesure est considérée par Duclos (2009) comme étant trop "contraignante". En effet, elle considère la croissance comme étant pro-pauvre que lorsque le taux de croissance de tous les centiles est supérieur au taux de croissance du revenu moyen. En second lieu, sur le plan pratique, il faut faire attention sur les résultats de cette mesure au niveau de la robustesse de l'évolution des taux de croissance aux queues de la distribution des revenus. Plus précisément, comme le souligne Ravallion and Chen (2003), la variabilité entre les deux périodes due à l'erreur de mesure est très importante dans les extrémités de la distribution des revenus.

1.1.3 Les mesures relatives de la croissance pro-pauvre

Selon la définition relative, la croissance est pro-pauvre lorsqu'elle réduit les inégalités de revenu.

Biais de pauvreté de croissance (BPC)– McCulloch and Baulch (1999)

Le biais de pauvreté de croissance s'intéresse à la variation de la pauvreté résultant du processus de croissance. Il correspond à l'écart entre le processus de croissance et une distribution neutre. Ainsi, il ne prend pas en compte l'effet croissance de Kakwani (2000). Rappelons que ce dernier décompose la variation de la pauvreté en deux composantes : une composantes croissance ΔP_g mesurant la variation de la pauvreté induite par une variation du revenu moyen

en supposant que les inégalités de revenu restent inchangées et une composante redistribution ΔP_i ou inégalité mesurant le changement de pauvreté engendré par une variation de la distribution des revenus pour un revenu moyen fixe :

$$\Delta P_t = \Delta P_g + \Delta P_i. \quad (1.11)$$

A l'opposé le Biais de Pauvreté de Croissance noté (BPC) est défini par :

$$BPC = -\Delta P_i. \quad (1.12)$$

La principale limite de cette approche tient au fait qu'un indice élevé (BPC élevé) ne correspond pas toujours à une importante réduction de la pauvreté. En effet, cette dernière dépend également de l'effet croissance qui n'est pas pris en compte dans cette approche. Ainsi, cet indice ne satisfait pas l'axiome de monotonie.

Indice de la croissance pro-pauvre (ICPP) – Kakwani et al. (2000)

Selon Kakwani et al. (2000) la croissance est pro-pauvre lorsque les pauvres en bénéficient de façon plus proportionnelle que les non pauvres.

Soit η , l'élasticité de la pauvreté à une augmentation de la croissance lorsqu'on suppose les inégalités de revenu inchangées.

Soit δ , l'élasticité globale de la pauvreté à la croissance, c'est-à-dire, la variation de la pauvreté, suite à une période de croissance économique.

La mesure δ peut être aussi définie comme étant la baisse de la pauvreté résultant d'une croissance économique d'un point. Elle peut être décomposée comme la somme de deux composantes : η (composante croissance) et ε (composante redistribution). Cette dernière représente la baisse de la pauvreté résultant d'une variation des inégalités de revenu (évidemment en supposant qu'il n'y ait pas de croissance économique) :

$$\delta = \eta + \varepsilon. \quad (1.13)$$

L'indice de croissance pro-pauvre (ICPP) est défini comme le ratio suivant :

$$\Psi = \frac{\delta}{\eta} = \frac{\eta + \varepsilon}{\eta}. \quad (1.14)$$

L'indice de croissance pro-pauvre est égal au rapport entre la réduction totale de

la pauvreté et la réduction de la pauvreté obtenue en supposant les inégalités de revenu inchangées.

Puisque la mesure η est toujours négative si le taux de croissance est positif (la pauvreté diminue), alors Ψ sera supérieur à un si et seulement si $\varepsilon < 0$, en effet :

$$\Psi = \frac{\delta}{\eta} = \frac{\eta + \varepsilon}{\eta} = \frac{\varepsilon}{\eta} + 1. \quad (1.15)$$

Ainsi, selon ces auteurs le processus de croissance sera qualifié de pro-pauvre lorsque ce ratio est supérieur à 1. Plus précisément, la croissance est dite pro-pauvre lorsque la modification de la distribution des revenus qui accompagne le processus de croissance réduit la pauvreté totale. Par conséquent, la croissance est pro-pauvre, si l'élasticité totale de la pauvreté est supérieure à l'élasticité de la pauvreté à la croissance. Plus généralement, cela signifie que les pauvres profitent plus que proportionnellement de la croissance que les riches. En revanche, lorsque $\Psi < 1$, la croissance est considérée comme non favorable aux pauvres.

Signalons que cet indice, tout comme le biais de pauvreté de croissance, n'est pas monotone. L'indice n'augmente pas de valeur lorsque le taux de réduction de la pauvreté est plus élevé. Il ne s'intéresse qu'à la distribution des bénéfices de la croissance entre les pauvres et les non-pauvres. Il ne tient pas non plus compte du niveau actuel de la croissance, puisque les modifications de la pauvreté dépendent aussi de l'effet croissance. Pour répondre à ces limites Kakwani and Son (2002) ont proposé une mesure alternative appelée le taux de croissance équivalent de la pauvreté.

Taux de croissance équivalent à la pauvreté (TCEP) - Kakwani and Son (2002)

Afin de surmonter les difficultés associées à l'indice de croissance pro-pauvre Kakwani and Son (2002) proposent le taux de croissance équivalent à la pauvreté. Ce taux est défini comme le taux de croissance qui engendre une baisse de la pauvreté dans la même proportion que celle obtenue avec le taux de croissance lorsque l'on suppose les inégalités de revenu inchangées, c'est-à-dire lorsque chaque individu reçoit la même proportion des bénéfices de la croissance. Pour définir le TCEP, les auteurs partent de l'idée que la réduction de la pauvreté est déterminée en grande partie par deux effets. Le premier est le taux de croissance. Ce dernier affecte le revenu moyen de la société. Le second est la manière dont

les bénéfices de la croissance sont distribués entre les membres de la société. Ce dernier correspond à l'indice de croissance pro-pauvre. Formellement :

$$\Delta \log(P_\alpha) = f(g, \Psi), \quad (1.16)$$

avec g le taux de croissance, P_α une mesure de la pauvreté et Ψ l'indice de croissance pro-pauvre. Par conséquent, en définissant le TCEP (noté ici g^*) comme le taux croissance qui engendre une baisse de la pauvreté dans la même proportion que celle obtenue avec le taux de croissance, lorsque l'on suppose les inégalités de revenu inchangées, on a :

$$f(g^*, 1) = f(g, \Psi). \quad (1.17)$$

Notons que lorsque $\Psi = 1$ alors chaque membre de la société reçoit la même proportion des bénéfices de la croissance. A partir de l'équation (1.16), nous avons :

$$\Delta \log(P_\alpha) = f(g, \Psi) = g\eta. \quad (1.18)$$

Selon les auteurs lorsque $\Psi = 1$ et $\eta = \delta$ alors :

$$f(g^*, 1) = \delta g^*. \quad (1.19)$$

Ainsi, en considérant ces deux équations précédentes et la définition de l'indice de croissance pro-pauvre $\Psi = \delta/\eta$ alors on a directement :

$$TCEP = g^* = \Psi g = (\delta/\eta)g. \quad (1.20)$$

De manière tout à fait équivalente on a :

$$TCEP = g^* = g + (\Psi - 1)g. \quad (1.21)$$

Le second terme de cette équation correspond au gain (perte) de croissance lorsque cette dernière est pro-pauvre. Par exemple⁸, considérons que l'élasticité à la pauvreté est de $2/3$, $\Psi = 2/3$. Un taux de croissance $g = 9$ correspond à un $g^* = 6[6 = 9 + (2/3 - 1) * 9]$. Ainsi, le taux de croissance efficace pour obtenir la même réduction de la pauvreté est plus faible que le taux de croissance actuel de trois points, ce qui montre que la croissance n'est pas pro-pauvre dans ce

8. Cet exemple est extrait du papier de Kakwani and Son (2002).

cas. La croissance sera alors considérée comme étant pro-pauvre (respectivement pro-riche) lorsque g^* est supérieur (inférieur) à g . Si g^* prend une valeur comprise entre 0 et g , la croissance conduit à une hausse des inégalités de revenu, mais la pauvreté diminue tout de même.

Le TCEP peut être défini comme "le taux de croissance efficace pour réduire la pauvreté" (Kakwani and Son, 2002). En outre, selon les auteurs la réduction de l'incidence de la pauvreté est une fonction croissante de g^* . Ainsi, maximiser l'ampleur de la réduction de la pauvreté est tout à fait équivalent à maximiser g^* . Par conséquent, pour atteindre une réduction rapide de la proportion des gens vivant dans l'extrême pauvreté, les pouvoirs publics devraient chercher à maximiser le TCEP plutôt que le taux de croissance du PIB. Malheureusement, même s'il respecte l'axiome de monotonie, ce taux reste très sensible au seuil de pauvreté choisi.

La Courbe de croissance de la pauvreté – Son (2004)

La courbe de croissance de la pauvreté a été proposée par Son (2004) en référence au théorème d'Atkinson. Selon ce théorème, un déplacement complet de la courbe de Lorenz généralisée vers le haut (bas) conduit sans risque d'erreur à une diminution (augmentation) de la pauvreté. Soit la courbe de Lorenz notée $CL(p)$:

$$CL(p) = \mu_p p / \mu. \quad (1.22)$$

Elle représente la part du revenu moyen de la population (μ) des $p\%$ les plus bas de la distribution, où μ_p est le revenu moyen des $p\%$ du bas de la distribution. En prenant le logarithme de cette équation, il s'ensuit :

$$\log(\mu_p) = \log(\mu CL(p)) - \log(p). \quad (1.23)$$

En considérant la différence première de cette dernière équation on a :

$$g(p) = \Delta CL(\mu_p) = \Delta \log(\mu CL(p)),$$

avec $\Delta \log(p) = 0$ et $g(p) = \Delta CL(\mu_p)$ le taux de croissance du revenu moyen des $p\%$ personnes en bas de la distribution de la population lorsque les individus sont ordonnés selon le revenu par tête. La courbe $g(p)$ est appelée la courbe de

croissance de pauvreté (Son, 2004) :

$$g(p) = \Delta \log(\mu CL(p)) = \Delta \log(\mu) + \Delta(CL(p)). \quad (1.24)$$

En posant $g = \Delta \log(\mu)$ il s'ensuit :

$$g(p) = g + \Delta(CL(p)); \quad (1.25)$$

où $g = \Delta \log(\mu)$ est le taux de croissance du revenu moyen de l'ensemble de la population. D'après le théorème d'Atkinson si $g(p) > 0$ la pauvreté va diminuer entre les deux dates. En outre, cette équation suggère que lorsque $g(p) > g$, pour l'ensemble des $p < 100\%$, la croissance est pro-pauvre puisque cela implique un déplacement complet de la courbe de Lorenz ($CL(p) > 0$ pour tous les p). Par contre, si $0 < g(p) < g$ pour l'ensemble des $p < 100\%$, la croissance réduit la pauvreté mais elle s'accompagne d'une hausse des inégalités de revenu ($CL(p) < 0$ pour tous les p). Ce qui correspond au phénomène du *trickle-down* : bien que la croissance réduit la pauvreté, les pauvres en bénéficient moins que les autres groupes.

Rappelons enfin que la courbe de croissance de la pauvreté diffère de la courbe d'incidence de croissance de Ravallion and Chen (2003). Tout d'abord, elle est dérivée à partir des conditions de dominance stochastique de deuxième ordre, contrairement à la CIC qui est déterminée à partir de la dominance stochastique de premier ordre.⁹ Ensuite, dans la détermination de la courbe de croissance de pauvreté, on utilise le taux de croissance du revenu au premier quantile – et non jusqu'au p ième pour la CIC Son (2004)

9. Cela signifie que davantage de considérations éthiques sont prises en compte. Précisément, l'ordre 2 de dominance stochastique est associé au respect des transferts de revenus de Pigou-Dalton.

1.1.4 Les mesures absolues de la croissance pro-pauvre

Une croissance est dite pro-pauvre, de manière absolue, lorsqu'elle engendre une diminution de la pauvreté.

Le taux de croissance pro-pauvre (TCPP) – Ravallion and Chen (2003)

En partant de leur courbe d'incidence de la pauvreté, Ravallion and Chen (2003) définissent une nouvelle mesure de la croissance pro-pauvre. Cette dernière appelée le taux de croissance pro-pauvre (que nous nommerons TCPP), correspond à la croissance moyenne de la consommation des pauvres. Le TCPP n'est rien d'autre que le taux de croissance équivalent à la pauvreté, lorsque ce dernier est calculé avec l'indice de Watts (que nous avons défini précédemment) :

$$TCPP = g_t^p = \frac{dW_t}{dW_t^*} \cdot \gamma_t. \quad (1.26)$$

Avec :

g_t^p le taux de croissance pro-pauvre du pays p au temps t ;

dW_t le changement réel de la pauvreté qui s'est produit en utilisant l'indice de Watts ;

dW_t^* le changement de la pauvreté qui se serait produit avec une distribution de la croissance neutre ;

γ_t le taux de croissance global au temps t .

Plus généralement le TCPP est défini par :

$$TCPP_t \equiv \frac{1}{H_{t-1}} \int_0^{H_{t-1}} g_t(p) dp. \quad (1.27)$$

Tout déplacement de la distribution en faveur (respectivement en défaveur) des pauvres engendrera un taux de croissance pro-pauvre supérieur (respectivement inférieur) au taux de croissance ordinaire.

Méthode de décomposition – Kraay (2006)

Selon Kraay (2006), il existe trois sources potentielles de croissance pro-pauvre : une forte hausse des revenus moyens ; une sensibilité accrue de la pauvreté à la croissance des revenus moyens ; enfin une évolution des revenus relatifs (réduction des inégalités de revenu) réduisant la pauvreté.

Soit $y_t(p)$ le revenu de la p ième quantile de la distribution du revenu à la date t . Il peut être réécrit comme une fonction des revenus moyens μ_t et de la courbe de Lorenz $CL_t(p)$:

$$y_t(p) = \mu_t \cdot \frac{dCL_t(p)}{dp}.$$

Pour une mesure de pauvreté P_t :

$$P_t = \int_0^1 f(y_t(p)) dp \quad (1.28)$$

En différenciant P_t , on obtient :

$$\frac{dP_t}{dt} = \int_0^1 \left(\frac{df(y_t(p))}{dy_t(p)} \cdot y_t(p) \right) \cdot \left(\frac{dy_t(p)}{dt} \cdot \frac{1}{y_t(p)} \right) dp. \quad (1.29)$$

Pour simplifier :

$$\eta_t(p) = \frac{df(y_t(p))}{dy_t(p)} \cdot y_t(p), \quad (1.30)$$

et

$$g_t(p) = \frac{dy_t(p)}{dt} \cdot \frac{1}{y_t(p)}, \quad (1.31)$$

où $\eta_t(p)$ représente l'élasticité de la mesure de pauvreté aux variations des revenus moyens et $g_t(p)$ correspond à la croissance des revenus moyens. D'où :

$$\frac{dP_t}{dt} = \int_0^1 \eta_t(p) \cdot g_t(p) dp. \quad (1.32)$$

En ajoutant et en retranchant $\left(\frac{\mu_t}{dt} \cdot \frac{1}{\mu_t} \right)$, on obtient :

$$\frac{dP_t}{dt} = \left(\frac{\mu_t}{dt} \cdot \frac{1}{\mu_t} \right) \int_0^1 \eta_t(p) \cdot dp + \int_0^1 \eta_t(p) \cdot dp \left[g_t(p) - \left(\frac{\mu_t}{dt} \cdot \frac{1}{\mu_t} \right) \right]. \quad (1.33)$$

Cette équation montre trois sources principales de croissance pro-pauvre : une croissance des revenus moyens (A), une forte sensibilité de la pauvreté à la croissance des revenus moyens (B), et enfin une évolution des revenus relatifs (C) :

$$\frac{dP_t}{dt} = \underbrace{\left(\frac{\mu_t}{dt} \cdot \frac{1}{\mu_t} \right) \int_0^1 \eta_t(p) \cdot dp}_A + \underbrace{\int_0^1 \eta_t(p) \cdot dp}_B \underbrace{\left[g_t(p) - \left(\frac{\mu_t}{dt} \cdot \frac{1}{\mu_t} \right) \right]}_C. \quad (1.34)$$

La composante A est communément appelée "la composante croissance" des évolutions de la pauvreté. Cette dernière correspond à la réduction de la pauvreté

qui aurait été observée s'il n'y avait pas eu de modification des revenus relatifs, c'est-à-dire pas de changement dans la distribution des revenus au cours de la période. La composante B capte la part de l'évolution de la pauvreté attribuable au changement des revenus relatifs.

Au niveau empirique, Kraay (2006) montre lorsque l'incidence de la pauvreté diminue, il est de loin très probable que la raison en soit une forte augmentation de la composante croissance. Autrement dit, selon lui cette dernière domine largement.

Avant de conclure cette sous-section, il faudra noter qu'à l'heure actuelle, il n'existe pas de mesure de la croissance pro-pauvre faisant l'unanimité. L'une des principales limites de ces indices est qu'ils se focalisent sur la dimension monétaire de la pauvreté. Ce qui peut paraître très restrictif bien que nous reconnaissons le lien entre la dimension monétaire et non monétaire de la pauvreté. En outre, comme nous allons le voir dans le chapitre 3, ces derniers peuvent aboutir à des conclusions contradictoires. C'est pourquoi, ces années récentes ont vu naître une approche alternative pour quantifier le caractère pro-pauvre de la croissance. Cette seconde approche utilise généralement des modèles économétriques afin d'établir une causalité entre croissance pro-pauvre et indicateurs de bien-être (monétaires et/ou non-monétaires).

1.1.5 L'approche économétrique de la croissance pro-pauvre

Les premiers travaux empiriques sur la croissance pro-pauvre remontent des 90. L'objectif principal était de tenter d'identifier la valeur de l'élasticité du taux de pauvreté du revenu moyen. Par exemple, Ravallion (1995) et Ravallion (1997) trouve une élasticité du taux de pauvreté, qui est définie comme la proportion des individus situés en dessous du seuil de pauvreté, au revenu comprise entre -2,4 et -3,1 selon la taille de l'échantillon retenu. Ainsi, d'après leurs résultats une hausse de 1% du revenu moyen se traduirait par une baisse comprise entre 2,4 % et 3,1% du taux de pauvreté. Toutefois, la principale limite de ces travaux est qu'ils ignorent les inégalités de revenu. En effet, l'élasticité du taux de pauvreté au revenu moyen dépend également de la distribution du revenu et du niveau de développement, qui peut être défini comme le rapport entre le revenu moyen et le seuil de pauvreté retenu.

Les travaux empiriques récents ont tenté de surmonter cet obstacle en intro-

duisant l'effet redistribution sur la valeur estimée de l'élasticité du taux de pauvreté. Ravallion (1997) régresse le taux de réduction de la pauvreté sur un ensemble de variables : l'indice de Gini, le taux croissance du revenu moyen, l'interaction entre ces deux variables et leur carré. Il trouve que c'est la croissance corrigée des inégalités de revenu qui importe le plus. En effet, ces résultats d'estimations indiquent que, plus l'indice de Gini est faible, plus est forte la valeur estimée de l'élasticité du taux de pauvreté : lorsque l'indice de Gini égal à 0,25, l'élasticité au revenu moyen du taux de pauvreté vaut en moyenne -3,3, contre -1,8 si le Gini est égal à 0,59. Toutefois, ces calculs ne prennent pas compte l'effet de la position de la ligne de pauvreté par rapport au revenu moyen. Plus récemment, Bourguignon (2003)¹⁰ trouve que la prise en compte du niveau de développement améliore sensiblement le pouvoir explicatif de la régression du taux de réduction de la pauvreté et que de forte inégalités réduit considérablement le caractère pro-pauvre de la croissance.

Globalement, la littérature empirique sur la croissance pro-pauvre s'attelle à montrer que la croissance joue significativement sur la réduction de la pauvreté. Malheureusement, ces résultats nous disent peu sur le rôle des inégalités sur la réduction de la pauvreté ? Encore moins sur la vitesse de réduction de la pauvreté ? Sur ce point, Kraay (2006) montre que pour lutter contre la pauvreté la priorité devrait être donnée aux politiques favorisant la croissance et il ne faudrait pas attendre beaucoup de résultats des politiques de redistribution des revenus.

En résumé, bien que l'approche économétrique présente l'avantage d'inclure la dimension non monétaire de la pauvreté, elle présente plusieurs faiblesses. En premier lieu, ces premiers travaux sur la croissance pro-pauvre ont été largement contestés par Bhalla (2002). Ce dernier montre que la plupart de ces travaux sous-estimeraient considérablement le rythme de réduction du taux de pauvreté absolue. En effet, il reproche à ces travaux d'utiliser des évaluations sur des estimations directes tirées des enquêtes auprès des ménages et donc sujettes à d'importantes erreurs de mesure. Son argumentaire est que la consommation moyenne estimée à partir des enquêtes (auprès des ménages) est généralement plus faible que celle trouvée dans la comptabilité nationale. Ainsi, en utilisant ces enquêtes comme seule source de données, ces travaux surestimeraient systématiquement le seuil de pauvreté dans les pays étudiés et aurait donc tendance à sous-estimer

10. Sur un échantillon plus large de 113 épisodes de croissance correspondant à 51 pays

le taux de réduction de la pauvreté. En outre, plusieurs travaux montrent que ces résultats sont sensibles aux formes fonctionnelles des modèles. En second lieu, si le concept d'élasticité totale par rapport à la croissance paraît intéressant, il faut néanmoins l'interpréter avec prudence étant donné la multitude de variables dont il dépend et sa sensibilité au ligne de pauvreté choisi Cling et al. (2004). Par ailleurs, estimer le taux de réduction de la pauvreté ne nous dit rien sur le caractère pro-pauvre de la croissance. Encore moins sur la répartition des bénéfices de la croissance sur les différentes couches de la populations. Sur ce point, White and Anderson (2001) ont proposé une approche innovante qui consiste à régresser le taux de croissance des différents quantiles sur la croissance, les inégalités et un ensemble d'autres variables de contrôle. Leurs résultats d'estimation suggèrent que les inégalités constituent bien une dimension importante du problème à ne pas négliger dans la lutte contre la pauvreté. Cette piste de recherche devrait être mieux explorée. En effet, pour déployer une stratégie de croissance pro-pauvre, il est crucial de comprendre les facteurs permettant d'accroître le revenu des quantiles inférieurs. En outre, cette technique présente plusieurs avantages : elle permet d'évaluer plus facilement le caractère pro-pauvre de la croissance et les estimations obtenues par cette approche sont insensibles au seuil de pauvreté choisi. Ce dernier point est particulièrement important. Comme nous l'avons souligné précédemment, estimer une ligne de pauvreté est un exercice risqué : les informations disponibles ne sont pas souvent suffisantes pour définir des seuils de privations rigoureusement fondés sur la méthode des coûts et des besoins de base.

1.1.6 Les politiques économiques pro-pauvres

Depuis les travaux fondateurs de Kuznets, l'approche traditionnelle a pendant longtemps soutenue que seule la croissance économique suffit, à terme, pour réduire la pauvreté et que les inégalités ne sont que passagères. Beaucoup d'études vont dans ce sens. Par exemple, Dollar and Kraay (2002) montrent qu'à long terme le revenu des pauvres progresse de la même manière que le revenu moyen. Toutefois, il faut noter que les années récentes ont vu la croissance évoluer de pair avec une forte augmentation des inégalités. Par ailleurs, Cogneau et al. (2002) montrent qu'il existe de nombreux facteurs qui influent à la fois sur la croissance et sur les inégalités et qu'il serait illusoire de vouloir trouver une

relation simple entre croissance et réduction de la pauvreté et inégalités. Ceci explique en grande partie les faibles performances en matière de réduction de la pauvreté enregistrées dans certains pays malgré une croissance élevée.

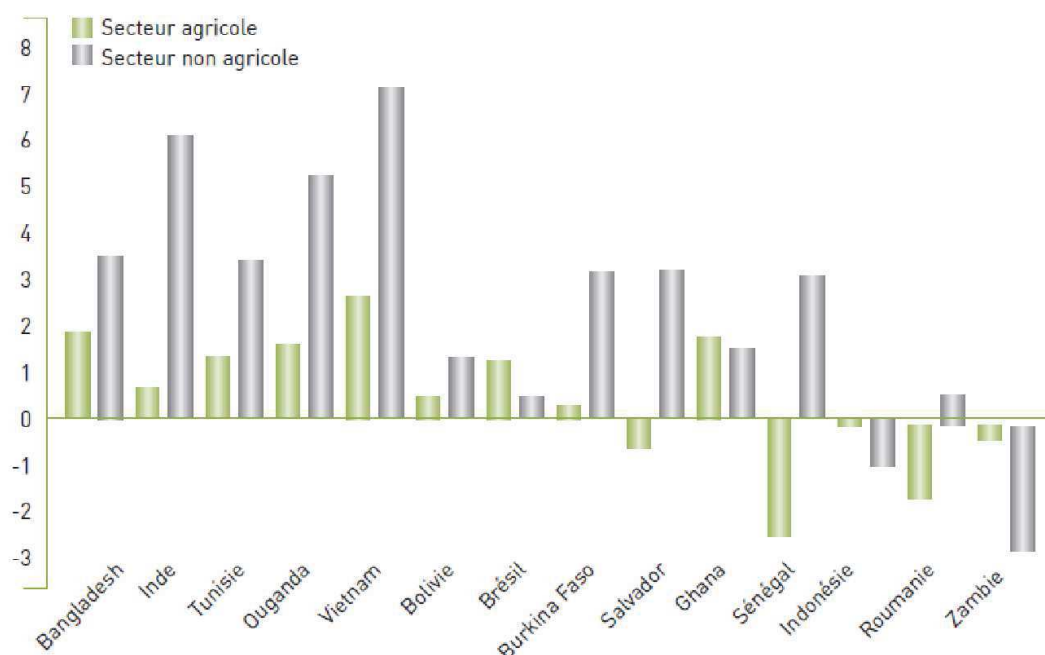
L'approche par la croissance inclusive cherche encore ses repères. Et comme le souligne Klasen (2005), cette approche ne tente nullement à réfuter l'idée selon laquelle la croissance réduit, à terme, la pauvreté. Bien au contraire, elle consiste à accroître l'influence de la croissance sur la réduction de la pauvreté. Promouvoir la croissance pro-pauvre revient donc à donner la priorité aux politiques ayant un impact favorable à la fois sur la croissance et sur la réduction des inégalités. Il n'existe pas de recette toute faite pour promouvoir une croissance inclusive. Néanmoins, certains grands principes se dégagent pour permettre une telle croissance. Généralement, les dépenses publiques qui ont le plus d'influence sur la réduction de la pauvreté sont les dépenses dans le domaine de l'agriculture, de l'éducation de la santé et d'infrastructures (Klasen, 2004; Lopez, 2004; Fan et al., 2004b). Chen and Ravallion (2004) observent qu'en Chine, la croissance du secteur primaire a un impact sur la pauvreté environ quatre fois plus important que celle des secteurs secondaire et tertiaire.

Plus concrètement, c'est par la création d'emploi que la croissance devrait conduire à une réduction soutenue de la pauvreté. Ceci semble plaider en faveur de l'utilisation de techniques intensives en travail comme l'agriculture. Cette dernière joue un rôle essentiel et favorable à la croissance pro-pauvre. Cela s'explique notamment du fait d'une élasticité de la réduction de la pauvreté à la croissance bien plus élevée dans le secteur agricole que dans les autres secteurs. En effet, dans les pays à faible revenu, les activités agricoles représentent la principale ressource des pauvres. Ainsi, au Burkina Faso et au Ghana plus de 70% des pauvres sont regroupés dans le monde rural et vivent essentiellement de l'agriculture (Klasen, 2004). En outre, l'histoire a montré que seuls les pays qui ont réussi à mettre en œuvre des politiques de promotion de l'agriculture ont obtenu une baisse significative de la pauvreté. La hausse de la productivité et des revenus dans le monde rural peut donc jouer un rôle très important à court terme. La croissance agricole augmente la demande rurale de produits non agricoles qui ont tendance à afficher un rapport travail-capital élevé, ce qui crée un cercle vertueux de croissance et de repli de la pauvreté Mellor (2000). D'ailleurs, Fan et al. (2004a) ont montré qu'à dépense égale, c'est la recherche agricole qui a le plus

permis de réduire la pauvreté au Vietnam suivie par l'éducation, l'électricité, et l'irrigation. Évidemment, il est important de promouvoir les facteurs permettant aux ménages de tirer parti des activités non agricoles en zone rurale et dans les villes. Comme nous pouvons le voir dans le graphique ci-dessous, dans les années 90, le PIB non agricole a progressé cinq fois plus vite que le PIB agricole. Généralement, une croissance soutenue de l'agriculture s'accompagne de la progression de l'emploi et des revenus non agricoles des ménages vivant dans les villes et dans les zones rurales les moins isolées. En outre, les ménages exerçant un emploi non agricole ont davantage de chances de sortir de la pauvreté que ceux travaillant dans l'agriculture, car la productivité est souvent plus forte dans les autres secteurs et la croissance généralement plus rapide. Au Vietnam, l'incidence de la pauvreté a reflué de 31% pour la main-d'œuvre agricole, de 41% pour les travailleurs du secteur informel et de 62% pour le secteur formel Lopez (2004).

FIGURE 1.3 – Évolution du PIB non agricole

Taux de croissance médian, par habitant et par an, année 90



Source : Lopez (2004)

Parallèlement, dans une optique de croissance pro-pauvre, il faut envisager de développer les infrastructures, l'éducation de base et la santé. Ces dernières,

en accroissant le capital humain des pauvres, devraient leur ouvrir de nouvelles opportunités. L'alphabétisation des femmes joue un rôle majeur dans la capacité de la croissance à faire reculer la pauvreté. En Inde, elle a été le principal responsable des différences entre les États en termes d'incidence de la croissance non agricole sur la pauvreté Ravallion and Datt (1996). Il faut toutefois souligné que les infrastructures directement utiles aux pauvres sont plus efficaces lorsqu'elles sont conçues dans une logique de réseaux au niveau national et régional. Évidemment, il est impératif de réduire les dysfonctionnements des infrastructures existants avant d'envisager de nouvelles réalisations Chataigner and Rafinot (2005). A mesure que le pays se développe, l'État peut également compléter le revenu perçu par les pauvres sur le marché par des transferts publics, ce qui permet aux ménages déshérités de profiter indirectement de la croissance grâce à une politique redistributive. Cependant, dans les pays pauvres avec des capacités financières et administratives limitées il serait difficile d'imaginer que les transferts publics puissent avoir une influence significative sur la réduction de la pauvreté. En outre, il est important de noter qu'effectuer des transferts de revenus à grande échelle n'est pas une réponse adaptée pour réduire la pauvreté. Tout d'abord, dans la réalité, il existe très peu d'exemple de cas où pareilles politiques ont entraîné une réduction significative de la pauvreté. D'ailleurs, plusieurs travaux empiriques (Wiggins and Higgins, 2008, par exemple) montrent que la réduction substantielle de la pauvreté dans les pays de l'OCDE au court du XXI ème siècle est attribuée à une croissance soutenue avec de modestes transferts de revenus. En outre, une telle politique redistributive pourrait ralentir la croissance en freinant l'incitation à investir, à innover et même à travailler. Par exemple Jha and Srinivasan (2001) ont montré que la plupart des subventions destinées à l'agriculture (notamment au riz et au blé) n'ont pas permis de faire reculer la malnutrition des enfants. En outre, une réduction des subventions en Andhra Pradesh (Inde) n'a eu qu'un effet mineur sur la situation nutritionnelle des enfants Tarozzi (2005).

TABLE 1.1 – Impact des transferts publics sur la répartition des revenus

| | Burkina | Salvador | Ghana | Ouganda | Vietnam | Zambie |
|------------------------------|------------------------|----------|---------|-----------|---------|--------|
| Année | 2003 | 2000 | 1999/99 | 1999/2000 | 1998 | 1998 |
| Après les transferts publics | 0.45 | 0.53 | 0.40 | 0.402 | 0.35 | 0.54 |
| Avant les transferts publics | 0.46 | 0.53 | 0.40 | 0.402 | 0.36 | 0.54 |
| Source | Source : (Lopez, 2004) | | | | | |

Avant de conclure cette section, il faut noter qu'aujourd'hui, il n'existe pas de méthode faisant l'unanimité pour évaluer le caractère pro-pauvre de la croissance. La recherche de nouveaux indicateurs doit donc être considérée comme une priorité. A notre avis, ces nouveaux indices devraient prendre en compte à la fois la dimension non-monnaire de la pauvreté et les interactions entre distribution du revenu et la croissance. Ce dernier point est un élément essentiel. En effet, lorsqu'on conçoit des stratégies de lutte contre la pauvreté qui associent des mesures favorisant la croissance et un changement dans la distribution du revenu, il importe de déterminer s'il existe une relation entre ces deux variables. Malheureusement, la littérature théorique est très partagée quant à la nature de cette relation et la littérature empirique n'a pas pu déceler de relation permanente entre ces deux variables. Ceci peut s'expliquer par les nombreuses difficultés méthodologiques rencontrées par les études empiriques. Ces dernières sont abordées dans la section suivante.

1.2 Les inégalités de revenu nuisibles ou favorables à la croissance

Les inégalités de revenu sont un nouveau problème que les économistes cherchent à comprendre. Depuis les travaux fondateurs de Kuznets dans les années 1950, le débat sur les interactions entre distribution du revenu et croissance ne cesse d'être alimenté par de nouveaux arguments.

Cette hypothèse de Kuznets, selon laquelle les inégalités de revenu aug-

mentent dans les premières phases du processus de croissance puis retombent avec le développement économique, largement corroborée par les travaux empiriques antérieurs (Adelman and Morris, 1973; Ahluwalia, 1976; Papanek and Kyn, 1986) a très vite laissé place à la controverse. Tout d'abord, Anand and Kanbur (1993b) montrent que sur le plan empirique l'apparition de la courbe de Kuznets dépendrait en grande partie de la forme fonctionnelle retenue. Ensuite et surtout peut être, dans les années 1990, un nouveau fait stylisé issu de l'économétrie révèle une relation de causalité inverse entre les inégalités de revenu et la croissance. Ces travaux sont l'œuvre de Persson and Tabellini (1994). De nombreux modèles théoriques ont été élaborés pour rendre compte de cette régularité. Nous pouvons citer à titre d'exemple les modèles d'économie politique Persson and Tabellini (1994); Alesina and Rodrik (1994), les modèles d'imperfection du marché du crédit (Banerjee and Newman, 1993; Galor and Zeira, 1993), les modèles portant sur l'instabilité sociopolitique que pourraient engendrer de fortes inégalités de revenu Alesina and Perotti (1996), etc. Malheureusement, cette nouvelle relation, tout comme la courbe de Kuznets d'ailleurs, a très vite montré sa fragilité. En effet, ces premiers travaux ont été largement remis en cause par les analyses sur données de panel. Ces dernières ont trouvé une relation positive, non linéaire, voire nulle, entre la distribution des revenus et la croissance (Deininger and Squire, 1998; Barro, 2000; Banerjee and Duflo, 2003; Forbes, 2000).

Comprendre la nature générale de cette relation requiert tout d'abord un réexamen attentif de cette littérature théorique (première sous-section) mais également d'analyser les résultats empiriques disponibles, obtenus à l'aide de modèles économétriques (deuxième sous-section).

1.2.1 Les approches théoriques

Aujourd'hui, il existe une multitude de modèles théoriques analysant les canaux de transmissions par lesquels une forte concentration de la richesse est susceptible d'influencer le processus de croissance dans le long terme.

Selon Galor and Tsiddon (1997) de fortes inégalités de revenu seraient favorables à la croissance. Tout d'abord, puisqu'il existe une relation de causalité positive entre le taux de croissance et l'épargne alors la croissance des économies inégalitaires serait plus importante car la propension marginale à épargner des riches est plus forte que celle des pauvres (Kaldor). Ensuite, comme la création

de nouvelles entreprises exige souvent des coûts fixes particulièrement importants, dans un contexte de marché financier imparfait, une forte concentration de la richesse serait indispensable pour l'innovation et la croissance Galor and Tsiddon (1997). Cependant, ces modèles ont été largement contestés par les travaux contemporains. Ces nouveaux travaux concluent à une relation négative entre les inégalités de revenu et la croissance de long terme (Galor and Zeira, 1993; Persson and Tabellini, 1994; Durlauf, 1994; Alesina and Perotti, 1996).

Les inégalités : un catalyseur de la croissance économique

La sagesse conventionnelle a pendant longtemps considéré que les inégalités de revenu sont très bonnes pour les incitations et donc pour l'investissement et la croissance. Elles ont été souvent comprises comme indispensable à toute performance économique. Car "l'inégalité est nécessaire à l'accumulation et qu'elle contient les germes de l'accroissement éventuel du revenu de chacun", Adelman et al. (1989). Cette idée selon laquelle les inégalités renforcent la croissance s'appuie sur deux arguments : le taux d'épargne et l'imperfection du marché du crédit.

L'argument du taux d'épargne.

Selon le modèle keynésien, l'épargne S est une fonction croissante du revenu R . Ce qui implique selon Kaldor que la propension marginale à épargner des riches (S_r) est plus forte que celle des pauvres (S_p). En supposant une causalité positive entre l'épargne et l'investissement, une forte concentration de la richesse serait favorable à la croissance économique :

$$S = f(R), \quad (1.35)$$

avec $f' > 0$ donc $S_r > S_p$. Plusieurs critiques peuvent être adressées à ce modèle. Tout d'abord, elle suppose une relation linéaire entre la propension marginale à épargner et le revenu. Ce qui n'est pas forcément le cas. D'ailleurs, cette relation aurait la forme d'une fonction en "U" inversée Ray (1998). En effet, la hausse de la propension marginale à épargner serait plus forte chez les revenus inférieurs et moyens. Elle diminuerait même chez les riches. Par conséquent, cette relation positive entre inégalité de revenu et croissance serait observée qu'à des niveaux de revenus faibles où une concentration de la richesse favorise l'accumulation.

Ensuite, lorsqu'on s'intéresse de près à ce modèle on se rend compte qu'il est plus adapté aux pays développés où la répartition des revenus est en grande partie constituée de salaires et de bénéfices Aghion et al. (1999). Enfin, ce modèle considère qu'une redistribution des ressources réduit l'épargne et donc est défavorable à la croissance. Ceci n'est vrai que dans le cas où l'investissement n'est financé que par l'épargne nationale.

L'argument de l'imperfection des marchés des capitaux.

Un second argument utilisé pour expliquer l'influence positive des inégalités de revenu sur la croissance s'appuie sur le caractère indivisible de l'investissement. Selon Galor and Tsiddon (1997), dans un contexte de marché financier imparfait, une très forte concentration de la richesse serait indispensable pour financer la création de nouvelles entreprises qui exigent souvent des coûts fixes irrécupérables particulièrement importants. En effet, les projets d'investissement et la mise en œuvre d'innovations technologiques impliquent souvent d'importants coûts fixes.

Ces théories économiques d'hier largement consensuelles sont remises en cause par les travaux contemporains. Aujourd'hui, la vision dominante serait celle d'un impact négatif des inégalités de revenu sur les performances économiques.

Les inégalités défavorables à la croissance

Aujourd'hui, de nombreux papiers traitent des canaux de transmissions par lesquels une forte concentration de la richesse défavorise la croissance économique. Parmi les explications développées dans la littérature figurent celle de l'instabilité socio-politique, de l'approche par les modèles d'économie politique, de l'arbitrage éducation/fécondité, etc.

Les modèles d'économie politique.

Dans une société démocratique lorsque le revenu moyen dépasse de loin le revenu médian un système de vote à la majorité simple serait favorable à une redistribution des ressources des riches vers les pauvres (transferts, dépenses publiques, etc.) ce qui pourrait conduire à des distorsions fiscales et détournerait ainsi la classe riche de l'investissement : Persson and Tabellini (1994), Alesina and Rodrik (1994). Ces modèles connus sous le nom de modèles d'économie politique

ont la particularité d'endogénéiser les politiques de redistribution. Cette redistribution peut prendre la forme de transferts, de dépenses publiques en faveur des pauvres (en éducation et santé), de réforme fiscale (faire varier le prix des biens) favorable aux plus pauvres. Ces dernières ont été mises en œuvre dans certains pays comme le Mexique et l'Égypte.

Selon Alesina and Rodrik (1994) cette relation négative entre inégalités de revenu et croissance se vérifie également dans les dictatures. En effet, le dictateur tient à la stabilité de son régime donc à la demande sociale et aux situations conflictuelles. D'ailleurs, Benabou (1996) a généralisé cette conclusion dans les sociétés non démocratiques. En outre, (Perotti, 1993; Saint Paul and Verdier, 1996) ont aussi développé leurs modèles d'économie politique où la distribution du revenu affecte l'investissement et la croissance.

En réalité, l'influence négative des inégalités de revenu sur le processus de croissance peut apparaître dans les modèles d'économie politique même en dehors de la redistribution. Les riches peuvent empêcher l'application des politiques de redistribution à travers des lobbies et l'achat de vote aux législateurs.¹¹ Malheureusement, ce type d'action peut conduire à un cercle vicieux. En effet, plus la concentration de la richesse est forte, plus la demande de redistribution sera importante. Ce qui conduit les lobbies à accroître leurs efforts pour empêcher la redistribution. Ainsi, l'intervention de ces groupes d'intérêt en favorisant la corruption affecte les performances économiques. Par conséquent, l'inégalité des revenus peut avoir un effet négatif sur la croissance d'une économie, même si aucune redistribution de revenu n'est observée.

Globalement deux mécanismes sont à l'œuvre dans ces modèles. Le premier conclut à une relation positive entre inégalités de revenu et taxation. Le second s'intéresse à l'effet désincitatif de la taxation sur l'investissement et par conséquent sur la croissance.

En même temps que cette vaste littérature théorique, l'intérêt pour la dynamique du processus de croissance et l'évolution de la distribution des revenus a suscité plusieurs analyses empiriques. Ces dernières sont abordées dans la sous-section suivante.

11. Dans certains pays ces groupes d'intérêt se permettent parfois de financer les campagnes électorales dans le seul but de bénéficier d'avantages fiscaux après les élections

1.2.2 Les approches empiriques

Les modèles d'économie politique

Tout d'abord, selon les modèles d'économie politique une distribution inégalitaire des ressources génère une forte demande de redistribution. En s'inspirant de Benabou (1996), nous avons répertorié les principaux résultats des travaux empiriques consacrés à l'examen de la relation entre inégalités de revenu et redistribution (la Table 1.2). Ces résultats sont particulièrement décevants. En effet, l'influence des inégalités de revenu sur la redistribution est rarement significative. En outre, le signe de cette relation varie d'une étude à l'autre en raison des nombreuses spécifications retenues.

Persson and Tabellini (1994) expliquent ces résultats par les difficultés inhérentes à la mesure de la redistribution. Ils font remarquer que la mesure de la redistribution est un exercice extrêmement complexe. La redistribution peut prendre plusieurs formes autres que les transferts comme par exemple : les impôts progressifs, les lois sur le salaire minimum, les restrictions commerciales, certaines dépenses publiques, la législation sur les brevets, la réglementation et la protection des droits de propriété, etc. Cela est certainement vrai. Cet argument est d'autant plus justifié que la plupart (mais pas tous) des travaux empiriques antérieurs répertoriés dans la Table 1.2 (ci-dessous) utilisent comme mesure de redistribution les taux de prélèvement moyen et marginaux. Cependant, cet argument n'est pas tout à fait convaincant. Premièrement, ces études incluent déjà certaines de ces variables comme les taux d'imposition ou les dépenses d'éducation. Deuxièmement, selon Benabou (1996), il est difficile de comprendre pourquoi l'augmentation de la pression pour la redistribution se manifesterait que par des voies indirectes et non par les effets directs.

Ensuite, le second mécanisme conclut à une relation négative entre la redistribution et la croissance. Ce lien traduit l'effet désincitatif de la taxation sur la croissance. A nouveau, les travaux empiriques rejettent largement cette hypothèse. Selon Lindert (1996), la redistribution serait favorable à la croissance, de plus la relation est statistiquement significative. L'analyse de la Table 1.2 révèle que ce résultat est loin d'être une exception. En effet, ces conclusions sont confirmées par plusieurs travaux comme ceux de Perotti (1996).

TABLE 1.2 – les résultats empiriques des modèles d'économie politique

| Articles | inégalité ↔ redistribution | | redistribution ↔ croissance | |
|--------------------------------|----------------------------|-----|-----------------------------|-----|
| | Trans, Tax | Exp | Trans, Tax | Exp |
| Devarajan <i>et al.</i> (1993) | | | + | (-) |
| Easterly and Rebelo (1993) | | + | (-) | |
| Lindert (1996) | ⊖ | (-) | (±) | ⊕ |
| Perotti (1992) | (-) | (±) | + | |
| Perotti (1994) | (-) | | (+) | |
| Perotti (1996) | (+) | (+) | ⊕ | ⊕ |
| Persson and Tabellini (1994) | (+) | | (-) | |

Exp : la part des dépenses publiques d'éducation dans le PIB.

Tax et Trans : un ensemble de transfert et taxe comme par exemple la part des dépenses publiques d'éducation et de santé dans le PIB (Devarajan *et al.*), les taux marginaux et moyennes d'imposition Easterly and Rebelo (1993); l'impôt sur le revenu, les dépenses publiques, la part des transferts dans le PIB etc.

⊖ ⊕ : signe consistant et généralement significatif.

(-) (+) : signe consistant mais non significatif.

- + : signe non consistant et généralement non significatif.

(±) : signe non consistant et généralement significatif.

Ces résultats surprenants traduisent-ils une sorte de causalité inverse entre redistribution et croissance? Autrement dit, la redistribution serait-elle un bien de luxe qui traduirait une loi de Wagner? ¹²

Perotti (1996) trouve que les prélèvements marginaux et moyens ont une influence positive sur la croissance. Encore mieux, ses travaux excluent l'idée d'une causalité inverse entre redistribution et croissance. En effet, dans son modèle à équations simultanées, le taux de croissance et le PIB par tête apparaissent comme des variables explicatives de la taxation. Ce qui exclut ainsi l'hypothèse que la redistribution serait un bien de luxe.

Si les travaux empiriques antérieurs remettent en cause l'ensemble des conclusions des modèles d'économie politique, ces modèles ont aussi été critiqués sur

12. Adolph Wagner (1835-1917) était un grand économiste Allemand connu pour ses travaux relatifs à l'étude des dépenses publiques dans les pays développés. Il avait observé une évolution croissante de la part des dépenses publiques dans le PIB. La loi de Wagner met en évidence une explication de ce phénomène et prédit qu'il devrait perdurer dans le temps. Autrement dit, cette loi considère que les pays développés pratiquent des politiques de redistribution plus généreuses.

le plan théorique. Tout d'abord, selon Saint Paul and Verdier (1996), les inégalités de revenu peuvent évoluer sans que la position médiane soit affectée. Cette évolution peut très bien être limitée à la population la plus pauvre. Ensuite Edsall (1984) et Conway (1991) montrent que les individus démunis (et les moins instruits généralement) ont un faible taux de participation dans les activités politiques (élection entre autre). Si un tel résultat est avéré, l'électeur décisif serait plus riche que l'électeur médian. Enfin, il semble que les dépenses sociales bénéficient plus à la classe moyenne et aux plus riches.

A côté des modèles d'économie politique, il existe d'autres arguments non moins pertinents qui considèrent que les inégalités de revenu ralentissent les performances économiques, à travers l'instabilité socio-politique qu'elles engendrent.

L'instabilité socio-politique

De fortes inégalités de revenu alimentent la violence et le mécontentement social. Certains individus se sentent marginalisés et exclus du "système" ; ce qui favorise l'apparition d'activités illicites et des manifestations risquant de provoquer émeutes et coups d'État Hibbs (1973). Les inégalités de revenu incitent les pauvres à s'engager dans des activités criminelles et d'autres actions subversives. La stabilité des institutions politiques peut même être menacée par la révolution. La participation des pauvres dans le crime et d'autres actions antisociales représente un véritable gaspillage des ressources. Le temps et l'énergie des criminels ne sont pas consacrés à la création de richesses, mais aussi les ressources utilisées pour lutter contre la criminalité pourraient servir plus efficacement dans la production. En outre, le non respect des droits de propriétés résultant de ce chaos entrave les décisions d'investissement. Ainsi, grâce à ces différentes dimensions de troubles socio-politiques, l'inégalité tend à réduire la productivité d'une économie.

Dans une approche semblable, Alesina and Perotti (1996) avancent qu'une forte concentration de la richesse favorise la criminalité, le vol et une instabilité sociale. Ces derniers perturbent les activités de marché et augmentent ainsi l'incertitude économique. Mais aussi, les manifestations répétitives, les guerres civiles, les troubles politiques et les menaces physiques sur les travailleurs et les entrepreneurs peuvent avoir des effets directs sur la production et globale-

TABLE 1.3 – Analyses empiriques des relations inégalité / instabilité et instabilité / croissance

| Articles | inég ↔ inst | inst ↔ croiss |
|----------------------------|-------------|---------------|
| Alesina and Perotti (1996) | ⊕ | - |
| Alesina et al (1996) | ⊕ | ⊖ |
| Benhabib-Spiegel (1996) | | (-) |
| Easterly-Rabello (1993) | | (±) |
| Levine and Renelt (1992) | | - |
| Perotti (1992) | ⊕ | ⊖ |
| Perotti (1996) | ⊕ | ⊖ |
| Svensson (1993) | + | ⊖ |
| Venieris and Gupta (1986) | | ⊖ |

inég : inégalités de revenu
 inst : instabilité socio-politique
 croiss : Croissance économique

ment sur l'activité économique en freinant sérieusement l'incitation à investir. Cette agitation sociale réduit l'horizon des institutions politiques, ce qui nourrit un sentiment d'incertitude. En effet, les investisseurs considèrent que des années d'efforts peuvent à tout moment être réduites à néant, et la propriété privée peut toujours être nationalisée par les révolutionnaires.

Pour étudier la pertinence de ces modèles, deux mécanismes ont été testés dans la littérature empirique. Le premier relie les inégalités de revenu et l'instabilité socio-politique. Le second examine l'effet néfaste de cette dernière sur la croissance. Toutefois, une telle étude n'est envisageable que lorsque l'instabilité est mesurable. Benabou (1996) distingue deux approches pour mesurer l'instabilité. La première s'intéresse aux facteurs socio-politiques comme les manifestations, les grèves, la violence de masse, les révolutions (Venieris and Gupta, 1986; Alesina and Perotti, 1996; Perotti, 1996). La seconde approche est liée aux facteurs institutionnels comme la qualité de la bureaucratie, la corruption, le respect des droits de propriété, etc.

Plusieurs études confirment la pertinence de ces modèles (Table 1.3).

A partir d'une analyse en composante principale, Alesina and Perotti (1996) construisent un nouvel indice d'instabilité sociopolitique appelé *SPI* (index of

Socio-Political Instability) :¹³

$$SPI = 1,86ASSASS + 1,28DEATH + 7,5COUP + 7,09UCOUP + 5,04DICT. \quad (1.36)$$

Les variables sont les suivantes :

ASSASS : le nombre d'assassinats politique ;

DEATH : le nombre de personne tué dans des violences de masse ;

COUP : les coups d'état réussis ;

UCOUP : le nombre de tentative de coup d'Etat mais sans succès ;

DICT : une variable indicatrice (avec 0 démocratie ; 0,5 semi démocratie et 1 les dictatures).

Les auteurs montrent à partir d'un modèle à équations simultanées, où les variables dépendantes sont l'investissement total et *SPI*, que l'instabilité socio-politique est défavorable à l'investissement.

L'éducation et la fécondité

L'arbitrage éducation-fécondité est aussi une autre explication avancée dans la littérature théorique et empirique pour rendre compte du lien négatif entre inégalités de revenu et croissance Becker et al. (1990); Perotti (1996). Le mécanisme est simple. Lorsqu'un individu s'engage à investir dans le capital humain (éducation), il accroît en même temps le coût d'opportunité d'élever un (des) enfant(s). Plus généralement, l'investissement en capital humain (éducation) évolue de manière inverse avec la fécondité. L'impact d'une baisse de la fécondité sur la croissance s'appréhende de manière intuitive. En effet, puisque le capital humain est un déterminant essentiel de la croissance, l'investissement accru dans l'éducation résultant d'une baisse de la fécondité se répercute positivement sur la croissance.

Les résultats des travaux empiriques effectués par Perotti (1996) confirment largement cette thèse. En effet, à partir d'un modèle à équations simultanées, il arrive aux conclusions suivantes. Tout d'abord, il existerait un lien positif entre inégalités de revenu-fécondité et un lien négatif entre fécondité-croissance. Notons que dans la première spécification de Perotti (1996), les variables dépendantes sont le taux de croissance et le taux de fécondité. Ensuite, dans la seconde spé-

13. Il est important de noter que cet indice n'est pas fondamentalement différent de celui de Hibbs (1973).

cification où les variables exogènes sont le taux de croissance et l'investissement en capital humain, il conclut à une relation négative entre inégalités de revenu et investissement en capital humain et positive entre investissement-croissance.

La courbe de Kuznets

De nombreuses études empiriques ont été développées pour analyser les interactions entre processus de croissance et la distribution du revenu. Globalement, les premiers modèles analysant l'impact des inégalités de revenu sur la croissance remontent aux années 1970. En effet, dans les années 1970, les données transversales réunies sur les PED par Ahluwalia (1976) semblaient corroborer l'hypothèse de Kuznets Ahluwalia (1976). Cependant, la robustesse de ces résultats fut de courte durée. Tout d'abord, Anand and Kanbur (1993b) montrent que l'apparition de la courbe de Kuznets dépendrait en grande partie de la forme fonctionnelle retenue. Ensuite, dans les années plus récentes, un nouveau fait stylisé issu de l'économétrie des inégalités et de la croissance révèle une relation de causalité inverse. Ces travaux trouvent que les inégalités sont défavorables à la croissance (Persson and Tabellini, 1994; Alesina and Rodrik, 1994). Malheureusement, la relation statistique exposée dans ces dernières études s'est révélée bien plus faible qu'elle ne la paraissait au premier abord et plusieurs études ont ensuite remis en question les résultats obtenus. En effet, ces premiers travaux ont été largement contestés par les analyses sur données de panel. Ces dernières ont trouvé une relation nulle, non linéaire, voire positive, entre les inégalités de revenu et la croissance (Deininger and Squire, 1998; Forbes, 2000; Barro, 2000; Banerjee and Duflo, 2003).

A partir d'un échantillon composé de 60 pays, Ahluwalia (1976) analyse la relation suivante :

$$Y_i = c + \alpha \log(\text{PIB}/\text{tête}) + \beta \log(\text{PIB}/\text{tête})^2 + \delta D + \varepsilon_i, \quad (1.37)$$

où,

Y_i est la part du revenu du groupe i ;

D une variable indicatrice (1 = pays socialiste et 0 = Non).

En estimant l'équation précédente, Ahluwalia arrive aux résultats suivants :

$$\hat{Y}_i(20\% \text{ les plus riches}) = -57,58 + 89,95 \log(\text{PIB}/\text{tête}) - 17,56 \log(\text{PIB}/\text{tête})^2 - 20,15D$$

$$\hat{Y}_i(40\% \text{ médians}) = 87,03 - 45,59 \log(\text{PIB}/\text{tête}) + 9,25 \log(\text{PIB}/\text{tête})^2 + 8,21D$$

$$\hat{Y}_i(20\% \text{ les plus pauvres}) = 27,31 - 16,97 \log(\text{PIB}/\text{tête}) + 3,06 \log(\text{PIB}/\text{tête})^2 + 5,54D.$$

Tous les coefficients estimés sont statistiquement significatifs, en plus il fournit les points de retournement pour chaque groupe. Rappelons que la possibilité d'une baisse de la part de revenu des groupes les plus pauvres soulève deux questions fondamentales quant à l'influence du processus de développement sur le bien-être. Plus précisément, la littérature théorique distingue deux façons d'envisager le processus sous-jacent à l'hypothèse de Kuznets. Selon Adelman and Morris (1973), ce schéma (la courbe de Kuznets) peut très bien refléter un processus d'appauvrissement absolu des groupes à faible revenu. Malheureusement, nous ne pouvons pas écarter cette possibilité a priori. En effet, l'apparition d'une courbe de Kuznets peut résulter d'une érosion des structures économiques traditionnelles sous l'impact d'un secteur moderne en pleine expansion qui fait concurrence aux secteurs traditionnels à la fois pour les marchés et les ressources. Il en résulterait une rupture du système économique existant qui conduirait à un appauvrissement absolu de certains groupes. Cependant, il y a une autre façon d'envisager l'apparition de la courbe de Kuznets. Dans cette seconde approche, la concentration de la richesse se renforce non pas parce que des groupes plus pauvres baissent en valeur absolue mais parce qu'ils évoluent (progressent) à un rythme moins rapide que ceux des tranches supérieures. Ainsi, si le développement se traduit par une croissance économique favorable (au départ) aux riches et s'il n'existe pas de lien très puissant entre le revenu de ces groupes et celui des tranches les plus pauvres, il doit en résulter un accroissement de l'inégalité relative qui ne va pas nécessairement de pair avec un appauvrissement absolu. La distinction essentielle entre ces deux approches réside dans la question de savoir si le bouleversement des activités économiques traditionnelles à faible revenu est une condition essentielle à la croissance du secteur moderne ou si le problème est simplement imputable à un manque d'intégration de deux secteurs. Dans ce dernier cas, le problème n'apparaît plus aussi grave car nous n'avons plus à faire face à une contradiction inhérente au processus de développement. Les résultats

de Ahluwalia (1976) ne nous permettent pas de choisir clairement l'une ou l'autre de ces possibilités.

En résumé les résultats de Ahluwalia (1976) confirment largement l'hypothèse de Kuznets. Il existe bien un processus global de croissance des inégalités initiales suivie (accompagner) d'un resserrement au-delà d'un certain revenu par tête. Cependant, ces résultats ne sont pas exempts de critiques. Tout d'abord, ils sont extrêmement sensibles à la forme fonctionnelle retenue mais aussi à la composition de l'échantillon de base Anand and Kanbur (1993b). En outre, comme le souligne Ahluwalia (1976) les résultats établis ci-dessus ne sont tout au plus que des descriptions de comportements moyens des pays de l'échantillon alors que des différences considérables peuvent exister entre les pays. Pour tenir compte de ces différences, Li et al. (1998) ont analysé la courbe de Kuznets en coupe instantanée. Malheureusement, leurs résultats mettent en évidence une certaine fragilité de la courbe de Kuznets. La relation dominante serait plutôt celle d'une relation inverse allant des inégalités vers la croissance. Ce dernier point fait l'objet de la sous-section suivante.

inégalités de revenu et croissance

Persson and Tabellini (1994) font partis des premiers auteurs à avoir souligné l'influence néfaste des inégalités de revenu sur la croissance. Même si le papier n'a été publié qu'en 1994, une version préliminaire circulait déjà en 1991. A l'aide d'un modèle d'économie politique, ils montrent qu'une forte concentration de la richesse induit une demande de redistribution des ressources des riches vers les pauvres. Cette dernière se fait à travers une taxe sur le revenu du capital (physique ou humain), et un transfert forfaitaire. En effet, pour accroître son revenu courant, l'électeur médian exige un taux de taxation sur le capital le plus élevé possible. Cependant, cela à un effet désincitatif sur l'accumulation du capital et donc déprime le taux de croissance de l'économie. Econométriquement, les auteurs cherchent à expliquer le taux de croissance. Pour ce faire, ils construisent deux échantillons : le premier est composé de 9 pays industriels et le second regroupe les données d'après-guerre de 56 pays sur la période 1960-1985. Les variables explicatives sont des mesures de l'inégalité des revenus et le degré de démocratisation. Ils incluent également dans toutes leurs régressions un effet de "convergence ", ici mesuré par le PIB initial par tête et le taux de scolarisa-

tion primaire. L'intérêt de cette dernière est de prendre en compte le capital humain. "N'apparaissent pas les variables endogènes mises en avant par la théorie, comme le taux d'investissement et le taux de scolarisation, au cours de la période analysée". A l'aide des MCO, ils trouvent une corrélation négative entre les inégalités de revenu et la croissance économique. Cependant, cette relation néfaste n'est vérifiée que dans les démocraties et non dans les pays non démocratiques. Les résultats semblent être robustes malgré les importantes erreurs de mesure.

Dans une approche semblable, Alesina and Rodrik (1994) montrent que le coefficient de Gini est corrélé négativement à la croissance économique. Cependant, contrairement à Persson and Tabellini (1994), ils ne trouvent pas d'interaction significative entre le degré de démocratisation et les inégalités de revenu. Cette différence peut être expliquée par le fait que les deux études reposent sur des séries distinctes de données pour les variables de distribution. En effet, les données de distribution dans Persson and Tabellini (1994) viennent de Paukert (1973), un pionnier dans l'étude de la courbe de Kuznets transversale. Alesina and Rodrik (1994) utilisent des sources différentes : (Jain et al., 1975; Lecaillon et al., 1984). La relation statistique exposée dans ces études s'est révélée bien plus faible qu'elle ne paraissait au premier abord et plusieurs études (Deininger and Squire, 1998; Barro, 2000; Forbes, 2000; Banerjee and Duflo, 2003) ont ensuite remis en question les résultats obtenus. Fishlow (1996) a montré la sensibilité de la corrélation à l'intégration de variables indicatrices régionales dans les variables explicatives. Il avance que l'inclusion d'une variable indicatrice pour les pays d'Amérique latine dans son échantillon rend le coefficient de Gini non significatif.

Dans un registre différent, Barro (2000) va plus loin que Fishlow (1996). En effet, il montre une sensibilité des résultats obtenus aux types d'échantillons retenus. Selon ses résultats, les inégalités de revenu seraient favorables à la croissance dans les pays riches et défavorables à la croissance dans les pays pauvres. Mais globalement, l'impact des inégalités de revenu sur la croissance et l'investissement est faible et statistiquement non significatif. Le modèle de Barro (2000) est une extension des modèles de croissance néoclassique.

Pour tester son modèle, Barro (2000) effectue une régression du taux de croissance moyen du PIB réel par habitant sur l'inégalité de revenu mesurée par l'indice de Gini et un ensemble d'autres variables explicatives en utilisant des données de panel pour 100 pays sur trois décennies 1965-75, 1975-85 et 1985-95. L'ana-

lyse empirique inclut des pays dont les niveaux de développement sont très différents, et sont exclus uniquement les pays dont les données sont manquantes. L'estimation de ce modèle a été faite par la méthode des triples moindres carrés ordinaires.

De son côté, Forbes (2000) trouve, à l'aide de l'estimateur GMM AB¹⁴, qu'une augmentation de dix points du coefficient de Gini est corrélée à un accroissement de 1,3% du taux moyen de croissance annuelle sur les cinq ans qui suivent. En outre, cette relation est robuste et statistiquement significative. Rappelons que le modèle de (Forbes, 2000) est très proche à celui de Perotti (1996). La seule différence est l'inclusion de variables dichotomiques (country et période) pour tenir compte des chocs sur la production et du biais dû aux variables manquantes.

Et parallèlement, Banerjee and Duflo (2003) proposent un modèle de "hold-up" où les opportunités de croissance sont plus ou moins efficacement exploitées, selon l'intensité des conflits distributifs qu'elles soulèvent. A l'aide d'une méthode non paramétrique, ils montrent que la relation entre les inégalités et la croissance est non linéaire. C'est seule la variation des inégalités (dans n'importe quelle direction) qui agit sur la croissance dans le long terme. En outre, selon ces auteurs, cette non linéarité est la principale explication des contradictions auxquelles on arrive sur le plan empirique.

Comme le montre la Table 1.4 ci-dessous, la plupart des papiers antérieurs s'accordent à l'idée qu'il existe une corrélation négative entre les inégalités de revenu et la croissance de long terme (par exemple (Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996) Venieris et Gutpa, 1996). Cependant, ce résultat est loin d'être consensuel. Par exemple, les estimations de Forbes (2000) font état d'une relation positive. Alors que pour Deininger and Squire (1996), cette relation reste faible et non significative. Ce dernier point à été confirmé par de nombreux travaux : Benhabib et Spiegel (1996), Fishlow (1996), entre autres. Toutefois, considérer que la distribution des revenus entre les individus n'agit pas sur les décisions d'investissement est une hypothèse envisageable mais difficile à accepter. D'ailleurs, comme nous l'avons souligné dans la section précédente, les inégalités jouent un rôle important dans l'évolution de la pauvreté.

14. L'estimateur d'Arrelano-Bond par la Méthode des Moments Généralisés.

TABLE 1.4 – Synthèse de quelques résultats empiriques

| Articles | Inégalité sur croissance ou investissement |
|------------------------------|--|
| Alesina and Rodrik (1994) | - |
| Barro (2000) | + - 0 |
| Benhabib et Spiegel (1996) | 0 |
| Brandolini et Rossi (1995) | 0 |
| Deininger and Squire (1998) | 0 |
| Forbes (2000) | + |
| Fishlow (1996) | 0 |
| Perotti (1996) | - |
| Persson and Tabellini (1994) | - |
| Venieris et Gutpa (1996) | - |

- : Les inégalités agissent négativement sur la croissance.

+ : Corrélation positive entre les inégalités et la croissance

0 : Relation faible et non significative entre les inégalités et la croissance.

1.3 Problèmes méthodologiques et qualités des données

Aujourd’hui, il est clairement admis dans la littérature que la relation entre les inégalités et la croissance est complexe. En effet, son signe varie d’une étude à l’autre, en raison non seulement des nombreuses spécifications retenues dans la littérature, mais également parce que les chercheurs ont pendant longtemps négligé les problèmes méthodologiques et les limitations inhérentes aux données. Dans cette section nous revenons en détail sur ces problèmes rencontrés dans la littérature empirique ainsi que leurs conséquences éventuelles sur les estimations.

1.3.1 La causalité inverse entre inégalités et croissance

La question de l'influence des inégalités de revenu sur le processus de croissance à toujours été considérée comme une préoccupation centrale des décideurs. Cependant, les interactions entre distribution du revenu et croissance sont complexes et loin d'être comprises d'ailleurs. La croissance affecte la distribution du revenu qui elle-même rétroagit sur la croissance et la pauvreté. En effet, le processus de développement d'une économie agit sur la distribution des ressources entre secteurs ; les rémunérations relatives des facteurs et les dotations des individus. Ces changements ont une influence directe sur l'évolution des inégalités, (et ceci) quel que soient les caractéristiques de la société concernée. La célèbre courbe de Kuznets s'identifie à cette idée. "Elle repose sur les différentiels de productivité entre secteurs et la modification de leur importance relative au cours du processus de développement, dans un contexte de marché du travail imparfait". Ainsi, ce différentiel de revenu résultant de ce processus crée les inégalités.

Dans les années plus récentes, une nouvelle approche est apparue sur la thématique des inégalités. Celles-ci ne sont plus considérées comme une conséquence de la croissance économique mais comme un déterminant majeur de la trajectoire de l'économie. Selon cette approche, les inégalités affectent le processus de croissance directement ou indirectement. A ce niveau d'ailleurs plusieurs thèses peuvent être avancées en guise d'exemple. Parmi lesquelles figure celle de l'instabilité socio-politique, du caractère indivisible de l'investissement, de l'effet désincitatif des politiques de redistribution etc. C'est cette causalité inverse qui est aujourd'hui clairement identifiée dans la littérature théorique qu'on appelle endogénéité.

Aujourd'hui, si des techniques d'estimations permettent d'isoler le sens de cette causalité, ce domaine reste peu exploré. En outre, les techniques de correction sont sensibles aux erreurs de mesure qui sont aussi non négligeables. Cette question de données est devenue un nouvel axe de recherche particulièrement intéressant dans le domaine des inégalités.

1.3.2 La qualité des données

"if science consists in a search for patterns in data, then the study of economic inequality suffers from an original sin : the job of measurement was badly done", Galbraith (2009).

En matière d'étude sur les inégalités de revenu la disponibilité et la qualité des données a pendant longtemps causé aux chercheurs de sérieux problèmes. Deininger and Squire (1996) sont probablement l'un des premiers auteurs à avoir soulevé cette question. Ils ont construit un nouvel ensemble de données améliorées et enrichies couvrant la période 1950-1997. Cependant, même dans leur base de données qui paraît être parmi les plus complètes à ce jour, ils identifient clairement ce problème. D'ailleurs, la particularité de leur travail est la ventilation des données en fonction de leur qualité. En effet, sur les 2600 observations recensées (2634 pour être plus précis) seule 682 ont été acceptées comme de bonne qualité¹⁵ soit environ 6% de l'échantillon. Par ailleurs, même sur ces données acceptées plusieurs auteurs refusent l'idée qu'il puisse exister une forte hétérogénéité entre les pays développés Atkinson and Brandolini (1999).

Aujourd'hui, les efforts visant à améliorer (ou élargir) les données de Deininger and Squire (1996) se poursuivent avec notamment les travaux du WIDER (World Institute for Development Research). Cependant, il est important d'insister sur le fait qu'aucune de ces séries de données n'est pleinement satisfaisante. Les données diffèrent entre les pays en termes de couverture géographique (couverture nationale, urbaine ou rurale), d'unités statistiques (familles, ménages ou individus) et de définition de revenu (dépenses de consommation, revenu disponible ou revenu brut). En outre, dans les pays de l'OCDE, les données de Deininger and Squire (1996) sont peu probables. Par exemple, d'une part l'Espagne apparaît comme un pays à faible inégalité malgré des années de dictature. D'autre part, les pays scandinaves malgré leur petite taille, des populations homogènes et une longue tradition de régime social-démocratique apparaissent avec des niveaux d'inégalités élevés Galbraith (2009).

Une erreur de mesure engendre une sous-estimation, en valeur absolue, de la vraie valeur du paramètre de la variable sujette à cette erreur de mesure. Cette

15. Selon Deininger and Squire (1996) les données sont de bonne qualité si elles remplissent les conditions suivantes : couverture nationale, source statistique clairement identifiée, méthode de calcul connue

sous-estimation est d'autant plus importante que la part de la variance de l'erreur de mesure dans la variance totale de la variable observée est élevée et/ou que le coefficient estimé est lui-même élevé. Et malheureusement ce problème s'empire dans les modèles de régression multiple. En effet, dans ce cas, cela induit une sous-estimation (respectivement une surestimation) du coefficient des autres variables, dès lors que leur covariance avec celle mesurée avec l'erreur est négative (respectivement positive).

1.3.3 L'hétérogénéité des données sur les inégalités de revenu

A l'heure actuelle, on doit admettre que les travaux empiriques sur les inégalités atteignent leurs limites. Premièrement, la disponibilité et la qualité des données continuent de poser de sérieux problèmes même si de réelles avancées ont été notées ces dernières années dans ce domaine avec notamment les travaux de Deininger and Squire (1996) et du WIDER. Deuxièmement, la littérature récente accepte largement l'idée d'une codétermination entre les inégalités et la croissance. Cependant, même si l'absence de données statistiques fiables et la codétermination des inégalités et la croissance (sont deux sources de problèmes qui) peuvent entamer sérieusement les résultats des estimations, ceux-ci ne peuvent pas tout expliquer. D'abord, parce que ces problèmes ne conduisent pas forcément à une divergence aussi prononcée. Ensuite et surtout parce qu'il existe aujourd'hui des techniques d'estimations assez efficaces pour contenir ces problèmes même si cette piste reste peu explorée. Actuellement, la véritable question qui mérite une attention particulière reste la forte hétérogénéité des données. La majorité des chercheurs s'accordent à l'idée que les estimations de la relation inégalités-croissance portent sur des données qui ne sont pas pleinement comparables. En outre, il est très probable que les contradictions auxquelles on arrive incombent en grande partie à cette forte hétérogénéité des données. Cet argument est d'autant plus justifié que les procédures d'estimations actuelles sont peu adaptées à ce type de données

Les données sur les inégalités sont critiquées par Atkinson and Brandolini (1999) en raison notamment de la très forte hétérogénéité des méthodologies de collectes. En d'autres termes, les données sont très déséquilibrées avec un minimum de 12,1 (chine 1982) et 79 (Zambie 1970). Par ailleurs, le classement de certaines zones par niveau d'inégalité semble peu plausible. Par exemple, les

moyennes d'inégalités pour l'Amérique latine et pour l'Afrique (respectivement 50,15 et 44) semblent trop élevées, alors que le niveau moyen de l'Asie de l'Est paraît trop faible 36. Enfin, "si l'on considère les différences entre pays, il est surprenant de constater la forte hétérogénéité des pays européens". Ces graves limitations ont conduit d'ailleurs Galbraith (2009) à construire un nouvel ensemble de données en se basant sur les différences de salaires dans le secteur industriel. Ces données de salaires ont été collectées dans pratiquement tous les pays du monde depuis le début des années 1960. Malheureusement, malgré une amélioration notable de la couverture et de la qualité des données, le jeu de données de Galbraith (2009) présente des limites, la plus grave est certainement les données retenues pour mesurer les inégalités. En effet, il est difficile de justifier comment les différences de salaires dans le secteur industriel puissent être utilisées pour tester l'impact de la distribution des ressources sur la croissance, puisqu'elles ne peuvent pas être considérées comme une variable indicatrice fiable pour les inégalités du revenu. Elles pourraient en revanche être utiles pour mesurer l'impact de la croissance sur les inégalités de revenu mesurées par les différences des salaires.

Face à ce problème d'hétérogénéité plusieurs pistes de recherches ont été envisagées, la plus sérieuse consiste à s'en tenir qu'aux seules données homogènes. Cette technique consiste à regrouper les pays (données) qui ont des caractéristiques assez proches en termes d'inégalité : exemple Barro (2000) montre que les inégalités sont favorables à la croissance dans les pays riches et défavorables dans les pays pauvres mais globalement non significatives. Ce type de traitement a été fréquemment repris dans la littérature empirique. Cependant, il n'est pas entièrement satisfaisant. En effet, ceci réduit considérablement la taille des échantillons ce qui n'est pas sans conséquence sur les résultats. Comme nous le savons l'estimateur GMM-Arellano -Bond largement utilisé dans les travaux empiriques antérieurs est biaisé vers le bas en petit échantillon. Ce qui fait que la significativité des paramètres estimés s'avère particulièrement douteuse. En effet, l'étude par simulation d'Arellano et Bond montre la possibilité d'une sous-estimation de 20 à 30% des écarts-types des coefficients en petit échantillon. Autrement dit quand vous divisez votre échantillon les T-stat "explosent" ce qui augmente la probabilité de rejeter l'hypothèse nulle alors qu'elle est vraie c'est-à-dire de conclure à

tort à la significativité des paramètres estimés¹⁶. En outre, il est important de noter que l'estimateur d'Arellano and Bond (1991) a été proposé dans le but d'éviter les deux causes d'inefficacité de l'estimateur en différence première connu sous le nom de l'estimateur d'Anderson et Hsiao. En effet, ce dernier est extrêmement sensible au biais de simultanéité et aux erreurs de mesure. Mais malheureusement, tout comme l'estimateur d'Anderson et Hsiao, les estimateurs d'Arellano et Bond souffrent de la perte d'information associée à l'écriture du modèle en différences premières ; perte d'information qui se traduit souvent par des estimateurs peu précis et parfois erratiques Baltagi (2008).

Par ailleurs, cette approche qui consiste à partitionner les échantillons en regroupant les données homogènes bien qu'intéressante ne résout en rien le problème lié aux points aberrants ("outliers"). Ce problème est bien connu en économétrie. Considérons un PED qui ne maîtrise pas les techniques de production et dont la production est agricole. Sa production peut tripler en une année à la suite d'une bonne pluviométrie, ou être divisée par deux après une période de sécheresse. Ces observations isolées ont souvent posé aux économètres de réelles difficultés. Traditionnellement, les chercheurs éliminent ces points "extrêmes" pour estimer ce type de modèle (méthode jugée non rigoureuse) ou construisent un panel pour limiter cette forte variabilité. En effet, un avantage décisif des données de panel par rapport aux autres types de données tient au fait qu'elles sont souvent extrêmement nombreuses. En panel, il est très facile de disposer d'échantillon de plusieurs milliers d'observation. Le principal avantage que présente la grande taille de ces échantillons et qu'ils conduisent systématiquement à des estimateurs dont les propriétés peuvent être assimilées aux propriétés asymptotiques des méthodes utilisées. Ainsi, la mise en oeuvre d'un estimateur convergent conduit à un estimateur très proche de la vraie valeur des paramètres. En réalité, le nombre élevé des observations permet de considérer que le biais et la variance des estimations tendent vers zéro. De fait, l'un des intérêts des estimateurs basés sur un très grand nombre d'observations est leur grande précision.

Toutefois, d'abord dans les données portant sur les inégalités de revenu, la variabilité est tellement forte que même en panel des pays entiers apparaissent

16. Notons au passage que Windmeijer (2005) a développé une méthode de correction de ces écarts-types en petit échantillon. Cependant, les contraintes rédactionnelles ne nous permettent pas d'intégrer cette méthode dans ce travail.

comme des outliers¹⁷. ” C’est le cas du Sénégal, Mali, Uganda, Pologne et Hongrie ” Deininger and Squire (1998). En outre, puisqu’un échantillon de données de panel contient un très grand nombre d’observations, on peut penser a priori que cet ensemble d’observations est capable de ”neutraliser” l’influence de quelques points aberrants. En réalité, ce n’est pas vraiment le cas. Huber (1980) a montré que seulement 3% de valeurs atypiques dans un ensemble d’observations suffisent à perturber significativement la qualité des estimations.

Pour conclure sur ces difficultés associées aux observations aberrantes, il faut noter que, jusqu’à une passé récente, il était d’usage d’éliminer toutes les observations relatives à un individu, dès lors que l’une d’elles apparaissait comme aberrante. Une telle pratique (qu’on appelle cylindrage) conduisait en effet souvent à éliminer une très forte proportion des observations et donc, à perdre beaucoup d’informations ; ceci induisant, au mieux, une moindre précision des estimations et au pire, un risque de biais. C’est pourquoi d’ailleurs que certains chercheurs ont préféré corriger ces points aberrants au lieu de les éliminer. Malheureusement, dans ce cas aussi il est indispensable que l’information disponible soit suffisante pour que la correction soit acceptable.

1.4 Conclusion

Il est aujourd’hui évident qu’un long travail est encore nécessaire sur cette question d’évaluation de la croissance pro-pauvre. En effet, malgré la grande diversité des mesures, à l’heure actuelle, il en n’existe pas une faisant l’unanimité. Tout d’abord, ces indices peuvent parfois conduire à des résultats contradictoires. En outre, bien que plusieurs travaux aient montré le caractère multidimensionnel de la pauvreté et l’importance de prendre en contact la dimension non monétaire (éducation, santé, inégalité genre etc), tous ces indices se basent sur la dimension monétaire de la pauvreté. Ce qui peut paraître très restrictif. Ces travaux (sur les indices de mesures de croissance pro-pauvre) font l’hypothèse qu’il existe un lien étroit entre la dimension monétaire et non monétaire de la pauvreté. En effet, ils soutiennent qu’en générale une croissance du revenu induit souvent une croissance non-monétaire et inversement. L’idée est qu’un accroissement du revenu des pauvres leurs permettent d’accéder à l’éducation et à la santé et inver-

17. The biggest outliers are Mali, Senegal, Uganda, Poland and Hungary.

sement une meilleure éducation (meilleur état de santé) favorisent les pauvres à accroître leurs revenus ou leurs consommations. Malheureusement, le lien entre les dimensions monétaire et non monétaire de la pauvreté n'est pas aussi simple (voir Drèze and Sen, 1989 ; Klasen, 2000, 2002a). En fait, comme plusieurs travaux le montrent il arrive souvent qu'un individu dispose des ressources monétaires nécessaires pour satisfaire ses besoins de bases, sans que certains biens ou infrastructures ne soient pas disponibles dans la localité où il demeure. Par ailleurs, la prise en compte de la dimension non-monétaire permet aux décideurs l'instauration de politiques plus adaptées favorisant la réduction des risques de distorsion par rapport à une intervention d'ordre général. C'est pourquoi, ces années récentes ont vu naître une approche alternative. Cette dernière utilise généralement des modèles économétriques afin d'établir une causalité entre croissance pro-pauvre et indicateurs de bien-être. Cette approche présente l'avantage d'inclure la dimension non monétaire de la pauvreté. Malheureusement, elle est aussi critiquable à plusieurs niveaux. En premier lieu, l'objectif principal était de tenter d'identifier la valeur de l'élasticité du taux de pauvreté au revenu moyen. En second lieu, la grande majorité des données sont de piètres qualités et sont sujettes à d'importantes erreurs de mesure. D'ailleurs, beaucoup de travaux ont montré la sensibilité des résultats aux formes fonctionnelles choisies.

Par ailleurs, étant donné que l'ampleur du recul de la pauvreté sous l'effet de la croissance varie considérablement d'un pays à l'autre et d'un épisode de croissance à l'autre, les meilleurs résultats sont obtenus là où les inégalités sont faibles et/ou là où les politiques en place améliorent la capacité des pauvres à bénéficier de la croissance. Considérer la croissance dans une perspective pro-pauvre impose donc de comprendre les interactions entre distribution de revenu et croissance. Il apparaît ainsi clair que de plus amples recherches seront nécessaires pour déterminer la nature de cette relation. Sur ce point la régression Gini semble être une approche particulièrement intéressante. En effet, comme nous allons le voir dans le prochain chapitre, ce type de régression présente plusieurs avantages : elle est moins sensible que les méthodes traditionnelles aux erreurs de mesures et aux points aberrants. Mais aussi, elle est plus robuste car les estimateurs obtenus par cette approche sont insensibles aux formes fonctionnelles retenues.

Chapitre 2

La méthodologie Gini

Durant ces dernières décennies, une attention particulière a été portée au coefficient de Gini, à la courbe de concentration et à la courbe de Lorenz. En effet, des chercheurs d'horizons différents ont redécouvert l'intérêt d'utiliser ces paramètres dans de nombreuses disciplines. Par exemple Atkinson (1970) souligne l'avantage d'utiliser la courbe de Lorenz dans le cas de la finance, (Kakwani, 1977, 1980) montre que la courbe de Lorenz et la courbe de concentration permettent d'estimer l'élasticité des biens de consommation de manière non paramétrique et plus récemment, Olkin and Yitzhaki (1992) puis (Yitzhaki and Schechtman, 2004, et 2013) ont développé et popularisé la régression Gini qui consiste à utiliser la différence moyenne de Gini (GMD) comme mesure de dispersion au lieu de la variance.

De toutes les mesures de dispersion la variance est la plus connue. En statistique comme dans beaucoup d'autres disciplines la variance et la covariance sont les deux principales mesures utilisées pour mener à bien une étude. La technique des MCO qui consiste à minimiser la variance des erreurs est un cas d'école en économétrie. Les estimateurs des MCO jouissent de très bonnes propriétés tant à distance finie qu'asymptotiquement : absence de biais, convergence et efficacité. Ceci ne signifie pas pour autant que le recours à d'autres estimateurs soit sans grand intérêt. En réalité, dans la littérature économétrique avancée, elle constitue un point de départ d'analyses plus approfondies. En effet, la bonne propriété des estimateurs des MCO repose sur des hypothèses assez fortes : l'hypothèse de linéarité, la stricte exogénéité, homoscedasticité, etc. En outre, son utilisation devient difficile à justifier lorsqu'il existe des points extrêmes dans l'échantillon

ou plus généralement lorsque l'hypothèse de normalité n'est plus satisfaite. Les estimateurs des MCO sont sensibles aux observations aberrantes pour la simple raison que cette méthode vise à minimiser la somme des carrés des résidus. Ceci fait que les grands résidus (positifs ou négatifs) reçoivent un poids plus important.¹ Par ailleurs, l'estimateur des MCO est sensible à la transformation monotone des données Yitzhaki (1990). Une approche alternative pour surmonter cette extrême sensibilité de l'estimateur des MCO aux observations isolées consiste à utiliser une autre mesure de la dispersion des résidus moins sensible aux valeurs aberrantes que la variance comme la différence moyenne de Gini (GMD) par exemple. Cette dernière a été introduite en 1912 par Corrado Gini, et depuis, plusieurs paramètres ont été déduits comme la covariance au sens de Gini (co-Gini), la corrélation au sens de Gini, l'analyse du Gini (ANOGI).

Comme nous le verrons dans ce chapitre, la GMD partage plusieurs propriétés avec la variance ; toutefois elle fournit plus d'information sur la distribution sous-jacente des données. En réalité, ces deux mesures sont très similaires, la différence fondamentale tient sur la norme utilisée. La variance se base sur la distance euclidienne (norme ℓ_2) tandis que la GMD utilise la "city block" distance (norme ℓ_1) (Ka and Mussard, 2015).

Il est intéressant de noter qu'il existe deux types de régression Gini. La première consiste à minimiser la GMD des erreurs. Les estimateurs obtenus par cette approche ont la propriété d'optimalité. La seconde approche consiste à construire l'estimateur en question de sorte qu'il soit moins sensible aux outliers. Cette seconde approche, puisqu'elle ne découle pas d'un programme de minimisation et par conséquent ne requiert pas de définir une forme fonctionnelle quelconque, est qualifiée de régression semi-paramétrique. L'estimateur obtenu par cette dernière approche est similaire à celui obtenu par les MCO ; excepté que chaque variance est remplacée par la GMD et chaque covariance par la co-Gini appropriée.

D'un point de vue théorique et empirique, un travail important est encore nécessaire sur la régression Gini notamment en panel. En effet, introduite par Olkin and Yitzhaki (1992), la réflexion sur la régression Gini a été étendue aux variables

1. Sur le plan pratique, l'existence d'observations extrêmes peut s'expliquer de deux façons. Le cas le plus simple à traiter est celui où une erreur s'est glissée au moment de l'encodage des données. Par exemple, ajouter des zéros ou déplacer la virgule peut, en effet, détériorer sensiblement les estimateurs par MCO. Elles peuvent également apparaître lorsque l'échantillon est tiré d'une petite population au sein de laquelle une ou plusieurs unités sont très différentes du reste de la population Wooldridge (2015).

instrumentales par Yitzhaki and Schechtman (2004) et aux séries temporelles par Serfling (2010) et par Shelef and Schechtman (2011). *Nous proposons de l'étendre aux données de panel.* C'est dans ce cadre que nous avons développé l'estimateur par effets fixes en utilisant la GMD comme mesure de dispersion Ka and Mussard (2016). *Nous proposons également le test de l'existence de spécificités individuelles et l'estimateur du modèle à effets aléatoires.* L'intérêt essentiel de développer la régression Gini en panel réside dans la volonté de pouvoir exploiter de mieux en mieux les avantages spécifiques de ces données. En particulier, la possibilité de prendre en compte les outliers. Puisqu'un échantillon de données de panel contient un très grand nombre d'observations, on peut penser a priori que cet ensemble d'observations est capable de "neutraliser" l'influence de quelques points aberrants. En réalité, ce n'est pas vraiment le cas. Huber (1980) a montré que seulement 3% de valeurs atypiques dans un ensemble d'observations suffisent à perturber significativement la qualité des estimations. Plus généralement, dans le cadre de l'économétrie sur données de panel, et comme nous le verrons d'ailleurs dans le chapitre suivant, quelques points atypiques peuvent suffire à modifier considérablement les résultats d'une estimation.

La première section de ce chapitre revient sur les généralités de la méthode Gini. Nous allons y présenter entre autres la différence moyenne du Gini, le co-Gini, la corrélation au sens de Gini et les différentes méthodes de décomposition du Gini. La seconde section s'intéresse aux régressions Gini, à leurs propriétés et aux hypothèses sur la matrice des moments. Une remarque additionnelle sera également portée au traitement des ex-æquo. Dans la troisième section nous présenterons le modèle à effets fixes dans le cas du Gini et le test d'existence de l'effet individuelle et enfin la dernière section constitue nos derniers développements théoriques.

2.1 Généralités sur la méthodologie de Gini

2.1.1 La différence moyenne de Gini et Co-Gini

Le bilan des travaux littéraires révèle qu'il existe plusieurs définitions de la différence moyenne du Gini. Il semble qu'il existe plus d'une douzaine de représentations alternatives du GMD. La (représentation) plus connue est celle qui dé-

fini la GMD comme l'écart espéré entre deux réalisations tirées au hasard (avec remise) d'une population. Soit x_i et x_j deux réalisations (d'une même variable aléatoire x) indépendantes et identiquement distribuées. Formellement, pour une variable aléatoire x à valeurs positives, la GMD se présente de la manière suivante :

$$GMD(\mathbf{x}) = \mathbb{E} \{ |x_i - x_j| \} \quad (2.1)$$

Une présentation alternative assez intéressante est donnée par les formes suivantes :

$$GMD(\mathbf{x}) = 2 \int \mathbf{F}_x(t) [1 - \mathbf{F}_x(t)] dt, \quad (2.2)$$

avec \mathbf{F}_x la fonction de répartition de la variable aléatoire x , ou encore,

$$GMD(\mathbf{x}) = 4\text{Cov} [x, \mathbf{F}_x(x)]. \quad (2.3)$$

L'un des attraits essentiels de cette dernière présentation est qu'elle ouvre la voie à plusieurs développements. En effet, la présentation du GMD en forme de covariance conduit naturellement à la corrélation au sens de Gini et plus généralement à la régression Gini – et à beaucoup d'autres développements théoriques comme la régression Gini avec variables instrumentales, les séries temporelles en Gini, etc.

En réalité, la GMD partage plusieurs propriétés avec la variance, la différence fondamentale tient sur la norme utilisée. La variance est définie par :

$$\text{Var}(x) = \frac{1}{2} \mathbb{E} (x_i - x_j)^2, \quad (2.4)$$

alors que la GMD est :

$$GMD(\mathbf{x}) = \mathbb{E} \{ |x_i - x_j| \}. \quad (2.5)$$

Ainsi, comme nous pouvons le remarquer, la principale différence entre ces deux mesures de dispersion est la norme utilisée. La variance se base sur la distance euclidienne (ℓ_2) tandis que la GMD utilise la "city block" distance (ℓ_1). La première des propriétés qu'elles partagent est le fait que ces deux mesures peuvent être écrites en covariance :

$$\text{Var}(\mathbf{x}) = \text{Cov}(\mathbf{x}, \mathbf{x}), \quad (2.6)$$

$$GMD(\mathbf{x}) = 4\text{Cov}(\mathbf{x}, \mathbf{F}_x(x)). \quad (2.7)$$

Ainsi, on peut exprimer la GMD comme la pente d'une régression par MCO. En effet, puisque $\mathbf{F}_x(x)$ est une fonction cumulative, elle est distribuée selon une loi uniforme sur $[0;1]$, alors $\text{Var}(\mathbf{F}_x(x)) = \text{Cov}(\mathbf{F}_x(x), \mathbf{F}_x(x)) = 1/12$. Par conséquent :

$$GMD(\mathbf{x}) = 4\text{Cov}(\mathbf{x}, \mathbf{F}_x(x)) = \frac{1}{3} \frac{\text{Cov}(\mathbf{x}, \mathbf{F}_x(x))}{\text{Var}(\mathbf{F}_x(x))}. \quad (2.8)$$

En outre, comme on peut le voir à partir de (2.4) et (2.5), les deux indices peuvent être définis sans référence à un paramètre de localisation spécifique, tel que la moyenne. Cette propriété peut expliquer pourquoi il est possible de contraindre les régressions MCO et Gini à passer par la médiane (Yitzhaki et al., 2003).

Une autre similarité entre ces deux mesures de dispersion est le fait que la décomposition de la variance peut être vue comme un cas particulier de la décomposition de la GMD. En effet, comme nous allons le voir dans la sous-section suivante, la décomposition de la GMD est très similaire à celle de la variance, la différence fondamentale tient au fait qu'elle fait intervenir un terme additionnel. En réalité, lorsque la distribution des données est normale la GMD et la variance sont quasiment identiques. Lorsque \mathbf{x} suit une loi normale on a :

$$\text{Var}(\mathbf{x}) = \frac{\sqrt{\pi}}{2} GMD(\mathbf{x}). \quad (2.9)$$

Cette dernière expression nous permet de comprendre pourquoi les estimateurs sont identiques dans le cas des MCO et du Gini lorsque la distribution des régresseurs est normale. Pour une présentation plus générale des similitudes et des différences entre la GMD et la variance le lecteur intéressé pourra se référer à (Yitzhaki et al., 2003) et (Yitzhaki and Schechtman, 2013, p. 21-22).

Plus généralement, la GMD peut être vue comme un élément d'un ensemble ou d'une famille de paramètres que l'on appelle l'indice de Gini généralisé ("the extended Gini") :

$$EG(\mathbf{x}; r) = \mathbb{E} |\mathbf{x}_i - \mathbf{x}_j|^r, \quad r > 0. \quad (2.10)$$

Lorsque $r = 1$, on retrouve la GMD et lorsque $r = 2$ nous retrouvons deux fois la variance ; ce qui nous permet de comprendre que la variance est un cas particulier de l'indice de Gini généralisé. Notons que ce dernier peut aussi s'écrire sous la forme suivante :

$$EG(\mathbf{x}; \nu) = -(\nu + 1)\text{Cov}[\mathbf{x}, (1 - \mathbf{F}_x(x))^\nu], \nu > 0. \quad (2.11)$$

Traditionnellement, le Gini généralisé est utilisé en finance (Shalit and Yitzhaki, 1984) ou pour modéliser la distribution des revenus² (Stark et al., 1988; Moyes, 2007; Tibiletti and Subramanian, 2015). Dans chaque cas, respectivement, ν représente le degré d'aversion au risque de l'investisseur (ou d'aversion aux inégalités du décideur public). Lorsque $\nu \rightarrow 0$, l'investisseur est riscophile (le décideur aime les inégalités), $\nu = 1$ traduit un décideur qui à un comportement d'aversion au risque (le décideur respecte le principe de transferts de Pigou-Dalton), et enfin $\nu \rightarrow \infty$ décrit un investisseur complètement averse au risque (un décideur Rawlsian qui ne redistribuerait des revenus qu'à l'individu le plus pauvre de la société).

L'avantage principal d'utiliser l'indice de Gini généralisé dans une régression réside dans la possibilité de tester l'ajustement aux données en changeant le paramètre ν . Il est à noter qu'il existe plusieurs autres définitions alternatives du GMD – et du Gini généralisé – que le lecteur intéressé pourra retrouver dans Yitzhaki (1998) et (Yitzhaki and Schechtman, 2013, chap2 et chap6). L'intérêt essentiel que présente la multiplicité de ces définitions est de montrer l'avantage d'utiliser la GMD. Chaque définition ouvre la voie à de nombreux développements ainsi qu'à de nombreuses interprétations. Par exemple, la définition basée sur la covariance permet de définir la covariance au sens de Gini (que nous nommerons co-Gini ou Cog). Cette dernière introduite par Schechtman and Yitzhaki (1987) n'est pas symétrique. En effet,

$$\text{Cog}(\mathbf{x}, \mathbf{y}) = \text{Cov}[\mathbf{x}, \mathbf{F}_y(y)] \quad (2.12)$$

et

$$\text{Cog}(\mathbf{y}, \mathbf{x}) = \text{Cov}[\mathbf{y}, \mathbf{F}_x(x)] \quad (2.13)$$

où $\text{Cog}(\mathbf{x}, \mathbf{y})$ et $\text{Cog}(\mathbf{y}, \mathbf{x})$ ne sont pas forcément égales. En réalité, elles permettent d'avoir une idée sur la dépendance entre les variables aléatoires \mathbf{x} et \mathbf{y} .

2. Notons que le Gini généralisé est très proche de l'indice d'Atkinson Stark et al. (1988).

Elles s'expriment aussi de la manière suivante :

$$\text{Cog}(\mathbf{x}, \mathbf{y}) = 2 \int \int x(2\mathbf{F}_{\mathbf{y}}(y) - 1)d\mathbf{F}(x, y), \quad (2.14)$$

$$\text{Cog}(\mathbf{y}, \mathbf{x}) = 2 \int \int y(2\mathbf{F}_{\mathbf{x}}(x) - 1)d\mathbf{F}(x, y). \quad (2.15)$$

Traditionnellement, deux approches ont été développées dans la littérature pour étudier la relation entre deux variables aléatoires \mathbf{x} et \mathbf{y} . La première approche se base sur la covariance entre ces deux variables ($\text{Cov}(\mathbf{x}, \mathbf{y})$). Cette approche a conduit naturellement à la variance et au coefficient de corrélation de Pearson. La seconde considère la covariance entre les fonctions cumulatives de \mathbf{x} et de \mathbf{y} : $\text{Cov}(\mathbf{F}_{\mathbf{x}}(x), \mathbf{F}_{\mathbf{y}}(y))$. Cette approche a conduit au coefficient de corrélation de Spearman. Plus récemment, des statisticiens ont proposé une troisième approche. Cette dernière se présente comme une combinaison des deux précédentes. En effet, elle se base sur la covariance entre la variable et la fonction cumulative de l'autre variable, c'est-à-dire la co-Gini 2.12 et 2.13

Avant de présenter la régression Gini, les recherches sur la décomposition en sous-groupes de l'indice de Gini doivent être présentées car elles permettent de découvrir des estimateurs robustes basés sur l'ANOVI, cf. Ka and Mussard (2015). De même, les propriétés de la corrélation au sens de Gini basée sur la co-Gini doivent être analysées.

2.1.2 Corrélation au sens de Gini : la G-corrélation

Il existe plusieurs mesures d'association dans la littérature. Les plus connues sont le coefficient de corrélation de Pearson, le coefficient de corrélation de Spearman et le tau de Kendall. Dans cette sous-section nous présenterons la corrélation au sens de Gini.

Le coefficient de corrélation de Pearson est :

$$\rho_{\mathbf{x}, \mathbf{y}} = \frac{\text{Cov}(\mathbf{x}, \mathbf{y})}{\text{Var}(\mathbf{x})\text{Var}(\mathbf{y})} = \frac{\mathbb{E}[\mathbf{x} - \mathbb{E}(\mathbf{x})] \mathbb{E}[\mathbf{y} - \mathbb{E}(\mathbf{y})]}{\text{Var}(\mathbf{x})\text{Var}(\mathbf{y})}. \quad (2.16)$$

Lorsqu'il est égal à $1(-1)$, on peut conclure qu'il existe une relation affine croissante (décroissante) entre les variables. Une corrélation égale à 0 signifie que les variables ne sont pas corrélées (la réciproque est fausse). En effet, le coefficient de Pearson indique une dépendance linéaire. Par conséquent, il est inap-

proprié pour étudier des relations qui ne seraient pas linéaire et non linéarisable (exponentielle, puissance, etc.). En outre, il est extrêmement sensible aux points aberrants.

D'une manière générale, un coefficient de corrélation proche de (± 1) est associé à une forte association entre les variables. Au contraire, un coefficient de corrélation proche de 0 indique une absence d'association. Malheureusement, ces deux assertions posent problèmes. Tout d'abord, un coefficient de corrélation proche de 0 ne correspond pas toujours à une absence d'association, *cf.* Denuit and Dhaene (2003). Par exemple, il peut arriver que deux variables aléatoires aient un coefficient de corrélation proche de 0 alors qu'elles sont liées par une transformation monotone. Ensuite, les distributions multivariées peuvent avoir naturellement des coefficients de corrélation compris dans l'intervalle $[-1/3; 1/3]$.

Le coefficient de corrélation de Spearman qui est une mesure non paramétrique permet de rendre compte de l'intensité de la relation entre des variables non linéaires (ou linéaires) et monotone. Il est fondé sur le rang des variables et est préférable au coefficient de Pearson lorsque la distribution comporte des valeurs extrêmes. En réalité, il est un cas particulier du coefficient de Pearson, en effet, c'est le coefficient de Pearson lui-même appliqué au rang des variables. Notons $R_i(\mathbf{x})$ le rang de la i ème observation de la variable aléatoire \mathbf{x} . Ainsi, le coefficient de corrélation de Spearman (R_s) entre les variables \mathbf{x} et \mathbf{y} est :

$$R_{s_{\mathbf{x},\mathbf{y}}} = \frac{\sum_{i=1}^n \left(R_i(\mathbf{x}) - \frac{n+1}{2} \right) \left(R_i(\mathbf{y}) - \frac{n+1}{2} \right)}{\frac{n(n^2-1)}{12}} = 1 - \frac{6 \sum_{i=1}^n (R_i(\mathbf{x}) - R_i(\mathbf{y}))^2}{n^2 - n}. \quad (2.17)$$

Le coefficient de corrélation de Spearman est compris entre -1 et 1 , il s'interprète comme le coefficient de Pearson, excepté qu'il tient compte des relations non linéaires entre les variables aléatoires \mathbf{x} et \mathbf{y} .

La corrélation au Sens de Gini a été introduite et popularisée par Schezhtman and Yitzhaki (1987). Elle peut être définie comme le rapport entre la co-Gini et la GMD appropriée. Comme nous l'avons vu précédemment, il existe deux co-Gini entre deux variables aléatoires \mathbf{x} et \mathbf{y} . Par conséquent, il existe deux Gini-corrélations (G-corrélation) lorsque les variables \mathbf{x} et \mathbf{y} sont permutées. Les G-corrélations sont définies comme suit :

$$\mathbf{G}_{x,y} = \frac{\text{Cov}(x, F_y(y))}{\text{Cov}(x, F_x(x))}; \mathbf{G}_{y,x} = \frac{\text{Cov}(y, F_x(x))}{\text{Cov}(y, F_y(y))}. \quad (2.18)$$

Lorsque que $\mathbf{G}_{x,y}$ ou $\mathbf{G}_{y,x}$ est égale à 1 ou (-1) , on peut dire que y est une fonction monotone croissante (décroissante) de x . Des coefficients égaux à zéro indiquent une indépendance statistique entre x et y . $\mathbf{G}_{x,y}$ est invariant à toute transformation strictement monotone de y . Pour terminer :

1. $-1 \leq \mathbf{G}_{x,y} \leq 1$.
2. Lorsque x est une fonction monotone croissante (décroissante) de y alors $\mathbf{G}_{x,y}$ et $\mathbf{G}_{y,x}$ sont égaux à $1(-1)$.
3. Si y et x sont indépendantes alors $\mathbf{G}_{x,y} = \mathbf{G}_{y,x} = 0$.
4. $\mathbf{G}_{x,y} = -\mathbf{G}_{x,-y} = -\mathbf{G}_{-x,y} = \mathbf{G}_{-x,-y}$.
5. $\mathbf{G}_{x,y}$ est invariant à toute transformation monotone stricte de y .
6. $\mathbf{G}_{x,y} = \mathbf{G}_{y,x}$ si et seulement si $\mathbf{G}_{x,y}$ est symétrique en (x, y) lorsque $ax + B$ et $cy + D$ sont permutées.
7. Lorsque (x, y) suit une loi normale bivariée alors $\mathbf{G}_{x,y} = \mathbf{G}_{y,x} = \rho_{x,y}$.

2.1.3 Décomposition de la différence moyenne du Gini

Aujourd'hui, la littérature sur la décomposition de la différence moyenne de Gini et plus généralement de l'indice de Gini est extrêmement élaborée. Comme nous le savons plusieurs méthodes statistiques telles que les MCO sont basées sur la décomposition de la variance. Cette dernière fréquemment utilisée présente une limite majeure. Elle n'est applicable que lorsque les populations échantillonnées sont distribuées selon la loi de Laplace-Gauss. Or, en réalité cette condition n'est qu'approximativement remplie, bien que certains travaux tentent de montrer que l'analyse de la variance est peu sensible à la non normalité. Il est souvent préférable d'utiliser une autre décomposition non soumise à l'hypothèse de normalité, comme l'ANOVI. Il existe deux méthodes de décomposition de la GMD : la décomposition d'une combinaison linéaire de variables aléatoires et la décomposition en sous-groupes (ANOVI).³

3. Notons qu'une combinaison de ces deux méthodes de décomposition est possible, cf. Muscard and Richard (2012).

Décomposition d'une combinaison linéaire de variable

Soit une variable aléatoire y à valeurs réelles décomposable linéairement comme suit :

$$\mathbf{y} = \sum_{k=1}^K \beta_k \mathbf{x}_k, \quad (2.19)$$

avec β_k un paramètre réel donné et \mathbf{x}_k des variables aléatoires à valeurs réelles. L'écart-type de \mathbf{x}_k est noté σ_k , ainsi la variance de y est donnée par l'expression suivante :

$$\text{Var}(\mathbf{y}) = \sum_{k=1}^K \beta_k^2 \mathbf{x}_k \text{Var}(\mathbf{x}_k) + \sum_{k \neq j} \beta_k \beta_j \sigma_k \sigma_j \rho_{jk}, \quad (2.20)$$

où ρ_{jk} représente le coefficient de corrélation linéaire (de Pearson) entre les variables \mathbf{x}_j et \mathbf{x}_k . En utilisant les propriétés de la covariance nous pouvons également décomposer la GMD. Il faudra toutefois noter que contrairement à la variance il existe deux types de décomposition qu'on appelle habituellement la décomposition marginale ("one step decomposition") et la décomposition complète ("two-step decomposition").

1. Décomposition marginale du GMD :

$$\frac{1}{4} GMD \left(\sum_{k=1}^K \beta_k \mathbf{x}_k \right) = \sum_{k=1}^K \beta_k \text{Cog}(\mathbf{x}_k, \mathbf{y}) \quad (2.21)$$

$$= \sum_{k=1}^K \beta_k \frac{\text{Cov}(\mathbf{x}_k, \mathbf{F}_y(y))}{\text{Cov}(\mathbf{x}_k, \mathbf{F}_{\mathbf{x}_k}(x_k))} \text{Cov}(\mathbf{x}_k, \mathbf{F}_{\mathbf{x}_k}(x_k)) \quad (2.22)$$

$$= \sum_{k=1}^K \beta_k \mathbf{G}_{\mathbf{x}_k, \mathbf{y}} \text{Cog}(\mathbf{x}_k, \mathbf{x}_k) \quad (2.23)$$

$$= \frac{1}{4} \sum_{k=1}^K \beta_k \mathbf{G}_{\mathbf{x}_k, \mathbf{y}} GMD(\mathbf{x}_k). \quad (2.24)$$

Ce type de décomposition du Gini est généralement utilisé dans le domaine de la finance et des inégalités et particulièrement dans le cas de la taxation. Par exemple, ce type de décomposition peut nous permettre d'appréhender l'impact de la variation du revenu suite à une taxation de l'une de ses composantes (salaires, primes, etc.).

2. Décomposition complète du GMD :

$$GMD^2 \left(\sum_{k=1}^K \beta_k \mathbf{x}_k \right) = \sum_{k=1}^K \beta_k^2 GMD^2(\mathbf{x}_k) + \sum_{k \neq j} \beta_k \beta_j GMD(\mathbf{x}_k) GMD(\mathbf{x}_j) \mathbf{G}_{\mathbf{x}_k, \mathbf{x}_j}. \quad (2.25)$$

Cette décomposition plus complexe que la première est basée sur une imitation de la décomposition de la variance, *cf.* Wodon et al. (2003).

Décomposition en sous-groupes

Soient y_i , μ et p_k représentant respectivement le revenu de l'individu i , la moyenne des revenus de la population et la part de population du sous-groupe k dans la population totale. Soit $S_k = p_k \times \mu_k / \mu$ la part que représente le revenu du groupe k dans le revenu total avec μ_k le revenu moyen du groupe k . La population est décomposée en k groupes exclusifs et exhaustifs tels que $k = 1, \dots, K$.

La décomposition du Gini de la population totale noté G_u peut être présentée de la manière suivante :

$$G_u = \sum_{k=1}^K S_k G_k O_k + G_b, \quad (2.26)$$

où,

$$G_b = \frac{2\text{Cov}(\bar{\mathbf{y}}, \overline{\mathbf{F}_y})}{\mu} \quad (2.27)$$

représente le Gini intergroupe (l'inégalité entre les groupes k), G_k représente le Gini intragroupe (l'inégalité dans le groupe k) et,

$$O_k = \frac{\text{Cov}(\mathbf{y}, \mathbf{F}_y(y))}{\text{Cov}(\mathbf{y}, \mathbf{F}_{y_k}(y_k))} \quad (2.28)$$

représente le chevauchement entre les densités de probabilités des différents groupes k . Comparée à l'analyse de la variance qui décompose la variance totale en variance intragroupe et variance intergroupe, l'ANOVI fait intervenir un terme additionnel qui représente en réalité le degré de stratification entre les distributions des différents groupes. Le terme d'interaction (stratification) est nul lorsque les distributions ne se chevauchent pas, par exemple, lorsque la population est divisée en quantiles. L'indice de chevauchement total est composé de l'indice de chevauchement du groupe k avec toutes les autres groupes, k inclus

également :

$$O_k = \sum_{j=1}^n p_j O_{kj} = p_k O_{kk} + \sum_{j \neq i} p_j O_{kj} = p_k + \sum_{j \neq i} p_j O_{kj} \quad (2.29)$$

avec,

$$O_{kj} = \frac{\text{Cov}(\mathbf{y}, \mathbf{F}(y))}{\text{Cov}(\mathbf{y}, \mathbf{F}_{y_k}(y))} \quad (2.30)$$

Le terme O_{kj} représente le chevauchement du groupe k avec le groupe j , et $O_{kk} = 1$ puisqu'il représente l'indice de chevauchement du groupe k avec lui-même. Pour une description plus complète de l'indice de chevauchement et ses propriétés voir Dagum (1997) et Frick et al. (2006). Nous pouvons toutefois noter que $O_{kj} \geq 0$, l'indice est égal à zéro lorsqu'il y a une stratification parfaite entre le groupe k et le groupe j , autrement dit, l'indice est nul lorsque les deux distributions ne se chevauchent pas.

Plus généralement, plus la variabilité intragroupe est faible plus la stratification est élevée (moins le terme d'interaction est important). Une dernière remarque peut être faite sur l'indice de Gini intergroupe (G_b). Ce dernier, introduit par Yitzhaki and Lerman (1991), est défini comme le double de la covariance entre le revenu moyen de chaque groupe et son rang moyen dans la population globale rapportée au revenu moyen. Par conséquent, l'indice de Gini intergroupe G_b diffère de celui qui est traditionnellement retenu dans la littérature, cf. (Pyatt, 1976, entre autres), Mookherjee and Shorrocks (1982), Silber (1989). La différence réside dans le rang des revenus choisis. En effet, l'indice de Gini intergroupe G_{bp} est traditionnellement défini comme étant deux fois la covariance entre le revenu moyen de chaque groupe et le rang des moyennes de chaque groupe rapporté au revenu moyen, tel que $G_{bp} \geq G_p$. Lorsque les sous-groupes ne se chevauchent pas, alors $G_{bp} = G_p$. Par ailleurs, G_b peut être négatif dans certaines situations, par exemple, lorsqu'un groupe est composé d'individus avec de faibles revenus et qu'un autre groupe possède des revenus extrêmement élevés. C'est pour cette raison, qu'il n'est pas véritablement un indice de Gini. La décomposition peut être réécrite de la manière suivante :

$$G_u = \sum_{k=1}^K S_k G_k + \sum_{k=1}^K S_k G_k (O_k - 1) + G_{bp} + (G_b - G_{bp}). \quad (2.31)$$

Pour conclure, l'indice de Gini est à la fois sensible à la distribution des reve-

nus et à la distribution des rangs, c'est-à-dire à la stratification. C'est principalement pour cette raison qu'on dit que la décomposition de l'indice de Gini est plus complète que celle de la variance. Par conséquent, l'ANOGI semble plus appropriée que l'ANOVA lorsque la stratification a des implications sur l'interprétation des résultats. L'ANOGI peut donc être pertinente pour l'analyse des régressions.

2.2 La régression Gini

La méthode des moindres carrés ordinaires (MCO) qui consiste à minimiser la variance des erreurs est une technique fondamentale. Les estimateurs des MCO jouissent de bonnes propriétés tant à distance finie qu'asymptotiquement. Néanmoins, le recours à d'autres estimateurs n'est pas sans intérêt, dans la mesure où les estimateurs des MCO reposent sur des hypothèses assez fortes : linéarité, la stricte exogénéité des régresseurs, homoscedasticité, etc. Les estimateurs des MCO sont par ailleurs extrêmement sensibles aux points aberrants et aux erreurs de mesures. Aussi, une transformation monotone des données peut changer le signe des estimateurs des MCO. Une approche alternative qui surmonte ces difficultés est la régression Gini, introduite par Olkin and Yitzhaki (1992). Elle consiste à utiliser la GMD comme mesure de dispersion.

Comme nous le verrons dans cette section, la régression Gini peut être vue comme une méthode par variable instrumentale où le rang des variables représente les instruments. L'idée d'utiliser le rang comme instrument n'est pas nouveau, en réalité, elle a été proposée par Durbin (1954). Malheureusement, ce dernier n'avait pas remarqué que l'utilisation du rang comme instrument était équivalent à utiliser la GMD comme mesure de dispersion des erreurs. Il existe deux types de régressions Gini. La première est basée sur la minimisation de la GMD des erreurs. Dans ce cadre, lorsque les erreurs sont indépendantes et normalement distribuées alors l'estimateur converge vers la vraie valeur du paramètre. Cette première approche correspond donc à une méthode paramétrique car l'estimateur découle de la minimisation de la dispersion choisie. Comme la régression par quantile, l'estimateur obtenu par cette approche n'a pas d'expression explicite, il ne peut être obtenu que de manière numérique. La seconde approche consiste à construire l'estimateur en question de sorte qu'il soit moins sensible aux outliers. Cette seconde approche, puisqu'elle ne découle pas d'un

programme de minimisation, ne requiert pas de définir une forme fonctionnelle quelconque et est qualifiée de régression semi-paramétrique. Une combinaison de ces deux approches permet de tester l'hypothèse de linéarité du modèle.

2.2.1 Sensibilité de l'estimateur des MCO aux transformations monotones et aux points aberrants

La régression par les MCO couramment utilisée dans les travaux empiriques a ses limites. Comme le montre Yitzhaki (1990), une transformation monotone des données peut affecter le signe de l'estimateur. Cependant, la limite la plus grave est certainement le fait que l'estimateur des MCO est extrêmement sensible aux points aberrants.

Rappelons que l'objectif d'une régression est de trouver la pente de la droite de régression. Plusieurs procédés sont possibles pour déterminer cette pente. Minimiser la variance des erreurs en est une. Cette technique, dans le cas d'une régression simple d'une variable y sur une variable de même taille x , conduit à l'estimateur suivant :

$$\hat{\beta} = \frac{\text{Cov}(x, y)}{\text{Cov}(x, x)}. \quad (2.32)$$

S'il existe, alors $\text{Cov}(x, x) > 0$, et le signe de $\hat{\beta}$ dépend de $\text{Cov}(x, y)$. Par ailleurs, étant donné que la covariance est une moyenne pondérée d'éléments positifs et négatifs, une transformation monotone $T(\cdot)$ peut croître ou décroître certaines valeurs, ce qui peut affecter le signe de $\text{Cov}(T(y), x)$ et donc du nouvel estimateur $\hat{\beta}$ qui en découle. L'exemple suivant tiré de Yitzhaki (1990) illustre bien ce fait.

Exemple 2.2.1 Soit $x = (-4, -3, -2, -1, 0, 10)$,
 $y = (2.48, 0.73, -0.43, -1.44, -1.32, 0)$ et $T(y) = \ln(y + 1.45) \equiv z$, une transformation monotone de y . Régressons y sur x puis z sur x . Ceci conduit à :
 $y = -0.076x + 0.008$ avec $R^2 = 0.07$;
 $z = 0.003x - 0.68$ avec $R^2 = 0.00$.

Cet exemple suffit à montrer que les MCO sont sensibles aux transformations monotones des données. Ceci cache un problème plus général qui est que les estimateurs des MCO sont fortement sensibles aux formes fonctionnelles choisies. Formellement, une mauvaise spécification de la forme fonctionnelle d'un modèle peut avoir de lourdes conséquences comme : l'autocorrélation des erreurs,

des estimateurs imprécis, des variances biaisées et donc des tests d'hypothèses probablement erronés. Cependant, dans certains cas, utiliser les logarithmes des variables ou ajouter des termes quadratiques est suffisant pour détecter un bon nombre de relations non linéaires. Toutefois, il peut être difficile dans certaines situations d'identifier la raison pour laquelle une forme fonctionnelle est mal spécifiée. C'est pour cette raison qu'il paraît préférable d'utiliser une technique moins sensible à la spécification du modèle.

Parallèlement, la pente de la droite des MCO, $\hat{\beta}$ peut être vue comme une moyenne pondérée de tangentes. Pour cela, il faut l'exprimer sous une nouvelle forme :

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(\mathbf{x}, \mathbf{y})}{\text{Cov}(\mathbf{x}, \mathbf{x})} = \frac{\sum_{i=1}^n (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{x}_i - \bar{\mathbf{x}})}{\sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})^2} \\ &= \sum_{i=1}^n \underbrace{\frac{(\mathbf{x}_i - \bar{\mathbf{x}})^2}{\sum (\mathbf{x}_i - \bar{\mathbf{x}})^2}}_{w_i} \cdot \underbrace{\frac{(\mathbf{y}_i - \bar{\mathbf{y}})}{(\mathbf{x}_i - \bar{\mathbf{x}})}}_{b_{i0}}.\end{aligned}\quad (2.33)$$

Ainsi, nous pouvons remarquer que lorsqu'une observation s'écarte de la moyenne d'une valeur égale à c , alors son poids sera sur-pondéré de c^2/n . Ceci accroît l'influence des poids extrêmes. Nous pouvons aller plus loin dans la décomposition de la covariance. En réalité, cette pente peut être vue comme une autre moyenne pondérée. En effet, comme le montrent Olkin and Yitzhaki (1992) :

$$\begin{aligned}\text{Cov}(\mathbf{x}, \mathbf{y}) &= \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{y}_i - \bar{\mathbf{x}} \bar{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{y}_i - \frac{1}{n^2} \sum_{i=1}^n \mathbf{x}_i \sum_{i=1}^n \mathbf{y}_i \\ &= \frac{1}{n} \sum \mathbf{x}_i \mathbf{y}_i - \frac{1}{n^2} \sum \mathbf{x}_i \mathbf{y}_i - \frac{2}{n^2} \sum_{i>j} \mathbf{x}_i \mathbf{y}_j \\ &= \frac{(n-1)}{n^2} \sum \mathbf{x}_i \mathbf{y}_i - \frac{2}{n^2} \sum_{i>j} \mathbf{x}_i \mathbf{y}_j = \frac{1}{n^2} \left[\sum \mathbf{x}_i \mathbf{y}_i (n-1) \right] - \frac{2}{n^2} \sum_{i>j} \mathbf{x}_i \mathbf{y}_j \\ &= \frac{1}{n} \sum [\mathbf{x}_i \mathbf{y}_i + \mathbf{x}_j \mathbf{y}_j] - \frac{1}{n} \sum [\mathbf{x}_i \mathbf{y}_j + \mathbf{x}_j \mathbf{y}_i] = \frac{1}{n} \sum [\mathbf{x}_i \mathbf{y}_i + \mathbf{x}_j \mathbf{y}_j - \mathbf{x}_i \mathbf{y}_j - \mathbf{x}_j \mathbf{y}_i] \\ &= \frac{1}{n} \sum (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{y}_i - \mathbf{y}_j).\end{aligned}\quad (2.34)$$

Par analogie avec la précédente démonstration on a :

$$\text{Cov}(\mathbf{x}, \mathbf{x}) = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j). \quad (2.35)$$

D'où :

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{y}_i - \mathbf{y}_j)}{\sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)} \\ &= \sum_{i=1}^n \underbrace{(\mathbf{x}_i - \mathbf{x}_j)^2}_{w_{ij}} \cdot \underbrace{(\mathbf{y}_i - \mathbf{y}_j)}_{b_{ij}}. \end{aligned} \quad (2.36)$$

L'estimateur $\hat{\beta}$ apparaît ainsi comme une moyenne pondérée de toutes les tangentes possibles entre les paires de points $(\mathbf{x}_i, \mathbf{y}_j)$. Une méthode de régression non paramétrique peut donc consister à trouver une pente $\hat{\beta}$ sans recourir aux méthodes d'optimisation.

Par ailleurs, d'après cette nouvelle formule de $\hat{\beta}$, on remarque bien que plus les données sont hétérogènes, c'est à dire plus les écarts $\mathbf{x}_i - \mathbf{x}_j$ sont élevés plus les poids w_{ij} seront volatiles, du fait que les $\mathbf{x}_i - \mathbf{x}_j$ sont élevés au carré. Ceci accroît l'influence exercée par les points aberrants, conduisant à des variances trop élevées et donc à des estimateurs instables.

Il existe deux approches pour surmonter cette extrême sensibilité de l'estimateur des MCO aux observations isolées. La première consiste à utiliser une autre mesure de la dispersion des résidus moins sensibles aux valeurs aberrantes que la variance. Dans ce cadre lorsque les erreurs sont indépendantes et normalement distribuées alors l'estimateur converge vers la vraie valeur du paramètre. Cette première approche correspond donc à une méthode paramétrique car l'estimateur découle de la minimisation de ce nouvel indicateur de dispersion choisi. La seconde approche consiste à construire l'estimateur en question de sorte qu'il soit moins sensible aux outliers. Cette seconde approche puisqu'elle ne découle pas d'un programme de minimisation – et par conséquent ne requiert pas de définir une forme fonctionnelle – est qualifiée de régression semi-paramétrique.

2.2.2 L'approche paramétrique

L'approche paramétrique de la régression Gini consiste à minimiser le Gini des erreurs. Soit le modèle linéaire simple suivant :

$$\mathbf{y} = \alpha + \beta \mathbf{x} + \varepsilon. \quad (2.37)$$

Supposons que l'estimateur passe par le point moyen et que $\mathbb{E}(\varepsilon) = 0$. Ainsi, en notant \mathbf{e} le vecteur des résidus, et $\mathbf{R}(x) = n\mathbf{F}(x)$ le vecteur rang de \mathbf{x} , l'estimateur $\hat{\beta}^P$ qui minimise la GMD des résidus est l'estimateur de la régression paramétrique suivant :

$$\hat{\beta}^P = \arg \min \left\{ \sum_{i=1}^n \sum_{j=1}^n |\mathbf{e}_i - \mathbf{e}_j| \right\} = \arg \min \left\{ \frac{1}{n} \text{Cov}(\mathbf{e}, \mathbf{R}(e)) \right\}. \quad (2.38)$$

Notons que cet estimateur ne peut être obtenu que numériquement et correspond en réalité à la R -régression de Jureckova (1971), Jaekel (1972), McKean and Hettmansperger (1976). En effet, minimiser la GMD des erreurs est équivalent à minimiser⁴ $\sum_{i=1}^n \mathbf{e}_i \mathbf{R}_i(e)$, ce qui montre que la R -régression est une régression Gini.

Néanmoins, puisque la GMD des erreurs n'est pas une fonction continue l'estimateur $\hat{\beta}^P$ qui le minimise n'est pas forcément unique. Pour simplifier supposons que cette dérivée existe et est unique :

$$\frac{\partial GMD(\mathbf{e}, \mathbf{R}(e))}{\partial \beta^P} = \frac{\partial \{ \text{Cov}(\mathbf{y}, \mathbf{R}(e)) - \beta^P \text{Cov}(\mathbf{x}, \mathbf{R}(e)) \}}{\partial \beta^P} = -\text{Cov}(\mathbf{x}, \mathbf{R}(e)). \quad (2.39)$$

Au minimum la dérivée est nulle, donc $\text{Cov}(\mathbf{x}, \mathbf{R}(e)) = 0$, ce qui correspond à la condition d'orthogonalité de la régression Gini. Pour conclure sur cette approche, notons que :

$$\text{Cov}(\hat{\mathbf{y}}, \mathbf{R}(e)) = \text{Cov}(\hat{\alpha} + \hat{\beta} \mathbf{x}, \mathbf{R}(e)) = \hat{\beta} \text{Cov}(\mathbf{x}, \mathbf{R}(e)) = 0, \quad (2.40)$$

$$\text{Cov}(\mathbf{y}, \mathbf{R}(e)) = \text{Cov}(\alpha + \beta \mathbf{x} + \mathbf{e}, \mathbf{R}(e)) = \text{Cov}(\mathbf{e}, \mathbf{R}(e)). \quad (2.41)$$

Au même titre que la régression par quantile, l'estimateur obtenu par cette

4. En effet,

$$\text{Cov}(\mathbf{e}, \mathbf{R}(e)) = \frac{1}{n} \sum_{i=1}^n \mathbf{e}_i \mathbf{R}_i(e) - \frac{1}{n^2} \sum_{i=1}^n \mathbf{e}_i \sum_{i=1}^n \mathbf{R}_i(e) = \frac{1}{n} \sum_{i=1}^n \mathbf{e}_i \mathbf{R}_i(e)$$

approche n'a pas d'expression explicite et il ne peut être obtenu que de manière numérique. Cet estimateur est toutefois convergent et a l'avantage d'être optimal puisqu'il découle d'un programme de minimisation.

2.2.3 L'approche semi-paramétrique

Comme nous l'avons souligné précédemment, la pente de la droite des MCO peut être vue comme une moyenne pondérée des pentes issues de toutes les paires de points $(\mathbf{x}_i, \mathbf{y}_i)$. En effet,

$$\begin{aligned}\beta &= \frac{\text{Cov}(\mathbf{x}, \mathbf{y})}{\text{Cov}(\mathbf{x}, \mathbf{x})} = \frac{1/2\mathbb{E}[(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{y}_j)]}{1/2\mathbb{E}\{(\mathbf{x}_i - \mathbf{x}_j)^2\}} \\ &= \frac{1}{\mathbb{E}\{(\mathbf{x}_i - \mathbf{x}_j)^2\}} \mathbb{E}\left\{\left(\frac{\mathbf{y}_i - \mathbf{y}_j}{\mathbf{x}_i - \mathbf{x}_j}\right)(\mathbf{x}_i - \mathbf{x}_j)^2\right\} \\ &= \mathbb{E}\left\{\frac{(\mathbf{x}_i - \mathbf{x}_j)^2}{\mathbb{E}\{(\mathbf{x}_i - \mathbf{x}_j)^2\}} \cdot \frac{\mathbf{y}_i - \mathbf{y}_j}{\mathbf{x}_i - \mathbf{x}_j}\right\}\end{aligned}\quad (2.42)$$

D'où, sur un échantillon, l'estimateur s'écrit :

$$\hat{\beta} = \sum_{i < j} \underbrace{\frac{(\mathbf{x}_i - \mathbf{x}_j)^2}{\sum_{i < j} (\mathbf{x}_i - \mathbf{x}_j)^2}}_{w_{ij}} \cdot \underbrace{\frac{(\mathbf{y}_i - \mathbf{y}_j)}{(\mathbf{x}_i - \mathbf{x}_j)}}_{b_{ij}}.\quad (2.43)$$

D'après cette formule, plus les données sont hétérogènes, c'est-à-dire plus les écarts $\mathbf{x}_i - \mathbf{x}_j$ sont élevés, plus les w_{ij} sont volatiles, puisque les $\mathbf{x}_i - \mathbf{x}_j$ sont élevés au carré. Ceci accroît l'influence des points aberrants, conduisant à des variances trop élevées et donc à des estimateurs peu fiables ou instables. Pour surmonter cette limite, Olkin and Yitzhaki (1992) proposent de substituer les w_{ij} par les poids v_{ij} moins volatiles :

$$v_{ij} = \frac{(\mathbf{x}_i - \mathbf{x}_j)}{\sum_{i < j} (\mathbf{x}_i - \mathbf{x}_j)}.\quad (2.44)$$

Ainsi, l'estimateur semi-paramétrique de la régression Gini (simple) est donné par l'expression suivante :

$$\hat{\beta}^{sp} = \sum_{i < j} v_{ij} b_{ij} = \sum_{i < j} \underbrace{\frac{(\mathbf{x}_i - \mathbf{x}_j)}{\sum_{i < j} (\mathbf{x}_i - \mathbf{x}_j)}}_{v_{ij}} \cdot \underbrace{\frac{(\mathbf{y}_i - \mathbf{y}_j)}{(\mathbf{x}_i - \mathbf{x}_j)}}_{b_{ij}}.\quad (2.45)$$

En réalité, la substitution de w_{ij} par v_{ij} revient à remplacer dans la formule des MCO chaque covariance par la co-Gini appropriée. En effet, soit le modèle de régression suivant :

$$\mathbf{y} = \alpha + \beta\mathbf{x} + \varepsilon. \quad (2.46)$$

En utilisant les propriétés de la covariance :

$$\text{Cov}(\mathbf{y}, \mathbf{x}) = \text{Cov}(\alpha + \beta\mathbf{x} + \varepsilon, \mathbf{x}) = \beta\text{Cov}(\mathbf{x}, \mathbf{x}) + \text{Cov}(\varepsilon, \mathbf{x}). \quad (2.47)$$

En supposant $\text{Cov}(\varepsilon, \mathbf{x}) = 0$ (condition d'exogénéité des régresseurs ou condition d'orthogonalité des MCO), on obtient l'estimateur des MCO :

$$\hat{\beta}^{MCO} = \frac{\text{Cov}(\mathbf{y}, \mathbf{x})}{\text{Cov}(\mathbf{x}, \mathbf{x})}. \quad (2.48)$$

La régression Gini consiste à adopter la même démarche, en remplaçant chaque covariance par la co-Gini appropriée et en imposant la condition d'orthogonalité au sens de Gini, $\text{Cov}(\varepsilon, \mathbf{R}(\mathbf{x})) = 0$:

$$\hat{\beta}^{sp} = \frac{\text{Cov}(\mathbf{y}, \mathbf{R}(x))}{\text{Cov}(\mathbf{x}, \mathbf{R}(x))}. \quad (2.49)$$

Dans le cas d'une régression multiple, l'estimateur du Gini est donné par l'expression suivante :

$$\hat{\beta}^{sp} = (\mathbf{R}(\mathbf{X})\mathbf{X})^{-1}(\mathbf{R}(\mathbf{X})\mathbf{y}), \quad (2.50)$$

où \mathbf{X} est la matrice des régresseurs de taille $n \times (K + 1)$, en tenant compte de la constante, et où $\mathbf{R}(\mathbf{X})$ représente la matrice rang de \mathbf{X} (matrice qui fait apparaître les rangs des régresseurs en colonne) telle que $\mathbf{R}(\mathbf{X}) = n\mathbf{F}(\mathbf{X})$. La Régression Gini peut être interprétée comme une régression par variables instrumentales où $\mathbf{R}(\mathbf{X})$ correspond à la matrice des instruments \mathbf{Z} . En effet,

$$\hat{\beta}^{sp} = (\mathbf{Z}\mathbf{X})^{-1}(\mathbf{Z}\mathbf{y}). \quad (2.51)$$

Ceci permet de montrer que la régression Gini est plus robuste que les MCO lorsque les instruments $\mathbf{R}(\mathbf{X})$ sont appropriés. Rappelons que l'idée d'utiliser le rang comme instrument n'est pas nouvelle. Elle a été proposée par Durbin (1954). Ce dernier n'avait pas remarqué que ceci correspondait à la régression Gini.

L'approche est très proche de celle des MCO. En réalité, elle se fonde sur une imitation des MCO. C'est pour cette raison d'ailleurs que certains concepts utilisés dans le cadre des MCO peuvent être transposés au cadre de la régression Gini. Parmi ces concepts, nous pouvons citer le coefficient de détermination R^2 . L'équivalent de ce dernier dans la régression Gini est noté GR^2 :

$$GR^2 = 1 - \left(\frac{\text{Cov}(\mathbf{e}, \mathbf{R}(e))}{\text{Cov}(\mathbf{y}, \mathbf{R}(y))} \right). \quad (2.52)$$

Toutefois, puisque la décomposition du Gini est plus complexe que la décomposition de la variance, les propriétés du GR^2 diffèrent de ceux du R^2 . Le GR^2 obtenu dans l'approche par minimisation est toujours plus élevé que celui obtenu par l'approche semi-paramétrique. Plus généralement, les estimateurs obtenus dans ces deux approches diffèrent, mais ils sont équivalents lorsque le modèle est linéaire. Dans ce cas les GR^2 sont aussi identiques. Par conséquent, nous pouvons tester l'hypothèse de linéarité en comparant les estimateurs, ou les GR^2 , des deux types de régression Gini. Lorsque les deux estimateurs convergent vers la même limite alors nous pouvons accepter l'hypothèse de linéarité du modèle. Olkin and Yitzhaki (1992) ont proposé une formalisation de ce test de linéarité. Reconsidérons le modèle de régression simple suivant :

$$\mathbf{y} = \alpha + \beta \mathbf{x} + \varepsilon. \quad (2.53)$$

Soient β^p et β^{sp} les estimateurs de β obtenu respectivement par l'approche par minimisation et par l'approche semi-paramétrique. Soit $\varepsilon_{sp} = \mathbf{y} - \alpha - \beta^{sp} \mathbf{x}$. L'approche semi-paramétrique implique que $\text{Cov}(\varepsilon_{sp}, \mathbf{R}(\mathbf{x})) = 0$, $\beta^p = \beta^{sp} = \beta$ si et seulement si ε_{sp} vérifie également la condition d'orthogonalité de l'approche par minimisation $\text{Cov}(\mathbf{R}(\mathbf{x}), \varepsilon_{sp}) = 0$, si le modèle est linéaire.

L'idée de ce test est relativement simple : on estime le modèle par l'approche semi-paramétrique et on vérifie si les résidus vérifient la condition d'orthogonalité de l'approche par minimisation. Plus formellement, Olkin and Yitzhaki (1992) proposent la procédure suivante pour tester la linéarité du modèle de régression.

(i) Utiliser l'approche semi-paramétrique de la régression Gini pour estimer les paramètres afin de récupérer les résidus et d'avoir la condition d'orthogonalité.

(ii) Utiliser ces résidus pour tester l'hypothèse :

$$\left\| \begin{array}{l} H_0 : \text{Cov}(\mathbf{x}, \mathbf{F}_\varepsilon(\varepsilon_{sp})) = 0 \\ H_1 : \text{Cov}(\mathbf{x}, \mathbf{F}_\varepsilon(\varepsilon_{sp})) \neq 0. \end{array} \right. \quad (2.54)$$

Lorsque H_0 est acceptée, l'hypothèse de linéarité du modèle est acceptée. Comme nous allons le voir dans la section suivante, $\text{Cov}(\mathbf{x}, \mathbf{F}_\varepsilon(\varepsilon_{sp}))$ est une \mathbb{U} -statistique. Elle peut être estimée de manière convergente par l'estimateur \mathbb{U}_n :

$$\mathbb{U}_n = \frac{1}{n(n-1)} \sum_{i=1}^n \mathbf{x}_i \mathbf{R}_i(e). \quad (2.55)$$

Pour conclure sur cette procédure de test nous pouvons souligner qu'il est possible d'inverser cette démarche en commençant par estimer le modèle par l'approche paramétrique puis vérifier la condition d'orthogonalité de l'approche semi-paramétrique. Toutefois, nous suggérons la procédure ci-dessus pour les raisons suivantes.

1. Tout d'abord, l'approche semi-paramétrique ne requiert pas d'hypothèse de linéarité.
2. Contrairement à l'approche par minimisation, avec cette approche il n'y a pas de problème d'unicité des paramètres estimés.
3. Les estimateurs de cette approche sont faciles à mettre en œuvre, alors que les estimateurs de l'approche par minimisation découlent de lourds programmes de minimisation.
4. Enfin les estimateurs de cette approche peuvent être obtenus de manière numérique et peuvent être vus comme des estimateurs par variables instrumentales.

2.2.4 Similarités et différences entre la régression Gini et les MCO

Dans cette sous-section nous revenons sur les similarités et les différences entre les régressions Gini et MCO.

La première ressemblance est que toutes les deux sont basées sur une approximation linéaire de la courbe de régression. Elles peuvent être vues comme des moyennes pondérées de tangentes. La différence fondamentale tient au poids. Tandis que le Gini utilise la distance entre les points adjacents, les MCO utilise

cette distance au carré. Autrement dit, dans le cas de la méthode Gini les poids sont issus de la norme ℓ_1 alors que dans le cas des MCO ils sont issus de la norme ℓ_2 . Par conséquent, la régression Gini donne de meilleurs résultats en présence de points aberrants.

Parallèlement, la décomposition en sous-groupes du Gini est plus informative que celle de la variance. En effet, le Gini est à la fois sensible à la distribution et à la stratification. En utilisant les propriétés de la covariance, nous allons tenter de décomposer les coefficients de régression du Gini et des MCO en estimateurs intragroupes et intergroupes. Soient :

- (y, x) un couple de variables aléatoires représentatif de la distribution totale d'une population et (y_m, x_m) avec $m = 1, \dots, M$, M couples de variables représentant les M sous-groupes disjoints de la population totale.
- (y_{mj}, x_{mj}) $m = 1, \dots, M$ et $j = 1, \dots, n_m$ un couple de variables aléatoires représentatif du groupe m .
- $N = \sum_{m=1}^M n_m$
- $\hat{\beta}^{sp}$ l'estimateur sémi-paramétrique obtenu dans la population totale et β_m le coefficient de régression obtenu dans le groupe m ,
- y_{mj} la j ème observation de y_m , $y..$ la moyenne de y et $y_m.$ la moyenne de y_m
- $F_x(x)$ la fonction cumulative de x et $F..$ la moyenne des $F_x(x)$.
- $p_m = \frac{n_m}{N}$

Ainsi, l'estimateur de la régression Gini peut être réécrit (voir preuve en annexe 3.3.2) :

$$\begin{aligned} \hat{\beta}^{sp} &= \frac{1}{\text{Cov}(x, F(x))} \left\{ \sum_{m=1}^M p_m \text{Cov}_m(y, F(x)) + \text{Cov}_B(y_{m.}, F_{m.}(x)) \right\} \\ &= \frac{1}{\text{Cov}(x, F(x))} \left\{ \sum_{m=1}^M p_m \frac{\text{Cov}_m(y, F(x))}{\text{Cov}_m(x, F(x))} \text{Cov}_m(x, F_m(x)) \frac{\text{Cov}_m(x, F(x))}{\text{Cov}_m(x, F_m(x))} \right. \\ &\quad \left. + \text{Cov}_B(y_{m.}, F_{m.}(x)) \right\}. \end{aligned} \quad (2.56)$$

Soit :

$$\hat{\beta}_m^{sp} = \frac{\text{Cov}_m(y, F(x))}{\text{Cov}_m(x, F(x))} \quad (2.57)$$

l'estimateur sémi-paramétrique dans le groupe m . Et

$$O_m = \frac{\text{Cov}_m(\mathbf{x}, \mathbf{F}(\mathbf{x}))}{\text{Cov}_m(\mathbf{x}, \mathbf{F}_m(\mathbf{x}))} \quad (2.58)$$

l'indice de chevauchement entre le groupe m et la population totale.

On a :

$$\hat{\beta}^{sp} = \sum_{m=1}^M V_n \beta_m^{sp} + V_b \beta_b^{sp} \quad (2.59)$$

Avec $V_n = \frac{p_m O_m GMD^m(\mathbf{x})}{GMD(\mathbf{x})}$ la part du groupe m dans la distribution totale. La première partie 2.59 peut être interprétée comme l'intra et la seconde partie l'inter.

La décomposition de l'estimateur des MCO est :

$$\beta = \sum_{m=1}^M W_n \beta_m + W_b \beta_b. \quad (2.60)$$

La régression Gini fait intervenir un terme additionnel qui est le terme d'interaction. Les deux régressions sont identiques dans les termes intragroupes et intergroupes. La différence fondamentale tient au terme d'interaction. Lorsque les distributions ne se chevauchent pas la décomposition du coefficient de régression Gini est identique à celle de la variance. Par contre lorsque les distributions se chevauchent, la part de la composante intergroupe dans le Gini est totalement différente de celle de la variance. Comme le note Frick et al. (2006), plus le chevauchement entre les groupes est important moins la composante intergroupe est faible, ce qui n'est pas le cas des MCO.

Après avoir introduit les (éléments) fondamentaux de la régression Gini, nous allons à présent nous intéresser aux conditions requises pour assurer le "bon comportement" de la régression Gini.

2.2.5 Les conditions de "Grenander modifiées"

Dans la littérature sur la régression Gini aucune condition particulière n'a été proposée sur la matrice rang $\mathbf{R}(\mathbf{X})$ alors que nous savons, qu'à l'évidence, certaines conditions sont nécessaires. Ces dernières, que nous nommons conditions de "Grenander modifiées", ont été proposées par Ka and Mussard (2016). Nous montrons aussi comment les ex-aequo doivent être traités.

Soit un modèle de régression à plusieurs régresseurs, avec \mathbf{X} la matrice des variables exogènes de taille $n \times (K + 1)$, $\mathbf{R}(\mathbf{X})$ la matrice rang associée et \mathbf{x}_k est la k ème colonne de \mathbf{X} .

(G'1) Aucune variable ne doit dégénérer en une séquence de zéros :

$$\lim_{n \rightarrow +\infty} \mathbf{R}'(\mathbf{x}_k)\mathbf{x}_k \neq 0, \forall k \in \{1, \dots, K\}. \quad (2.61)$$

Autrement dit, $\mathbf{R}'(\mathbf{X})\mathbf{X}$ continue à augmenter avec la taille de l'échantillon.

(G2) Lorsque $n \rightarrow \infty$, les observations individuelles deviennent moins importantes. De plus, aucune valeur ne doit dominer $\mathbf{x}'_k\mathbf{x}_k$. Dans le cas de la régression Gini basée sur la norme ℓ_1 , les moments d'ordre 2 ne sont pas nécessaires. Ainsi, la condition Grenander classique **(G2)** n'est pas sine qua non.

(G'3) La matrice \mathbf{X} ne doit pas être comonotone. Deux vecteurs \mathbf{x} et \mathbf{y} sont comonotones si et seulement si $\mathbf{R}(\mathbf{x}) = \mathbf{R}(\mathbf{y})$. Lorsque deux régresseurs sont comonotones alors la matrice $\mathbf{R}'(\mathbf{X})\mathbf{X}$ n'est pas inversible. Notons que l'absence de comonotonie implique l'hypothèse de rang plein mais la réciproque n'est pas vraie. Soit \mathcal{M}^c l'ensemble des matrices comonotones avec au moins deux vecteurs comonotones, alors **(G'3)** devient :

$$\mathbf{X} \notin \mathcal{M}^c, \text{ pour tout } \mathbf{X} \in \mathbb{R}^{n \times (K+1)}.$$

Exemple 2.2.2 Prenons un exemple de matrice comonotone :

$$\mathbf{X} = \begin{pmatrix} 2 & 3 \\ 7 & 4 \\ 10 & 11 \end{pmatrix} \implies \mathbf{R}(\mathbf{X}) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{pmatrix}. \quad (2.62)$$

Par conséquent :

$$\mathbf{R}'(\mathbf{X})\mathbf{X} = \begin{pmatrix} 46 & 44 \\ 46 & 44 \end{pmatrix}. \quad (2.63)$$

Donc

$$\det(\mathbf{R}'(\mathbf{X})\mathbf{X}) = 0, \text{ et donc } \mathbf{R}'(\mathbf{X})\mathbf{X} \text{ est non inversible.} \quad (2.64)$$

2.2.6 Traitement des ex aequo

Une autre remarque importante peut être faite sur la manière de traiter les ex aequo. Dans la pratique trois possibilités ont été avancées pour traiter cette question.

(i) La première possibilité est de prendre le rang supérieur. Malheureusement cette procédure pose problème dans la mesure où elle conduit à une sur-pondération des poids et donc un biais vers le haut des estimateurs :

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 5 \\ 2 \\ 3 \end{pmatrix} \implies \mathbf{R}(\mathbf{x}) = \begin{pmatrix} 3 \\ 3 \\ 6 \\ 4 \\ 5 \end{pmatrix}. \quad (2.65)$$

(ii) La seconde possibilité est de prendre le rang inférieur. Cette option entraîne une sous-pondération des poids et donc des estimateurs biaisés vers le bas :

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 5 \\ 2 \\ 3 \end{pmatrix} \implies \mathbf{R}(\mathbf{x}) = \begin{pmatrix} 1 \\ 1 \\ 5 \\ 3 \\ 4 \end{pmatrix}. \quad (2.66)$$

(iii) La dernière option consiste à prendre le rang du point moyen. L'intérêt essentiel que présente cette dernière démarche est qu'elle conduit à des estimateurs non biaisés contrairement aux deux précédentes Yitzhaki and Schechtman (2013) :

$$\mathbf{x} = \begin{pmatrix} 2 \\ 2 \\ 5 \\ 3 \\ 4 \end{pmatrix} \implies \mathbf{R}(\mathbf{x}) = \begin{pmatrix} 1.5 \\ 1.5 \\ 5 \\ 3 \\ 4 \end{pmatrix}. \quad (2.67)$$

Autre exemple :

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} \implies \mathbf{R}(\mathbf{x}) = \begin{pmatrix} 1.5 \\ 1.5 \\ 3 \end{pmatrix} \quad (2.68)$$

Ce dernier cas correspond à ce qui se fait avec la régression Gini. Notons que le rang n'est qu'une approximation de la fonction cumulative. Cependant, seul le dernier cas permet de retrouver $F_{\mathbf{x}}(x) = \mathbf{R}(x)/n$. Rappelons tout de même que des pistes de recherche peuvent être effectuées sur les approches non paramétriques par noyau avec les estimateurs d'Epanechnikov notamment. Mais ceci n'est pas l'objet de la thèse.

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} 30 \\ 30 \\ 100 \\ 60 \\ 80 \end{pmatrix} = \mathbf{R}(\mathbf{x})/n. \quad (2.69)$$

2.2.7 La régression Gini-VI

Yitzhaki and Schechtman (2004) ont proposé la régression Gini avec variables instrumentales, nommée Gini-VI. Elle permet de prendre en compte les points atypiques mais aussi de corriger simultanément les erreurs de mesures ou problèmes d'endogénéité. Considérons le modèle linéaire suivant :

$$\mathbf{y} = \alpha + \beta\mathbf{x} + \varepsilon. \quad (2.70)$$

Supposons par ailleurs que la matrice colonne \mathbf{x} soit endogène et que la matrice colonne \mathbf{z} soit un instrument valide de ce modèle. De sorte que le modèle se réécrit comme suit :

$$\mathbf{y} = \alpha + \beta_1\mathbf{x} + \beta_2\mathbf{z} + \varepsilon. \quad (2.71)$$

Selon Yitzhaki and Schechtman (2004), l'estimateur Gini avec variables instrumentales est :

$$\hat{\beta}^{GIV} = \frac{\text{Cov}(\mathbf{y}, \mathbf{F}_{\mathbf{z}}(\mathbf{z}))}{\text{Cov}(\mathbf{x}, \mathbf{F}_{\mathbf{z}}(\mathbf{z}))}, \quad (2.72)$$

où $\mathbf{F}_{\mathbf{z}}(z)$ représente la fonction cumulative de \mathbf{z} , qui peut être approximée par la

matrice rang de \mathbf{z} . Par ailleurs, tout comme l'estimateur $\hat{\beta}^G$ de la régression Gini, l'estimateur de la régression Gini avec variables instrumentales $\hat{\beta}^{GIV}$ peut être exprimé comme une moyenne pondérée :

$$\hat{\beta}^{GIV} = \sum_{i=1}^n V_i b_i. \quad (2.73)$$

La seule différence concerne les poids V_i . Ces derniers s'écrivent dans le cas de la régression Gini-VI comme suit :

$$V_i = \frac{p\Delta\mathbf{x}_i}{n\text{Cov}(\mathbf{x}, \mathbf{R}(\mathbf{z}))} \left[\frac{n+1}{2} - \bar{\mathbf{R}}(\mathbf{z}) \right], \quad (2.74)$$

d'où :

$$\hat{\beta}^{GIV} = \frac{p\Delta\mathbf{x}_i}{n\text{Cov}(\mathbf{x}, \mathbf{R}(\mathbf{z}))} \left[\frac{n+1}{2} - \bar{\mathbf{R}}(\mathbf{z}) \right] \frac{(\mathbf{y}_{i+1} - \mathbf{y}_i)}{(\mathbf{x}_{i+1} - \mathbf{x}_i)}, \quad (2.75)$$

avec $p = i/n$, $\Delta\mathbf{x}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ et $\bar{\mathbf{R}}(\mathbf{z})$ représentant la moyenne des rangs.

Dans le cas de la régression multiple, l'estimateur Gini-VI est donné par la formule suivante :

$$\hat{\beta}^{GIV} = (\mathbf{R}'(\mathbf{Z})\mathbf{X})^{-1}\mathbf{R}'(\mathbf{Z})\mathbf{Y}, \quad (2.76)$$

où $\mathbf{R}(\mathbf{Z})$ représente la matrice rang des instruments.

L'estimateur Gini-VI est moins sensible aux observations extrêmes et à la violation de l'hypothèse de linéarité que les méthodes standards. Ceci découle directement des caractéristiques de la régression Gini. Cette propriété (non sensibilité à la violation de l'hypothèse de linéarité) est aussi très importante car d'après la littérature empirique sur les inégalités de revenu, les résultats des estimations dépendraient fortement de la forme fonctionnelle retenue. La régression Gini permet de contourner ce problème puisqu'elle ne requiert pas de définir une forme fonctionnelle précise. En outre, une transformation monotone des variables instrumentales \mathbf{z} n'affecte pas l'estimation, contrairement aux méthodes standards, où une transformation monotone de \mathbf{z} modifie considérablement les résultats. Par ailleurs, comme nous allons le voir dans la sous-section suivante, tous les estimateurs utilisés dans la régression Gini sont basés sur la théorie des \mathbb{U} -statistiques. Par conséquent, ils sont convergents (après normalisation appropriée).

2.2.8 Inférence et variance des estimateurs

L'objectif d'une régression ne se limite pas simplement à trouver l'estimateur des paramètres de la régression mais également de déterminer ses propriétés et de s'assurer notamment de sa convergence. D'après Yitzhaki and Schechtman (2013), tous les estimateurs développés dans le cadre de la régression Gini sont des \mathbb{U} -statistiques. La théorie des \mathbb{U} -statistiques a été introduite par Hoeffding (1948).⁵ L'un des principaux avantages de montrer que les paramètres découlant de la régression Gini sont des \mathbb{U} -statistiques, ou des fonctions de \mathbb{U} -statistiques, réside dans le fait que cette propriété garantit l'absence de biais et la convergence de l'estimateur et permet de calculer facilement sa variance. Nous ne développerons pas en détail la théorie des \mathbb{U} -statistiques ci-après, le lecteur intéressé par cette littérature pourra se référer à Denker (1985), (Serfling, 2009, Chapitre 5) et (Lehmann, 1999, Chapitre 6). Nous nous intéressons dans cette section aux propriétés des \mathbb{U} -statistiques et leurs liens avec la GMD.

Soit $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ une séquence de n variables aléatoires i.i.d. d'une famille P . Les estimateurs de cette famille peuvent s'écrire

$$\vartheta = \mathbb{E}[h(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)], \quad (2.77)$$

avec h une fonction symétrique, que l'on appelle le noyau. Dans ce cas, il est possible de montrer qu'un estimateur sans biais de ϑ est donné par l'expression suivante :

$$\mathbb{U}_n = \binom{n}{m}^{-1} \sum_{i_1, i_2, \dots, i_m} \dots \sum h(\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}). \quad (2.78)$$

La statistique \mathbb{U}_n est une \mathbb{U} -statistique de noyau h et de degré m . Par ailleurs, \mathbb{U}_n est le seul estimateur symétrique sans biais de ϑ et sa variance est la plus faible parmi les autres estimateurs de ϑ , cf. (Lehmann, 1999, Chapitre 6). Par ailleurs, il converge asymptotiquement vers la loi normale.

Sans perte de généralité, nous pouvons étendre cette statistique à l'ordre 2, voire à l'ordre K . Pour cette dernière voir (Lehmann, 1999, Chapitre 6). Soient $\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}$ et $\mathbf{y}_{i_1}, \mathbf{y}_{i_2} \dots \mathbf{y}_{i_m}$ des variables aléatoires indépendamment distri-

5. On pourra noter que \mathbb{U} désigne Unbiased (absence de biais).

buées. Considérons la statistique

$$\vartheta = \mathbb{E} [h(\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}, \mathbf{y}_{i_1}, \mathbf{y}_{i_2} \dots \mathbf{y}_{i_m})] \quad (2.79)$$

avec,

$$h(\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}, \mathbf{y}_{i_1}, \mathbf{y}_{i_2} \dots \mathbf{y}_{i_m}) \quad (2.80)$$

son noyau. Alors :

$$\mathbb{U}_{n_1, n_2} = \binom{n_1}{m_1}^{-1} \binom{n_2}{m_2}^{-1} \sum_{i_1, i_2, \dots, i_m} \dots \sum h(\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}, \mathbf{y}_{i_1}, \mathbf{y}_{i_2} \dots \mathbf{y}_{i_m}). \quad (2.81)$$

\mathbb{U}_{n_1, n_2} est une "two" \mathbb{U} -statistique de noyau h et de degré 2.

Exemple 2.2.3 Prenons quelques exemples simples de \mathbb{U} -statistiques.

(i) *L'espérance* : $\vartheta = \mathbb{E}(\mathbf{X})$, $m = 1$ et $h(\mathbf{x}) = \mathbf{x}$. La \mathbb{U} -statistique correspondante est la moyenne arithmétique :

$$\mathbb{U}_1 = \binom{n}{1}^{-1} \sum \mathbf{x}_i = \frac{1}{n} \sum \mathbf{x}_i. \quad (2.82)$$

(ii) *La Variance* : $\vartheta = \frac{1}{2} \mathbb{E}(\mathbf{x}_i - \mathbf{x}_j)^2$, $m = 2$ et $h(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^2$. La \mathbb{U} -statistique correspondante est :

$$\mathbb{U}_2 = \frac{2}{n(n-1)} \sum_{i < j} \frac{(\mathbf{x}_i - \mathbf{x}_j)^2}{2} = \frac{1}{(n-1)} \left(\sum_{i=1}^n \mathbf{x}_i^2 - n\mathbb{U}_1^2 \right). \quad (2.83)$$

(iii) *La GMD* : $\vartheta = \mathbb{E} |\mathbf{x}_i - \mathbf{x}_j|$ avec $h(\mathbf{x}_i, \mathbf{x}_j) = |\mathbf{x}_i - \mathbf{x}_j|$ et $m = 2$. La \mathbb{U} -statistique correspondante est :

$$\mathbb{U}_3 = \frac{2}{n(n-1)} \sum_{i < j} |\mathbf{x}_i - \mathbf{x}_j|. \quad (2.84)$$

A ce niveau, il faudra noter qu'il existe plusieurs autres définitions alternatives de \mathbb{U}_3 . En effet,

$$\mathbb{U}_3 = \frac{1}{4 \binom{n}{2}} \sum_{i=1}^n (2i-1-n) \mathbf{x}_{(i)} = \frac{1}{4 \binom{n}{2}} \sum_{i=1}^n i \mathbf{x}_{(i)} - \frac{n(n+1)}{2} \bar{\mathbf{x}}. \quad (2.85)$$

avec $\mathbf{x}_{(i)}$ la statistique d'ordre i de \mathbf{x} . Autrement dit, $\mathbf{x}_{(i)}$ est la i -ième plus petite valeur de \mathbf{x} cf. Yitzhaki and Olkin (1991). Pour plus d'exemples de \mathbb{U} -statistique voir annexe 3.3.1

La variance d'une \mathbb{U} -statistique, $\text{Var}(\mathbb{U})$, pour le paramètre ϑ est donnée par :

$$\text{Var}(\mathbb{U}) = \binom{n}{m}^{-1} \sum_{i=1}^m \binom{m}{i} \binom{n-m}{m-i} \xi_i, \quad (2.86)$$

où $\xi_i = \text{Var}[h_i(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)] = \mathbb{E}[h_i^2(x_1, x_2 \dots x_n)] - \theta_i^2$ et où θ_i est donné par ,

$$\theta_i = \mathbb{E}[h_i(x_1, x_2, \dots, x_n)]. \quad (2.87)$$

Ainsi,

$$\text{Var}(\mathbb{U}_{n_1, n_2}) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \frac{\binom{m_1}{i} \binom{n_1-m_1}{m_1-i}}{\binom{n_1}{m_1}^{-1}} \frac{\binom{m_2}{j} \binom{n_2-m_2}{m_2-j}}{\binom{n_2}{m_2}^{-1}} \xi_{ij}.$$

avec,

$$\xi_{ij} = \text{Var}[h(\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \dots \mathbf{x}_{i_m}, \mathbf{y}_{i_1}, \mathbf{y}_{i_2} \dots \mathbf{y}_{i_m})]. \quad (2.88)$$

L'expression de cette variance est assez complexe et son utilisation pratique est de ce fait très délicate. En effet, elle peut conduire à des formes complexes surtout lorsque $m > 1$. Une approche alternative pour obtenir la variance asymptotique d'une \mathbb{U} -statistique consiste à utiliser la méthode du Jackknife. En effet, Arvesen (1969) a montré que cette dernière permet d'obtenir une estimation convergente de la variance d'une \mathbb{U} -statistique :

$$\text{Var}(\mathbb{U}) = \frac{N-1}{N} \sum_{i=1}^N \left[\mathbb{U}_{-i} - \frac{1}{N} \sum_{i=1}^N \mathbb{U}_{-i} \right]^2,$$

avec \mathbb{U}_{-i} l'estimateur de \mathbb{U} basé sur l'échantillon de taille N sans la i ème observation.

Comme nous l'avons souligné précédemment, les \mathbb{U} -statistiques jouissent de bonnes propriétés à savoir l'absence de biais et la convergence asymptotique. Malheureusement, la principale difficulté est de montrer qu'une statistique est bien une \mathbb{U} -statistique et de déterminer son degré et son noyau. Heureusement, Lenth (1983) d'abord (pour l'ordre 1) puis Schechtman and Schechtman (2002) (généralisation à l'ordre m) ont proposé une manière très simple de vérifier si une statistique est une \mathbb{U} -statistique. En effet, selon eux \mathbb{U}_n en est une, si et seulement si, pour n suffisamment grand et \mathbb{U}_n une fonction symétrique, cette dernière est invariante par Jackknife. Intuitivement, puisque la technique du Jackknife permet de réduire le biais et qu'une \mathbb{U} -statistique est toujours sans biais alors, tout

U-statistique est invariante au Jackknife.

Après avoir introduit les généralités de la régression Gini, nous allons à présent proposer une extension de cette méthode à l'économétrie des données de Panel.

2.3 L'estimateur Gini en Données de Panel

Le recours grandissant à l'utilisation des données de panel est l'un des aspects marquants de l'évolution de l'économétrie au cours de ces dernières années. Elles sont de plus en plus utilisées en économie appliquée, en sociologie, en finance, etc. Comme le souligne Hsiao (2003), l'utilisation de données de panel présente plusieurs avantages. La double dimension des données de panel et leurs grands nombres permettent d'obtenir plus de variabilité, plus de degré de liberté, moins de colinéarité et permettent également de contrôler l'hétérogénéité des comportements. Mais surtout cette double dimension permet de tenir compte de l'influence de caractéristiques non observables des individus sur leur comportement, cf. Baltagi (2008). En outre, un avantage décisif des données de panel par rapport aux autres types de données tient au fait qu'elles sont souvent extrêmement nombreuses. En panel, il est très facile de disposer d'échantillons de plusieurs milliers d'observations. L'intérêt essentiel que présente la grande taille de ces échantillons et qu'ils conduisent à des estimations dont les propriétés peuvent être assimilées aux propriétés asymptotiques des méthodes utilisées. Ainsi, la mise en œuvre d'un estimateur convergent conduit à un estimateur très proche de la vraie valeur des paramètres. En effet, le nombre élevé des observations permet alors de supposer que le biais et la variance de l'estimateur tendent vers zéro. De fait, l'une des caractéristiques importantes des estimations basées sur un très grand nombre d'observations est leur grande précision.

Puisqu'un échantillon de données de panel contient un très grand nombre d'observations, on peut penser a priori que cet ensemble d'observations est capable de "neutraliser" l'influence de quelques points aberrants. En réalité, ce n'est pas vraiment le cas. Huber (1980) a montré que seulement 3% de valeurs atypiques dans un ensemble d'observations suffisent à perturber significativement la qualité des estimations. Plus généralement, tous ceux qui ont pratiqué l'économétrie sur données de panel (et comme nous le verrons d'ailleurs dans le chapitre

suivant) savent que quelques points atypiques peuvent suffire à modifier considérablement les résultats d'une estimation.

Par ailleurs, il faut signaler que, jusqu'à une période récente, il était d'usage d'éliminer toutes les observations relatives à un individu, dès lors que l'une d'elles apparaissait comme aberrante. Une telle pratique (qu'on appelle cylindrage) conduisait en effet souvent à éliminer une très forte proportion d'observations et donc, à perdre beaucoup d'informations ; ceci induisant, au mieux, une moindre précision des estimations et au pire, un risque de biais. C'est pourquoi d'ailleurs certains chercheurs ont préféré corriger ces points aberrants au lieu de les éliminer. Malheureusement, dans ce cas aussi il est indispensable que l'information disponible soit suffisante pour que la correction soit acceptable.

Pour toutes ces raisons évoquées, il serait tout-à-fait préférable d'utiliser une technique d'estimation moins sensible aux points aberrants comme la régression Gini. En effet, la correction des points aberrants quoi qu'intéressante conduit souvent à des erreurs de mesure dommageables à la qualité des estimations. En outre, le fait de développer la régression Gini en panel présente plusieurs autres avantages : la régression Gini est une méthode par variable instrumentale donc plus robuste que les méthodes standards : moins sensible aux erreurs et à la spécification du modèle. Ainsi, cette section se propose d'étendre la réflexion sur la régression Gini en panel, notamment, nous y développons l'estimateur par effet fixe en utilisant la GMD comme mesure de dispersion, *cf.* (Ka and Mussard (2016)). Nous proposons également le test de l'existence de spécificités individuelles.⁶

6. Koenker and Bassett Jr (1978) avait déjà proposé une technique d'estimation particulièrement adaptée lorsque les données sont contaminées par des points aberrants. Il s'agit de la régression par quantile. L'objectif est de minimiser une fonction de perte absolue différente en positive ou en négative :

$$\min \left\{ \sum q |y_i - x_i' \beta| + \sum (1 - q) |y_i - x_i' \beta| \right\}. \quad (2.89)$$

Cette méthode peut être vue comme un cas particulier de l'approche paramétrique de la régression Gini, *cf.* Olkin and Yitzhaki (1992).

2.3.1 L'estimateur Gini du modèle à effets fixes

Dans cette sous-section nous proposons une nouvelle technique de régression, la régression Gini en panel. Considérons le modèle à effets fixes suivant :

$$\mathbf{y}_{nt} = \beta_0 + \beta_n + \beta \mathbf{x}_{nt} + \varepsilon_{nt}, \quad (2.90)$$

avec n représentant la dimension individuelle et t la dimension temporelle, \mathbf{y}_{nt} est la variable expliquée, \mathbf{x}_{nt} la variable explicative, β_0 une constante, β_n les effets fixes individuels et ε_{nt} les perturbations aléatoires. Nous supposons que ces dernières sont i.i.d., d'espérance nulle et de variance constante.

L'estimateur des MCO du modèle précédent est donné par :

$$\hat{\beta}_{WMCO} = \left[\sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{n.})(\mathbf{x}_{nt} - \mathbf{x}_{n.})' \right]^{-1} \left[\sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{n.})(\mathbf{y}_{nt} - \mathbf{y}_{n.}) \right]. \quad (2.91)$$

Il s'agit de l'estimateur des MCO appliqué au modèle écrit en écarts par rapport aux moyennes individuelles :

$$\mathbf{y}_{nt} - \mathbf{y}_{n.} = \beta'(\mathbf{x}_{nt} - \mathbf{x}_{n.}) + \varepsilon_{nt} - \varepsilon_{n.}, \quad (2.92)$$

avec

$$\mathbf{y}_{n.} = \frac{1}{T} \mathbf{y}_{nt} \quad \mathbf{x}_{n.} = \frac{1}{T} \mathbf{x}_{nt} \quad \text{et} \quad \varepsilon_{n.} = \frac{1}{T} \varepsilon_{nt}. \quad (2.93)$$

L'estimateur par effet fixe couramment appelé l'estimateur "intra" résulte de la décomposition de la variabilité totale en variabilité intragroupe et en variabilité intergroupe. Une décomposition semblable peut être envisagée dans le cas du Gini.

Soit, $\mathbf{x}_{..}$ le vecteur ($K \times 1$) des moyennes des variables explicatives contenu dans \mathbf{X} soit $\mathbf{R}_x =: (\mathbf{R}'_{11}(\mathbf{X}), \dots, \mathbf{R}'_{nt}(\mathbf{X}), \dots, \mathbf{R}'_{NT}(\mathbf{X}))$ la matrice rang de \mathbf{X} avec $\mathbf{R}'_{nt}(\mathbf{X})$ la nt -ième ligne \mathbf{R}_x , qui est, un vecteur $1 \times K$. Soient $\mathbf{r}_{n.}(\mathbf{X})$ le vecteur rang des moyennes individuelles $K \times 1$, $\mathbf{r}_{..}(\mathbf{X})$ le vecteur des rangs moyens (de dimension $K \times 1$), $\mathbf{r}_{n.}(\mathbf{X}) = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{nt}(\mathbf{X})$, le vecteur rang des moyennes individuelles et $\mathbf{r}_{..}(\mathbf{X}) = \frac{1}{NT} \sum_{n=1}^N \sum_{t=1}^T \mathbf{R}_{nt}(\mathbf{X})$, le vecteur rang des moyennes totales.

Alors, la décomposition de la matrice des co-Gini, $\mathbf{R}'_x \mathbf{X}$, est :

$$\begin{aligned}
\mathbf{G}_{xx}^{total} &= \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{..}) (\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{..}(\mathbf{X}))' & (2.94) \\
&= \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} + \mathbf{x}_{n.} - \mathbf{x}_{n.} - \mathbf{x}_{..}) \mathbf{R}'_{nt}(\mathbf{X}) \\
&= \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{n.}) \mathbf{R}'_{nt}(\mathbf{X}) + \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{n.} - \mathbf{x}_{..}) \mathbf{R}'_{nt}(\mathbf{X}) \\
&= \underbrace{\sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{n.}) [\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{n.}(\mathbf{X})]'}_{\text{Variabilité intragroupe : } \mathbf{G}_{xx}^{intra}} + \underbrace{\sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{n.} - \mathbf{x}_{..}) [\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{..}(\mathbf{X})]'}_{\text{Variabilité intergroupe : } \mathbf{G}_{xx}^{inter}}.
\end{aligned}$$

De manière toute à fait équivalente nous pouvons décomposer \mathbf{R}'_y en variabilité intra-individuelle et inter-individuelle :

$$\begin{aligned}
\mathbf{G}_{xy}^{total} &= \sum_{n=1}^N \sum_{t=1}^T (y_{nt} - y_{n.}) [\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{n.}(\mathbf{X})] & (\mathbf{G}_{xy}^{intra}) \\
&+ \sum_{n=1}^N \sum_{t=1}^T (y_{n.} - y_{..}) [\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{..}(\mathbf{X})]. & (\mathbf{G}_{xy}^{inter})
\end{aligned}$$

Comme dans le cas des MCO, il est possible de vérifier que :

$$\mathbf{G}_{xx}^{total} = \mathbf{G}_{xx}^{intra} + \mathbf{G}_{xx}^{inter} \quad \text{et} \quad \mathbf{G}_{xy}^{total} = \mathbf{G}_{xy}^{intra} + \mathbf{G}_{xy}^{inter}. \quad (2.95)$$

La décomposition précédente conduit aux estimateurs intragroupe et intergroupe :

$$\hat{\beta}^{WG} = [\mathbf{G}_{xx}^{intra}]^{-1} [\mathbf{G}_{xy}^{intra}]. \quad (2.96)$$

Soit \mathbf{R}^c la matrice rang de taille $N \times K$ des données centrées

$$\mathbf{R}^c = ((\mathbf{R}_{11}(\mathbf{X}) - \mathbf{R}_{1.}(\mathbf{X}))', \dots, (\mathbf{R}_{nt}(\mathbf{X}) - \mathbf{R}_{n.}(\mathbf{X}))', \dots, (\mathbf{R}_{NT}(\mathbf{X}) - \mathbf{R}_{N.}(\mathbf{X}))'). \quad (2.97)$$

L'estimateur intragroupe peut être réécrit comme suit :

$$\hat{\beta}^{WG} = (\mathbf{R}^{c'} \mathbf{X}^c)^{-1} \mathbf{R}^{c'} \mathbf{y}^c. \quad (2.98)$$

Et l'estimateur inter est donné par la forme suivante :

$$\hat{\beta}^{BG} = [\mathbf{G}_{xx}^{inter}]^{-1} [\mathbf{G}_{xy}^{inter}]. \quad (2.99)$$

Introduisons les notations suivantes :

$$\mathbf{F}^{intra} = [\mathbf{G}_{xx}^{intra} + \mathbf{G}_{xx}^{inter}]^{-1} \mathbf{G}_{xx}^{intra} \quad (2.100)$$

$$\mathbf{F}^{inter} = [\mathbf{G}_{xx}^{intra} + \mathbf{G}_{xx}^{inter}]^{-1} \mathbf{G}_{xx}^{inter}. \quad (2.101)$$

L'estimateur Gini du modèle total peut être décomposé comme suit :

$$\begin{aligned} \hat{\beta}^G &= [\mathbf{G}_{xx}^{total}]^{-1} [\mathbf{G}_{xy}^{total}] = [\mathbf{G}_{xx}^{intra} + \mathbf{G}_{xx}^{inter}]^{-1} [\mathbf{G}_{xy}^{intra} + \mathbf{G}_{xy}^{inter}] \\ &= [\mathbf{G}_{xx}^{intra} + \mathbf{G}_{xx}^{inter}]^{-1} [\mathbf{G}_{xx}^{intra} \hat{\beta}^{WG} + \mathbf{G}_{xx}^{inter} \hat{\beta}^{BG}] \\ &= \mathbf{F}^{intra} \hat{\beta}^{WG} + \mathbf{F}^{inter} \hat{\beta}^{BG}. \end{aligned} \quad (2.102)$$

Comme nous l'a suggéré Yihzhaki dans une communication privée, une seconde famille d'estimateurs intragroupes et intergroupes peuvent être déterminée avec le Gini. En effet, puisque la co-Gini n'est pas symétrique, $(\text{Cog}(x, y) = \text{Cov}(y, \mathbf{F}(x)) \neq \text{Cov}(x, \mathbf{F}(y)) = \text{Cog}(y, x))$, nous pouvons décomposer aussi \mathbf{G}_{yx}^{total} et donc déduire les estimateurs intragroupes et intergroupes de la seconde approche de la régression Gini. Toutefois, pour simplifier la présentation de nos résultats, cette seconde approche ne sera pas présentée, nous pouvons néanmoins souligner que ces deux estimateurs convergent vers la même limite lorsque la taille de l'échantillon est importante.

2.3.2 Inférence sur l'estimateur intragroupe

Comme nous l'avons souligné précédemment, tous les estimateurs basés sur la régression Gini sont des \mathbb{U} -statistiques. Deux stratégies d'estimation de la variance sont donc possibles : l'estimation directe (la variance d'une \mathbb{U} -statistique) et l'estimation indirecte (en utilisant la méthode du Jackknife). La première (l'estimation directe), $\text{Var}(\mathbb{U})$, pour un paramètre ϕ^* de degré m est donnée par :

$$\text{Var}(\mathbb{U}) = \binom{N}{m}^{-1} \sum_{i=1}^m \binom{m}{i} \binom{N-m}{m-i} \xi_i.$$

L'estimation indirecte qui consiste à estimer la variance de la \mathbb{U} -statistique en utilisant la méthode du Jackknife est :

$$\text{Var}(\mathbb{U}) = \frac{N-1}{N} \sum_{i=1}^N \left[U_{-i} - \frac{1}{N} \sum_{i=1}^N U_{-i} \right]^2, \quad (2.103)$$

avec \mathbb{U}_{-i} l'estimateur basé sur l'échantillon de taille N privé de la i ème observation. L'estimation indirecte est en pratique assez souvent retenue dans la mesure où elle est plus simple d'utilisation surtout lorsque le degré de la \mathbb{U} -statistique est supérieur à 1.

Afin de montrer que $\hat{\beta}^{WG}$ est une fonction de \mathbb{U} -statistiques, considérons \mathbf{R}_k^c la k ème colonne de la matrice \mathbf{R}^c et \mathbf{x}_k^c la k ème colonne de la matrice \mathbf{X}^c , $k = 1, \dots, K$. Puisque l'estimateur intragroupe est $\hat{\beta}^{WG} = (\hat{\beta}_1^{WG}, \dots, \hat{\beta}_K^{WG})$ alors,

$$\mathbf{y}^c = \hat{\beta}_1^{WG} \mathbf{x}_1^c + \dots + \hat{\beta}_K^{WG} \mathbf{x}_K^c + \boldsymbol{\varepsilon}, \quad (2.104)$$

ainsi l'identité suivante est vérifiée :

$$\text{Cov}(\mathbf{y}^c, \mathbf{R}_1^c) = \hat{\beta}_1^{WG} \text{Cov}(\mathbf{x}_1^c, \mathbf{R}_1^c) + \dots + \hat{\beta}_K^{WG} \text{Cov}(\mathbf{x}_K^c, \mathbf{R}_1^c) + \text{Cov}(\boldsymbol{\varepsilon}, \mathbf{R}_1^c) \quad (2.105)$$

$$\text{Cov}(\mathbf{y}^c, \mathbf{R}_k^c) = \hat{\beta}_1^{WG} \text{Cov}(\mathbf{x}_1^c, \mathbf{R}_k^c) + \dots + \hat{\beta}_K^{WG} \text{Cov}(\mathbf{x}_K^c, \mathbf{R}_k^c) + \text{Cov}(\boldsymbol{\varepsilon}, \mathbf{R}_k^c) \quad (2.106)$$

$$\text{Cov}(\mathbf{y}^c, \mathbf{R}_K^c) = \hat{\beta}_1^{WG} \text{Cov}(\mathbf{x}_1^c, \mathbf{R}_K^c) + \dots + \hat{\beta}_K^{WG} \text{Cov}(\mathbf{x}_K^c, \mathbf{R}_K^c) + \text{Cov}(\boldsymbol{\varepsilon}, \mathbf{R}_K^c). \quad (2.107)$$

Soient $\hat{\beta}_{\varepsilon j} := \frac{\text{Cov}(\boldsymbol{\varepsilon}, \mathbf{R}_j^c)}{\text{Cov}(\mathbf{x}_j^c, \mathbf{R}_j^c)}$, $\hat{\beta}_{0j} := \frac{\text{Cov}(\mathbf{y}^c, \mathbf{R}_j^c)}{\text{Cov}(\mathbf{x}_j^c, \mathbf{R}_j^c)}$ et $\hat{\beta}_{kj} := \frac{\text{Cov}(\mathbf{x}_k^c, \mathbf{R}_j^c)}{\text{Cov}(\mathbf{x}_j^c, \mathbf{R}_j^c)}$, en divisant les trois équations précédentes respectivement par, $\text{Cov}(\mathbf{x}_1^c, \mathbf{R}_1^c)$, $\text{Cov}(\mathbf{x}_k^c, \mathbf{R}_k^c)$ et $\text{Cov}(\mathbf{x}_K^c, \mathbf{R}_K^c)$ alors :

$$\hat{\beta}_{01} = \hat{\beta}_1^{WG} + \dots + \hat{\beta}_K^{WG} \hat{\beta}_{K1} + \hat{\beta}_{\varepsilon 1} \quad (2.108)$$

$$\hat{\beta}_{0k} = \hat{\beta}_1^{WG} \hat{\beta}_{1k} + \dots + \hat{\beta}_K^{WG} \hat{\beta}_{Kk} + \hat{\beta}_{\varepsilon k} \quad (2.109)$$

$$\hat{\beta}_{0K} = \hat{\beta}_1^{WG} \hat{\beta}_{1K} + \dots + \hat{\beta}_K^{WG} + \hat{\beta}_{\varepsilon K}. \quad (2.110)$$

Soit le vecteur colonne suivant $\hat{\mathbf{b}}_0 := (\hat{\beta}_{01}, \dots, \hat{\beta}_{0K})$ and $\hat{\mathbf{b}}_\varepsilon := (\hat{\beta}_{\varepsilon 1}, \dots, \hat{\beta}_{\varepsilon K})$, alors :

$$\begin{pmatrix} \hat{\beta}_1^{WG} \\ \vdots \\ \hat{\beta}_K^{WG} \end{pmatrix} = \begin{pmatrix} 1 & \hat{\beta}_{21} & \dots & \hat{\beta}_{K1} \\ \vdots & \vdots & \dots & \vdots \\ \hat{\beta}_{1K} & \hat{\beta}_{2K} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_{01} - \hat{\beta}_{\varepsilon 1} \\ \vdots \\ \hat{\beta}_{0K} - \hat{\beta}_{\varepsilon K} \end{pmatrix} =: \hat{\mathbf{B}}^{-1} [\hat{\mathbf{b}}_0 - \hat{\mathbf{b}}_\varepsilon]. \quad (2.111)$$

L'estimateur intra-individuel est une fonction de coefficients de la régression Gini semi-paramétrique. D'après (Yitzhaki and Schechtman, 2013, chapitre 9), $\hat{\beta}_{0k}$, $\hat{\beta}_{\varepsilon k}$ et $\hat{\beta}_{kh}$ sont des \mathbb{U} -statistiques. En outre, puisque $\hat{\mathbf{B}}$ est une matrice de rang plein, alors $\hat{\beta}^{WG}$ est une fonction de \mathbb{U} -statistiques. En utilisant le théorème de Slutsky, $\hat{\beta}^{WG}$ est un estimateur convergent de β^{WG} , d'où :

$$\hat{\beta}^{WG} \underset{a}{\sim} \mathcal{N}(\beta, \text{Var}(\mathbb{U})). \quad (2.112)$$

2.3.3 Simulations

Dans l'objectif d'illustrer l'avantage de nos estimateurs intragroupe et inter-groupe nous proposons une simulation dont l'algorithme est décrit ci-dessous. Elle consiste à estimer le modèle (2.90) et comparer à chaque fois les MSE.

Simulation de référence

- Boucle de $b = 1, \dots, B = 10,000$;
 - ↪ les régresseurs sont générés selon une loi normale multivariée $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$ tel que $\mu = (0, 10, 4)$, $\Sigma = \begin{pmatrix} 1 & 0.5 & 0.2 \\ & 1 & 0.15 \\ & & 1 \end{pmatrix}$, $\varepsilon \sim \mathcal{N}(0, 1)$, $\beta = (0.7, 1.23, 0.13)$ et l'effet fixe $\beta_n \sim \mathcal{N}(0, 5)$, $\beta_0 = 0.75$;
 - ↪ la variable dépendante y est déduite de (2.90) ;
 - ↪ y est régressé sur \mathbf{x} [MCO (2.92) et Gini (2.98)] ;
 - ↪ $[MSE_j = \frac{1}{B} \sum_{j=1}^{NT} (\hat{\beta}_j - \beta)^2]$;
 - ↪ Des points aberrants sont introduits (\mathbf{x}^o), puis y est régressé sur \mathbf{x}^o [MCO (2.92) et Gini (2.98)] : les estimateurs $\hat{\beta}_b^{MCO}$ et $\hat{\beta}_b^G$ sont déduits pour tout b avec $b = 1, \dots, B$;
- Fin b ;
- Les moyennes des estimateurs du Gini et des MCO ainsi que les MSE sont calculés. La moyenne des estimateurs Gini et MCO sont respectivement $\bar{\hat{\beta}}^{WG}$ et

$\bar{\hat{\beta}}^{WMCO}$ ($\bar{\hat{\beta}}^{BG}$ et $\bar{\hat{\beta}}^{BMCO}$ sont les moyennes des estimateurs inter enfin $\bar{\hat{\beta}}^G$ et $\bar{\hat{\beta}}^{MCO}$ la moyenne des estimateurs globaux).

Comme nous pouvons le voir, les estimations globales (table2.1) et intergroupes (table2.2) sont biaisées. Ce résultat est très connu dans la mesure où la variabilité inter-individuelle représente l'essentiel de la variabilité totale en panel. Ainsi, même si la régression Gini semble être meilleure, l'estimation reste de très mauvaise qualité et peu fiable. Ce qui nous paraît intéressant, ce sont les estimations avec contamination. Ces dernières soulignent que la régression Gini est plus robuste à l'influence des points aberrants.

TABLE 2.1 – Estimateur Gini Global $\hat{\beta}^G$

| Estimates → $\beta =$ | Sans points extrêmes | | Sans points extrêmes | |
|--------------------------|---------------------------------|-----------------------------|---------------------------------|-----------------------------|
| | $\bar{\hat{\beta}}^{MCO}$ (MSE) | $\bar{\hat{\beta}}^G$ (MSE) | $\bar{\hat{\beta}}^{MCO}$ (MSE) | $\bar{\hat{\beta}}^G$ (MSE) |
| 0.75 | 5.0423 (226.478) | 4.78504 (256.621) | 6.856 (443.18) | 5.0832 (301.12) |
| 0.7 | 0.773 (5.116) | 0.7537 (5.559) | 1.1022 (10.147) | 0.8998 (6.961) |
| 1.23 | 1.273 (0.886) | 1.259 (1.047) | 1.3430 (1.945) | 1.2896 (1.094) |
| 0.13 | 0.1221 (0.211) | 0.127 (0.253) | 0.1267 (0.1909) | 0.1195 (0.839) |

TABLE 2.2 – Estimateur Gini intergroupe $\hat{\beta}^{BG}$

| Estimates → $\beta =$ | Sans points extrêmes | | Avec points extrêmes | |
|--------------------------|----------------------------------|--------------------------------|----------------------------------|--------------------------------|
| | $\bar{\hat{\beta}}^{BMCO}$ (MSE) | $\bar{\hat{\beta}}^{BG}$ (MSE) | $\bar{\hat{\beta}}^{BMCO}$ (MSE) | $\bar{\hat{\beta}}^{BG}$ (MSE) |
| 0.7 | 0.7873 (0.03926) | 0.7677 (0.03905) | 0.6390 (0.0523) | 0.6891(0.0301) |
| 1.23 | 1.2534 (0.0797) | 1.25470 (0.0784) | 1.0315 (0.082) | 1.1929 (0.049) |
| 0.13 | 0.14570 (0.030) | 0.1455 (0.033) | 0.09517 (0.056) | 0.1195(0.021) |

A présent nous nous intéressons aux estimateurs non biaisés. Comme nous le savons l'estimateur intragroupe est sans biais. Nous pouvons remarquer une forte similarité entre les estimations obtenues par l'application des MCO et du Gini en l'absence de points extrêmes. Ce résultat est usuel puisque lorsque la distribution des régresseurs est normale multivariée, la régression Gini est équivalente à celle des MCO. Par contre, les estimations avec contamination sont notablement différentes. Une dernière remarque peut être faite quant à la procédure

de contamination. Dans la table 2.3 ci-dessous, nous avons contaminé $p = 1\%$ des observations au hasard en remplaçant ces p observations par 2 fois le maximum des régresseurs. Même si cette procédure conduit à des résultats biaisés vers le haut, le Gini est plus robuste (voir les MSE).

TABLE 2.3 – Estimateur Gini intragroupe $\hat{\beta}^{WG}$

| Estimateurs → $\beta =$ | Sans points extrêmes | | Avec points extrêmes | |
|----------------------------|----------------------------|--------------------------|----------------------------|--------------------------|
| | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) |
| 0.7 | 0.70022 (0.000388) | 0.70013 (0.000432) | 0.95150 (0.01724) | 0.7514 (0.001012) |
| 1.23 | 1.22991 (0.000062) | 1.22989 (0.000067) | 1.31870 (0.26722) | 1.26013 (0.00253) |
| 0.13 | 0.13018 (0.000119) | 0.13023 (0.000124) | 0.22851 (0.17871) | 0.140152(0.001948) |

TABLE 2.4 – Estimateur Gini intragroupe : outliers positifs

| Estimateurs → $\beta =$ | Sans points extrêmes | | Avec points extrêmes | |
|----------------------------|----------------------------|--------------------------|----------------------------|--------------------------|
| | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) |
| 0.7 | 0.70011 (0.000263) | 0.70012 (0.000279) | 0.9702 (0.02014) | 0.78135 (0.00100) |
| 1.23 | 1.22993 (0.000442) | 1.22989 (0.000475) | 1.3243 (0.10242) | 1.2679 (0.00351) |
| 0.13 | 0.13017 (0.000864) | 0.13021 (0.000909) | 0.2391 (0.19280) | 0.1471(0.00315) |

Dans la table 2.5 ci-dessous une valeur aberrante négative de 50 est ajoutée à 1% de l'échantillon tiré au hasard. Il est évident que toutes les estimations intragroupes sont biaisées vers le bas. Il est également important de noter que les MSE des MCO sont très élevés par rapport à ceux de Gini.

TABLE 2.5 – Estimateur Gini intragroupe : outliers négatifs

| Estimateurs → $\beta =$ | Sans points extrêmes | | Avec points extrêmes | |
|----------------------------|----------------------------|--------------------------|----------------------------|--------------------------|
| | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) | $\hat{\beta}^{WMCO}$ (MSE) | $\hat{\beta}^{WG}$ (MSE) |
| 0.7 | 0.70011 (0.000263) | 0.70012 (0.000279) | 0.5189 (0.45601) | 0.6598 (0.00287) |
| 1.23 | 1.22993 (0.000442) | 1.22989 (0.000475) | 0.9978 (0.5223) | 1.1589 (0.00843) |
| 0.13 | 0.13017 (0.000864) | 0.13021 (0.000909) | 0.0851 (0.287650) | 0.10991(0.00498) |

Enfin, nous nous intéressons aux distributions non gaussienne (table 2.7). ,20) ;

- $x_2 \sim \text{Weibull}(10, 15)$;
- $x_2 \sim \text{Exponentielle}(15)$;
- l'effet individuelle $\beta_n \sim U[5, 20]$.

TABLE 2.6 – **Distribution non normale : sans outlier**

| Estimates \rightarrow | $\hat{\beta}^{WG}$ | $\hat{\beta}^{WOLS}$ |
|-------------------------|--------------------|----------------------|
| 0.7 | 0.7032 (0.0099) | 0.8501 (1.07564) |
| 1.23 | 1.2312 (0.0081) | 1.29012 (2.74206) |
| 0.13 | 0.1308 (0.0531) | 0.199 (4.4869) |

Ces résultats constituent une bonne illustration de l'avantage de la régression Gini. Toutefois, il ne faudrait surtout pas réduire la régression Gini à une simple méthode de traitement des points aberrants. En effet, comme nous l'avons vu précédemment cette technique de régression jouit de plusieurs autres avantages : estimateurs sans biais, convergents, moins sensibles à la spécification du modèle et aux erreurs de mesures.

En réalité, l'insensibilité des estimations de la régression Gini aux points aberrants n'est qu'une propriété de la régression Gini. Et cette propriété découle du fait que la GMD est plus adaptée lorsque l'hypothèse de normalité n'est pas remplie et plus généralement lorsque le processus générateur des données a une queue épaisse (Carcea and Serfling, 2015). Par ailleurs, la régression Gini est une méthode par variable instrumentale⁷.

2.3.4 Test de l'existence de spécificités individuelles

L'utilisation du modèle à effets fixes et plus généralement de la structure de panel suppose l'existence d'une hétérogénéité inter-individuelle des comportements. Cette dernière est prise en compte par l'adjonction d'effets fixes individuels au modèle. Les estimateurs en données de panel sont généralement biaisés lorsque l'hypothèse d'homogénéité des comportements est rejetée. Nous proposons donc une procédure de test afin de prouver l'existence d'une telle hypothèse

7. more precise variable is obtained by ranking the x 's in order, and putting z equal to the rank order. If the ranking is relatively unaffected by errors the rank should be a good instrumental variable. Even if the errors are large, the bias should be a good deal smaller than for the original observations Durbin (1954)

dans le cadre des régressions sur indice de Gini. Ceci revient à discriminer entre le modèle,

$$\mathbf{y}_{nt} = \beta_0 + \sum \beta_k \mathbf{x}_{nt}^k + \varepsilon_{nt}, \quad (2.113)$$

et le modèle avec effets individuels,

$$\mathbf{y}_{nt} = \alpha_n + \sum \beta_k \mathbf{x}_{nt}^k + \varepsilon_{nt}. \quad (2.114)$$

Il s'agit d'un test joint :

$$\left\| \begin{array}{l} H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_n (= \beta_0) \\ H_1 : \exists i, j \in \{1, \dots, n\} \text{ tel que } \alpha_i \neq \alpha_j. \end{array} \right. \quad (2.115)$$

Dans le cas des MCO, cette hypothèse H_0 est testée en utilisant un test de Fisher. Ce dernier conduit à comparer la statistique F ci-dessous au fractile de la loi de Fischer à $(N - 1, N(T - 1) - K)$ degrés de liberté :

$$F = \frac{SCR_o - SCR}{SCR} \cdot \frac{N(T - 1) - k}{N - 1} \sim \mathcal{F}(N - 1, N(T - 1) - K), \quad (2.116)$$

où SCR_o représente la somme des carrés résiduels associée à l'estimateur du modèle contraint (2.113), sous l'hypothèse H_0 , et SCR celle associée au modèle à effets individuels (2.114). L'une des limites de ce test est que les statistiques de tests sont basées sur la norme ℓ_2 . Le test est donc sensible aux points aberrants et à la transformation monotone des données. Ici nous proposons un test alternatif basé sur la GMD.

Formellement, deux modèles sont identiques lorsque les variabilités expliquées sont égales. Ainsi, tester H_0 est équivalent à tester l'égalité de la variabilité expliquée dans deux modèles. La variabilité totale peut être décomposée, pour chaque modèle, en variabilité expliquée et résiduelle en suivant l'analyse du Gini (ANOGI) :

$$\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}_c(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})) + \text{Cov}_c(\mathbf{e}, \mathbf{F}_y(\mathbf{y})), \quad (2.117)$$

$$\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})) + \text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y})), \quad (2.118)$$

où Cov_c représente la covariance issue du modèle avec contrainte sur les para-

mètres, *i.e.*, le modèle (2.113). Tester l'égalité de la variabilité expliquée entre les deux modèles revient à tester :

$$\begin{cases} H'_0 : \text{Cov}_c(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})) \\ H'_1 : \text{Cov}_c(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})) \neq \text{Cov}(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{y})). \end{cases} \quad (2.119)$$

L'égalité H'_0 implique :

$$\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) - \text{Cov}_c(\mathbf{e}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) - \text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y})). \quad (2.120)$$

Puisque sous H_0 , $\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))$, on a :

$$\text{Cov}_c(\mathbf{e}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y})). \quad (2.121)$$

Notons que \mathbf{e} et \mathbf{y} peuvent être permutés, voir Olkin and Yitzhaki (1992), alors pour le modèle (2.114) :

$$\frac{\text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y}))}{\text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{e}))} = \frac{\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{e}))}{\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))}, \quad (2.122)$$

d'où

$$\text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y})) = \frac{\text{Cov}(\hat{\mathbf{y}}, \mathbf{F}_y(\mathbf{e})) + \text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{e}))}{\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))} \text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{e})). \quad (2.123)$$

Puisque $\text{Cov}(\hat{\mathbf{y}}, \mathbf{R}(\mathbf{e})) = 0^8$ alors la précédente équation peut être réécrite comme suit :

$$\text{Cov}(\mathbf{e}, \mathbf{F}_y(\mathbf{y})) = \frac{\{\text{Cov}(\mathbf{e}, \mathbf{F}(\mathbf{e}))\}^2}{\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))}. \quad (2.124)$$

De même pour le modèle contraint (2.113) :

$$\text{Cov}_c(\mathbf{e}, \mathbf{F}_y(\mathbf{y})) = \frac{\{\text{Cov}_c(\mathbf{e}, \mathbf{F}(\mathbf{e}))\}^2}{\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))}. \quad (2.125)$$

Puisque sous H_0 , $\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))$ alors tester $\text{Cov}(\varepsilon, \mathbf{F}_y(\mathbf{y})) = \text{Cov}_c(\varepsilon, \mathbf{F}_y(\mathbf{y}))$, *i.e.* $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_i = \beta_0$, est tout à fait équivalent à tester $\text{Cov}_c(\varepsilon, \mathbf{F}(\varepsilon)) = \text{Cov}(\varepsilon, \mathbf{F}(\varepsilon))$. Ainsi nous pouvons construire un test basé

8. $\text{Cov}(\hat{\mathbf{y}}, \mathbf{R}(\mathbf{e})) = \text{Cov}(\hat{\mathbf{a}} + \hat{\beta}\mathbf{x}, \mathbf{R}(\mathbf{e})) = \hat{\beta}\text{Cov}(\mathbf{x}, \mathbf{R}(\mathbf{e})) = 0$.

sur la différence entre les GMD des deux modèles :

$$\left\| \begin{array}{l} H_0'' : \text{Cov}_c(\varepsilon, F(\varepsilon)) - \text{Cov}(\varepsilon, F(\varepsilon)) = 0 \\ H_1'' : \text{Cov}_c(\varepsilon, F(\varepsilon)) - \text{Cov}(\varepsilon, F(\varepsilon)) \neq 0. \end{array} \right. \quad (2.126)$$

Notons que $\mathbb{U}_1 = \text{Cov}_c(\mathbf{e}, R(\mathbf{e}))$ et $\mathbb{U}_2 = \text{Cov}(\mathbf{e}, R(\mathbf{e}))$ sont deux \mathbb{U} -statistiques. Il s'agit d'estimateurs sans biais de respectivement, $\text{Cov}_c(\varepsilon, \mathbf{F}(\varepsilon))$ et $\text{Cov}(\varepsilon, \mathbf{F}(\varepsilon))$, dont la forme est donnée pour chacun des modèles par :

$$\mathbb{U} = \frac{1}{n(n-1)} \sum_{i=1}^n \mathbf{e}_i R_i(\mathbf{e}). \quad (2.127)$$

Par conséquent nous pouvons estimer la variance de leur différence, $\mathbb{U}_1 - \mathbb{U}_2 = \hat{S}$, en utilisant la méthode du Jackknife. La statistique de test est alors donnée par l'expression suivante :

$$T = \frac{\hat{S}}{\sqrt{\hat{V}(\hat{S})}}. \quad (2.128)$$

Comme dans le cas du test de Fischer nous pouvons montrer que tester l'égalité des GMD des erreurs est équivalent à comparer les coefficients de détermination de Gini ($G\rho^2$) des deux modèles ($G\rho^2$ et $G_c\rho^2$). Le GR^2 est un estimateur de $G\rho^2$. Nous pouvons montrer qu'il est une fonction de \mathbb{U} -statistiques, par conséquent il s'agit d'une \mathbb{U} -statistique. Notons que $\text{Cov}(\mathbf{e}, \mathbf{R}(\mathbf{y})) - \text{Cov}_c(\mathbf{e}, \mathbf{R}(\mathbf{y})) = 0$ implique :

$$\hat{Z} = \left(\frac{\text{Cov}(\mathbf{e}, \mathbf{R}(\mathbf{e}))}{\text{Cov}(\mathbf{y}, \mathbf{R}(\mathbf{y}))} \right)^2 - \left(\frac{\text{Cov}_c(\mathbf{e}, \mathbf{R}(\mathbf{e}))}{\text{Cov}_c(\mathbf{y}, \mathbf{R}(\mathbf{y}))} \right)^2 = 0. \quad (2.129)$$

Aussi la statistique précédente issue de la population est :

$$Z = \left\{ 1 - \left(\frac{\text{Cov}_c(\varepsilon, \mathbf{F}(\varepsilon))}{\text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))} \right)^2 \right\} - \left\{ 1 - \left(\frac{\text{Cov}(\varepsilon, \mathbf{F}(\varepsilon))}{\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))} \right)^2 \right\} \quad (2.130)$$

$$= G_c\rho^2 - G\rho^2. \quad (2.131)$$

Ainsi, une troisième possibilité de test est donnée par :

$$\left\| \begin{array}{l} H_0''' : G_c\rho^2 - G\rho^2 = 0 \\ H_1''' : G_c\rho^2 - G\rho^2 \neq 0. \end{array} \right. \quad (2.132)$$

Pour réaliser le test, il faut un estimateur convergent de Z . Puisque $\text{Cov}(\mathbf{y}, \mathbf{F}_y(\mathbf{y})) = \text{Cov}_c(\mathbf{y}, \mathbf{F}_y(\mathbf{y}))$, ce dernier se réécrit :

$$\hat{Z} = \frac{\mathbb{U}_2^2 - \mathbb{U}_1^2}{\mathbb{U}_3^2}, \quad (2.133)$$

avec $\mathbb{U}_3 = \binom{n}{2}^{-1} \sum_{i < j} \sum |\mathbf{y}_i - \mathbf{y}_j|$. Donc la statistique de test permettant de tester H_0''' est donnée par :

$$T' = \frac{\hat{Z}}{\sqrt{\hat{V}(\hat{Z})}}, \quad (2.134)$$

avec $\hat{V}(\hat{Z})$ la variance de la \mathbb{U} -statistique \hat{Z} estimée par Jackknife. En effet T' est une fonction de U -statistiques, il s'agit donc d'une \mathbb{U} -statistique convergente. Ainsi, pour un échantillon suffisamment large, la normalité asymptotique est obtenue :

$$T' = \frac{\hat{Z}}{\sqrt{\hat{V}(\hat{Z})}} \stackrel{a}{\sim} \mathcal{N}(0, 1). \quad (2.135)$$

2.4 Les perturbations non sphériques

Depuis les travaux fondateurs de Olkin and Yitzhaki (1992), l'utilisation de la méthodologie Gini à été employée dans de nombreux développements théoriques et économétriques. Comme nous l'avons vu précédemment cette réflexion sur la régression Gini a été étendue aux variables instrumentales par Yitzhaki and Schechtman (2004), aux séries temporelles par Carcea and Serfling (2015) et par Shelef and Schechtman (2011) et plus récemment au test de racine unitaire par Shelef (2014). D'après cette dernière, le test de Dickey et Fuller fournit de faibles puissances lorsque les données sont contaminées par des valeurs aberrantes. Sur le plan empirique, ce test semble être à la fois robuste et valide pour une large gamme de distributions. Toutefois, il semble que sa puissance soit faible lorsque l'hypothèse d'absence d'autocorrélation des erreurs n'est plus satisfaite. Dans cette section, nous proposons l'estimateur Aïtken Gini. L'une des particularités de cet estimateur est justement de prendre en compte les problèmes d'hétéroscédasticité et d'autocorrélation. Soit le modèle suivant :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g, \quad (2.136)$$

où \mathbf{X} est de taille $n \times (K + 1)$. Notons $\text{Var}(\varepsilon_g)$ la variance du terme d'erreur de telle sorte que $\text{Var}(\varepsilon_g) = \sigma_g^2 \mathbf{\Omega}$, avec $\mathbf{\Omega}$ une matrice de taille $n \times n$, donnée par $\mathbf{\Omega} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}'$ où $\mathbf{\Omega}^{-1} = \mathbf{P}'\mathbf{P}$ et $\mathbf{P}' = \mathbf{C}\mathbf{\Lambda}^{-\frac{1}{2}}$. Soit $\mathbf{y}^* := \mathbf{P}\mathbf{y}$, $\mathbf{X}^* := \mathbf{P}\mathbf{X}$ et $\varepsilon_g^* := \mathbf{P}\varepsilon_g$ il en résulte que :

$$\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{X}\boldsymbol{\beta}_g + \mathbf{P}\varepsilon_g \iff \mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}_g + \varepsilon_g^*. \quad (2.137)$$

Ainsi

$$\mathbb{E}[\varepsilon_g^* \varepsilon_g^{*'}] = \mathbf{P}\sigma_g^2 \mathbf{\Omega}\mathbf{P}' = \sigma_g^2 \mathbb{I}_n. \quad (2.138)$$

De plus $\mathbf{X}^* := \mathbf{P}\mathbf{X}$, $\mathbf{R}^* := \mathbf{R}'_{\mathbf{X}^*}$, $\mathbf{y}^* := \mathbf{P}\mathbf{y}$ et $\widetilde{\mathbf{X}}' := \mathbf{R}^*\mathbf{P}$. L'application de la régression Gini semi-paramétrique donne estimateur Aïtken Gini :

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{R}'_{\mathbf{X}^*} \mathbf{X}^*)^{-1} \mathbf{R}'_{\mathbf{X}^*} \mathbf{y}^* = (\mathbf{R}^* \mathbf{P}\mathbf{X})^{-1} \mathbf{R}^* \mathbf{P}\mathbf{y} = (\widetilde{\mathbf{X}}' \mathbf{X})^{-1} \widetilde{\mathbf{X}}' \mathbf{y}. \quad (2.139)$$

Comme nous l'a suggéré Yitzhaki lui-même, cet estimateur dépend de la variance. Un réel estimateur de type Aïtken devrait être obtenu en utilisant l'indice de Gini comme mesure de variabilité. C'est ce que nous proposons un peu plus loin.

L'objectif principal de cette section est d'illustrer l'intérêt de cette méthode par deux exemples d'applications théoriques : améliorer le test de Shelef et proposer l'estimateur à effets aléatoires.

2.4.1 Test de Shelef (2014)

Shelef (2014) a montré que le test de Dickey-Fuller fournit de faibles puissances en présence de valeurs aberrantes. Elle a proposé un test de racine unitaire basé sur l'estimation du processus AR(1) à l'aide de la régression Gini. Soit le processus autorégressif d'ordre 1 suivant :

$$\mathbf{y}_t = \beta_0 + \beta_1 \mathbf{y}_{t-1} + \varepsilon_t. \quad (2.140)$$

L'estimateur de la régression Gini de l'équation (2.140) est donné par :

$$\hat{b}_1^G = \frac{\text{Cov}(\mathbf{y}_t, \mathbf{R}_y(\mathbf{y}_{t-1}))}{\text{Cov}(\mathbf{y}_{t-1}, \mathbf{R}_y(\mathbf{y}_{t-1}))}. \quad (2.141)$$

Puisque \hat{b}_1^G est une \mathbb{U} -statistique, nous pouvons estimer son écart-type de manière

convergente en utilisant la procédure du Jackknife :

$$\hat{SD}(\hat{b}_1^G) = \sqrt{\frac{n-1}{n} \sum_i (\hat{b}_{1(-i)}^G - \hat{b}_{1(\cdot)}^G)^2}, \quad (2.142)$$

avec $b_{1(\cdot)}^G$ l'estimateur de la régression Gini calculé sur l'ensemble de l'échantillon et $b_{1(-i)}^G$ l'estimateur de la régression Gini calculé après suppression d'une observation ($i = 1, \dots, n$). La statistique de test de Shelef (2014) est :

$$GDF = \frac{\hat{b}_1^G - 1}{\hat{SD}(\hat{b}_1^G)}. \quad (2.143)$$

Pour réussir à faire une approche pratique du test, nous avons décidé de partir sur un modèle autorégressif d'ordre 1 avec l'hypothèse que le processus d'innovation est un bruit blanc, comme la procédure de Shelef l'indique. Le but étant de vérifier la puissance de ce test par le biais de plusieurs étapes détaillées de la façon suivante :

- Génération d'une série $\mathbf{y}_t = \beta_0 + \beta_1 \mathbf{y}_{t-1} + \varepsilon_t$ avec ε_t suivant une loi normale.
- Calcul de la statistique de test GDF .
- Calcul des résidus de la régression Gini $e_t^{G1} = \mathbf{y}_t - (\beta_0 + \hat{\beta}_1^{G1} \mathbf{y}_{t-1})$.
- Via la méthode du rééchantillonnage consistant à faire des tirages successifs avec remise, on prélève des échantillons Y^B sur la base desquels on calcule le bootstrap des résidus.
- Sur la base de la nouvelle série obtenue Y^B à l'aide de la méthode bootstrap, on constitue la série Y^* de sorte que $\mathbf{y}_t^{G1*} = \mathbf{y}_{t-1}^{G1*} + e_t^{G1*}$.
- Calcul de la statistique GDF^* .
- Répétition de la procédure M fois, ce qui nous permet d'aboutir à M statistiques GDF^* .
- Calcul de la statistique de test dont la règle de décision est la suivante. Si $GDF < w_\alpha^*$, on rejette l'hypothèse nulle : pas de présence de racine unitaire, avec w_α^* le fractile associé au risque α . Le fractile est obtenu à partir des M GDF^* que l'on classe par ordre croissant et ainsi on trouve le w_α^* au rang de la distribution GDF^* . Par exemple $w_{.10}^*$ est obtenu au rang 10 de la distribution GDF^* .
- On répète M fois la totalité des étapes citées ci-dessus pour calculer la proportion de rejet de l'hypothèse nulle.

Notons que les valeurs critiques sont obtenues en calculant les résidus de la régression Gini e_t^{G1*} de telle sorte que la série bootstrap est ensuite calculée sous l'hypothèse nulle H_0 avec $\beta = 1$:

$$\mathbf{y}_t^{G1*} = \mathbf{y}_{t-1}^{G1*} + e_t^{G1*}, \quad (2.144)$$

où \mathbf{y}_0^{G1*} est la valeur médiane de l'échantillon initial. La procédure de rééchantillonnage consistant à faire des tirages avec remise de l'ensemble des résidus estimés conduit à construire un processus qui permet de calculer l'estimateur bootstrap. Ainsi la statistique de test bootstrap basée sur la régression Gini est :

$$GDF^* = \frac{\hat{\beta}_1^{G1*} - 1}{\hat{SD}(\hat{\beta}_1^{G1*})}, \quad (2.145)$$

où $\hat{\beta}_1^{G1*}$ est l'estimateur de régression Gini, calculé sur l'échantillon bootstrap,

$$\hat{\beta}_1^{G1*} = \frac{\text{Cov}(\mathbf{y}_t^*, \mathbf{R}(\mathbf{y}_{t-1}^*))}{\text{Cov}(\mathbf{y}_{t-1}^*, \mathbf{R}(\mathbf{y}_{t-1}^*))}. \quad (2.146)$$

Cette procédure GDF^* est employée pour imiter la distribution de test sous l'hypothèse nulle ($H_0 : \beta_1 = 1$) afin d'obtenir les valeurs critiques.

Globalement, ce test semble être à la fois robuste et valide pour une large gamme de distributions. Toutefois, comme nous pouvons le voir dans le tableau ci-dessous, ce test souffre d'une perte de puissance lorsque l'hypothèse d'absence d'autocorrélation n'est plus satisfaite. La prise en compte de l'autocorrélation impacte directement la puissance du test. Soit le modèle auto-corrélé

$$\mathbf{y}_t = \phi_0 + \phi_1 \mathbf{y}_{t-1} + \varepsilon_t, \quad (2.147)$$

avec $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$. On montre ainsi que la puissance n'est plus constante, elle baisse au fur et à mesure que le coefficient de corrélation ρ augmente, passant ainsi de 0.07 à 0.05. Ces résultats nous permettent d'observer les limites du test de Shelef en présence d'autocorrélation. En effet, la significativité du test s'en voit considérablement altérée. Les résultats auxquels nous aboutissons avec différents coefficients d'autocorrélation sont les suivants.

TABLE 2.7 – Perte de puissance du Test de Shelef

| coefficient de corrélation | Puissance du test |
|----------------------------|-------------------|
| $\rho = 0.2$ | 0.075 |
| $\rho = 0.5$ | 0.0655 |
| $\rho = 0.7$ | 0.0501 |

2.4.2 Extension du test de Shelef (2014)

Considérons le processus autorégressif ci-dessous :

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + u_t, \quad (2.148)$$

avec $\mathbb{E}(u_t) = 0$, $\mathbb{E}(u_t^2) = \sigma_u^2$ et $\text{Cov}(u_t, u_s) = 0$ pour tout $t \neq s$. Traditionnellement, le coefficient d'autocorrélation habituel ρ est tel que :

$$\rho_1 = \frac{\text{Cov}(\varepsilon_t, \varepsilon_{t-1})}{\sqrt{\text{Var}(\varepsilon_t)\text{Var}(\varepsilon_{t-1})}}. \quad (2.149)$$

Suivant Shelef and Schechtman (2011), si ε_t est *strictement stationnaire*, deux fonctions d'autocorrélation Gini peuvent être définies à partir de deux co-Gini : $\text{Cog}(\varepsilon_t, \varepsilon_{t-s}) := \text{Cov}(\varepsilon_t, \mathbf{F}(\varepsilon_{t-s}))$ et $\text{Cog}(\varepsilon_{t-s}, \varepsilon_t) := \text{Cov}(\varepsilon_{t-s}, \mathbf{F}(\varepsilon_t))$. La première est la suivante :

$$\rho_s^{G_1} = \frac{\text{Cov}(\varepsilon_t, \mathbf{F}(\varepsilon_{t-s}))}{\text{Cov}(\varepsilon_t, \mathbf{F}(\varepsilon_t))}, \quad (2.150)$$

La seconde est donnée par :

$$\rho_s^{G_2} = \frac{\text{Cov}(\varepsilon_{t-s}, \mathbf{F}(\varepsilon_t))}{\text{Cov}(\varepsilon_{t-s}, \mathbf{F}(\varepsilon_{t-s}))}. \quad (2.151)$$

Comme le soulignent Carcea and Serfling (2014), la prise en considération de ces deux fonctions d'autocorrélation au sens de Gini présente un avantage considérable : si un moment de toute série (ε_t ou ε_{t-s}) est inconnu, l'identification du processus ARMA reste possible ; si le second moment de ε_t n'existe pas où si $\text{Var}(\varepsilon_t) > \infty$ à cause des valeurs aberrantes, le recours à la fonction d'autocorrélation Gini apparaît comme une solution.

Hypothèse 2.4.1 – Stationnarité au sens de Gini – Carcea and Serfling (2015)

Le processus ε_t est dit stationnaire au sens de Gini si :

$$\begin{aligned}\text{Cog}(\varepsilon_{t+s}, \varepsilon_t) &= \text{Cog}(\varepsilon_{1+s}, \varepsilon_1) \text{ pour tout } s = 0, \pm 1, \pm 2, \dots \\ \text{Cog}(\varepsilon_t, \varepsilon_{t+s}) &= \text{Cog}(\varepsilon_1, \varepsilon_{1+s}) \text{ pour tout } s = 0, \pm 1, \pm 2, \dots \\ &= \text{Cog}(\varepsilon_{1-s}, \varepsilon_1) \text{ pour tout } s \geq 1.\end{aligned}$$

Hypothèse 2.4.2 – Indépendance au Sens de Gini –

Les processus ε_t et u_t sont indépendants au sens de Gini si :

$$\text{Cog}(u_t, \varepsilon_r) = 0 \quad \forall t, r = 1, \dots, T.$$

Il est intéressant de noter que l'estimateur de la régression Gini du processus auto-régressif est donné par :

$$\frac{\text{Cog}(u_{t-r}, \varepsilon_{t-s})}{\text{Cog}(\varepsilon_{t-s}, \varepsilon_{t-s})} =: b_{r,s}^G, \quad (2.152)$$

avec

$$\varepsilon_t = [\rho_1^G]^s + \sum_{r=0}^{s-1} [\rho_1^G]^r u_{t-r}. \quad (2.153)$$

Par conséquent, en posant $\text{Cog}(\varepsilon_t, \varepsilon_t) = \text{Cov}(\varepsilon_t, \mathbf{F}(\varepsilon_t)) =: g_0$ on obtient :

$$\text{Cog}(\varepsilon_t, \varepsilon_{t-s}) = \mathbb{E}(\varepsilon_t' \mathbf{F}(\varepsilon_{t-s})) = [\rho_1^G]^s g_0 + \sum_{r=0}^{s-1} \text{Cog}(u_{t-r}, \varepsilon_{t-s}) = [\rho_1^G]^s g_0. \quad (2.154)$$

Maintenant, nous pouvons définir la matrice d'autocorrélation :

$$\mathbb{E}(\boldsymbol{\varepsilon} \mathbf{F}'(\boldsymbol{\varepsilon})) \equiv \boldsymbol{\Omega}^G = g_0 \begin{pmatrix} 1 & \rho_1^G & [\rho_1^G]^2 & \dots & [\rho_1^G]^{T-1} \\ \rho_1^G & 1 & \rho_1^G & \dots & [\rho_1^G]^{T-2} \\ \vdots & & & & \vdots \\ [\rho_1^G]^{T-1} & [\rho_1^G]^{T-2} & [\rho_1^G]^{T-3} & \dots & 1 \end{pmatrix}. \quad (2.155)$$

Rappelons que $\text{Var}(\boldsymbol{\varepsilon}_g) := \sigma_g^2 \boldsymbol{\Omega}$, avec $\boldsymbol{\Omega}$ étant une matrice $n \times n$. Il en résulte que :

$$\boldsymbol{\Omega}^{-1/2} \mathbf{y} = \boldsymbol{\Omega}^{-1/2} \mathbf{X} \boldsymbol{\beta}_g + \boldsymbol{\Omega}^{-1/2} \boldsymbol{\varepsilon}_g \iff \mathbf{y}^* = \mathbf{X}^* \boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g^*. \quad (2.156)$$

Ainsi l'application de la régression Gini semi-paramétrique conduit à l'estimateur Aïtken Gini :

$$\hat{\beta}_g = (\mathbf{R}'_{\mathbf{x}^*} \mathbf{X}^*)^{-1} \mathbf{R}'_{\mathbf{x}^*} \mathbf{y}^* = (\mathbf{R}'_{\mathbf{x}^*} \boldsymbol{\Omega}^{-1/2} \mathbf{X})^{-1} \mathbf{R}'_{\mathbf{x}^*} \boldsymbol{\Omega}^{-1/2} \mathbf{y} = (\widetilde{\mathbf{X}}' \widetilde{\mathbf{X}})^{-1} \widetilde{\mathbf{X}}' \mathbf{y}. \quad (2.157)$$

Ainsi, nous obtenons un estimateur semblable à ceux des variables instrumentales. Cet estimateur dérivé de la régression Gini permet de tenir compte du problème que pose l'autocorrélation et les valeurs aberrantes. Par conséquent, l'application du test de Shelef (2014) avec processus autocorrélé ne pose plus aucun problème de perte de puissance. Toutefois, cet estimateur nécessite une estimation convergente de $\boldsymbol{\Omega}$. Nous proposons d'estimer $\boldsymbol{\Omega}$ par $\boldsymbol{\Omega}_G$ qui est une U-statistique. Avec

$$\hat{\rho}_1^G = \frac{\text{Cov}(e_t, \mathbf{R}(e_{t-s}))}{\text{Cov}(e_t, \mathbf{R}(e_t))} = \frac{\mathbb{U}'_1}{\mathbb{U}'_2} \quad (2.158)$$

\mathbb{U}'_1 et \mathbb{U}'_2 sont les U-statistiques, correspondant respectivement aux estimateurs convergents et sans biais de $\text{Cov}(e_t, \mathbf{F}(e_{t-s}))$ et $\text{Cov}(e_t, \mathbf{F}(e_t))$, données par les expressions suivantes :

$$\mathbb{U}'_1 = [1/(T-1)T] \sum_{t=1}^T \mathbf{R}_t(e_{t-s})e_t \quad \text{et} \quad \mathbb{U}'_2 = [1/(T-1)T] \sum_{t=1}^T \mathbf{R}_t(e_t)e_t. \quad (2.159)$$

Ainsi, l'estimateur Aïtken Gini, $\hat{\beta}_g$, est donné par l'expression suivante :

$$\hat{\beta}_g = (\mathbf{R}_G^* \boldsymbol{\Omega}_G^{-1/2} \mathbf{X})^{-1} \mathbf{R}_G^* \boldsymbol{\Omega}_G^{-1/2} \mathbf{y} = (\mathbf{R}_G^* \mathbf{X}^*)^{-1} \mathbf{R}_G^* \mathbf{y}^* \quad (2.160)$$

avec $\mathbf{R}_G^* = \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X})$.

TABLE 2.8 – Correction du Test de Shelef

| coefficient de corrélation | Puissance du test |
|----------------------------|-------------------|
| $\rho = 0.2$ | 0.873 |
| $\rho = 0.5$ | 0.867 |
| $\rho = 0.7$ | 0.858 |

Ainsi, comme nous pouvons l'observer dans la table 2.8 ci-dessus, l'application du test de Shelef (2014) avec processus autocorrélé ne pose plus aucun pro-

blème de perte de puissance lorsque les erreurs sont autocorrélées.

2.4.3 L'estimateur Gini du modèle à effets aléatoires

Dans la section précédente, nous avons introduit l'estimateur à effets fixes au sens de Gini. Dans cette sous-section nous nous intéressons aux modèles à effets aléatoires. Puisque dans ce modèle, les spécificités individuelles apparaissent au niveau de la variance des variables endogènes l'estimateur du Gini serait inapproprié. Ainsi, nous utiliserons l'Aïtken Gini.

Soit le modèle élémentaire à erreurs composées :

$$\mathbf{y}_{nt} = \beta_0 + \sum_{k=1}^K \beta \mathbf{x}_{knt} + \varepsilon_{nt}, \quad (2.161)$$

avec

$$\varepsilon_{nt} = U_n + W_{nt} \quad (2.162)$$

où \mathbf{y}_{nt} est la variable expliquée, \mathbf{x}_{nt} le vecteur des variables explicatives, β_0 la constante et ε_{nt} le terme d'erreur composé de deux éléments : une perturbation idiosyncratique W_{nt} (de variance σ_w^2) et un effet individuel U_n (de variance σ_u^2). Ces effets sont supposés non autocorrélés. On suppose également qu'ils sont non corrélés avec les \mathbf{x}_{nt} . Ainsi, ce modèle apparaît comme un modèle de régression dont les perturbations exhibent une forme particulière d'autocorrélation. Il en résulte que l'estimateur des MCO serait inapproprié. En supposant que la matrice de variance covariance Ω est connue, l'estimateur des moindres carrés généralisés (MCG) est donné alors par l'expression suivante :

$$\hat{\beta}_{MCG} = (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}(\mathbf{X}'\Omega^{-1}\mathbf{y}). \quad (2.163)$$

Avec

$$\Omega = \mathbb{E}(\varepsilon\varepsilon') = \begin{pmatrix} \mathbb{E}(\varepsilon_1\varepsilon_1') & \mathbb{E}(\varepsilon_1\varepsilon_2') & \dots & \mathbb{E}(\varepsilon_1\varepsilon_N') \\ \mathbb{E}(\varepsilon_2\varepsilon_1') & \mathbb{E}(\varepsilon_2\varepsilon_2') & \dots & \mathbb{E}(\varepsilon_2\varepsilon_N') \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(\varepsilon_N\varepsilon_1') & \mathbb{E}(\varepsilon_N\varepsilon_2') & \dots & \mathbb{E}(\varepsilon_N\varepsilon_N') \end{pmatrix} \quad (2.164)$$

L'estimateur MCG peut être vu comme une combinaison linéaire des variabilités inter et intra-individuelles des observations. En effet

$$\hat{\beta}_{MCG} = \left\{ \mathbf{S}_{xx}^{intra} + \theta \mathbf{S}_{xx}^{inter} \right\}^{-1} \left\{ \mathbf{S}_{xy}^{intra} + \theta \mathbf{S}_{xy}^{inter} \right\}, \quad (2.165)$$

avec \mathbf{S}_{xx}^{intra} et \mathbf{S}_{xx}^{inter} les variabilités intra et inter-individuelles :

$$\mathbf{S}_{xx}^{intra} = \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_{n.})(\mathbf{x}_{nt} - \mathbf{x}_{n.}) \quad (2.166)$$

$$\mathbf{S}_{xx}^{inter} = \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{n.} - \mathbf{x}_{..})(\mathbf{x}_{n.} - \mathbf{x}_{..}) \quad \text{et} \quad \theta = \frac{\sigma_w^2}{\sigma_w^2 + T\sigma_u^2} \quad (2.167)$$

Ceci montre que $\hat{\beta}_{MCG}$ est une moyenne pondérée des estimateurs intra et inter-individuels. En effet

$$\hat{\beta}_{MCG} = \mathbf{F}^1 \hat{\beta}^{intra} + \mathbf{F}^2 \hat{\beta}^{inter}, \quad (2.168)$$

avec

$$\mathbf{F}^1 = \left[\mathbf{S}_{xx}^{intra} + \theta \mathbf{S}_{xx}^{inter} \right]^{-1} \mathbf{S}_{xx}^{intra} \quad (2.169)$$

$$\mathbf{F}^2 = \left[\mathbf{S}_{xx}^{intra} + \theta \mathbf{S}_{xx}^{inter} \right]^{-1} \left(\theta \mathbf{S}_{xx}^{inter} \right). \quad (2.170)$$

L'estimateur des MCG considère la même pondération pour les variabilités intragroupes et intergroupes, *cf. e.g.* Baltagi (2008). Bien que l'estimateur des MCG jouit de très bonnes propriétés, il ne résout en rien les problèmes des points aberrants et peut être affecté par une transformation monotone des données. En effet, \mathbf{S}_{xx}^{intra} et \mathbf{S}_{xy}^{intra} (respectivement \mathbf{S}_{xx}^{inter} et \mathbf{S}_{xy}^{inter}) restent très sensibles aux outliers. Comme nous le savons une approche alternative pour surmonter cette difficulté serait d'utiliser une autre mesure de dispersion comme la GMD. Reconsidérons le modèle à erreurs composées précédent :

$$\mathbf{y}_{nt} = \beta_0 + \sum_{k=1}^K \beta \mathbf{x}_{knt} + \varepsilon_{nt}. \quad (2.171)$$

Comme l'estimateur MCG, l'estimateur Aïtken Gini s'obtient en appliquant la régression Gini au modèle transformé suivant :

$$\Omega^{-1/2} \mathbf{y}_{nt} = \Omega^{-1/2} \beta \mathbf{x}_{nt} + \Omega^{-1/2} \varepsilon_{nt}, \quad (2.172)$$

Ce qui conduit à :

$$\hat{\beta}_{AG} = \left\{ \mathbf{R}(\boldsymbol{\Omega}^{-1/2} \mathbf{x}) \boldsymbol{\Omega}^{-1/2} \mathbf{x} \right\}^{-1} \left\{ \mathbf{R}(\boldsymbol{\Omega}^{-1/2} \mathbf{x}) \boldsymbol{\Omega}^{-1/2} \mathbf{y} \right\}. \quad (2.173)$$

Comme nous l'a recommandé Yithzaki un estimateur convergent de $\boldsymbol{\Omega}$ peut être obtenu en utilisant la matrice $\boldsymbol{\Omega}_G$ suivante que nous nommerons matrice de Gini-Cog :

$$\boldsymbol{\Omega}_G = \mathbb{E}(\boldsymbol{\varepsilon} \mathbf{F}(\boldsymbol{\varepsilon}')) = \begin{pmatrix} \mathbb{E}(\boldsymbol{\varepsilon}_1 \mathbf{F}(\boldsymbol{\varepsilon}'_1)) & \mathbb{E}(\boldsymbol{\varepsilon}_1 \mathbf{F}(\boldsymbol{\varepsilon}'_2)) & \dots & \mathbb{E}(\boldsymbol{\varepsilon}_1 \mathbf{F}(\boldsymbol{\varepsilon}'_N)) \\ \mathbb{E}(\boldsymbol{\varepsilon}_2 \mathbf{F}(\boldsymbol{\varepsilon}'_1)) & \mathbb{E}(\boldsymbol{\varepsilon}_2 \mathbf{F}(\boldsymbol{\varepsilon}'_2)) & \dots & \mathbb{E}(\boldsymbol{\varepsilon}_2 \mathbf{F}(\boldsymbol{\varepsilon}'_N)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(\boldsymbol{\varepsilon}_N \mathbf{F}(\boldsymbol{\varepsilon}'_1)) & \mathbb{E}(\boldsymbol{\varepsilon}_N \mathbf{F}(\boldsymbol{\varepsilon}'_2)) & \dots & \mathbb{E}(\boldsymbol{\varepsilon}_N \mathbf{F}(\boldsymbol{\varepsilon}'_N)) \end{pmatrix} \quad (2.174)$$

Ainsi, l'estimateur $\hat{\beta}_{AG}$ du modèle à effets aléatoires s'écrit comme suit :

$$\hat{\beta}_{AG} = \left\{ \mathbf{R}'(\boldsymbol{\Omega}_G^{-1/2} \mathbf{x}) \boldsymbol{\Omega}_G^{-1/2} \mathbf{x} \right\}^{-1} \left\{ \mathbf{R}'(\boldsymbol{\Omega}_G^{-1/2} \mathbf{x}) \boldsymbol{\Omega}_G^{-1/2} \mathbf{y} \right\} = (\mathbf{R}'_{\mathbf{x}^*} \mathbf{x}^*)^{-1} \mathbf{R}'_{\mathbf{x}^*} \mathbf{y}^* \quad (2.175)$$

Avec $\mathbf{x}^* = \boldsymbol{\Omega}_G^{-1/2} \mathbf{x}$; $\mathbf{y}^* = \boldsymbol{\Omega}_G^{-1/2} \mathbf{y}$ et $\mathbf{R}'_{\mathbf{x}^*} = \mathbf{R}'(\boldsymbol{\Omega}_G^{-1/2} \mathbf{x})$. Nous prouvons dans les lignes qui suivent que $\hat{\beta}_{AG}$ peut être vu comme une combinaison linéaire des variabilités inter et intra-individuelles des observations. Soient J_T une matrice carrée ($T \times T$) composée uniquement de 1 et \mathbf{x}_n le vecteur ($T \times 1$) des observations de \mathbf{x} pour le nième individu. Ainsi,

$$\mathbf{x} - \mathbf{x}_n = \begin{pmatrix} \mathbf{x}_{n1} - \mathbf{x}_n \\ \mathbf{x}_{n2} - \mathbf{x}_n \\ \vdots \\ \mathbf{x}_{nT} - \mathbf{x}_n \end{pmatrix} = \left(I_T - \frac{J_T}{T} \right) \mathbf{x}_n$$

Il s'en suit que

$$\mathbf{X} - \mathbf{X}_n = \left[I_N \otimes \left(I_T - \frac{J_T}{T} \right) \right] \mathbf{X}_n.$$

Ainsi, $\mathbb{W}_n = \left[I_N \otimes \left(I_T - \frac{J_T}{T} \right) \right]$ permet de calculer les écarts de chaque observation aux moyennes individuelles qu'on appelle habituellement "l'opérateur intra-individuel". Par analogie, nous pouvons déduire "l'opérateur inter-individuel" \mathbb{B}_n

donné par l'expression suivante :

$$\mathbb{B}_n = I_{NT} - \mathbb{W}_n = I_N \otimes \frac{J_T}{T}$$

. En utilisant ces deux opérateurs nous pouvons réécrire $\hat{\beta}_{AG}$ comme suit (voir annexe 3.3.3 pour la démonstration) :

$$\beta_{AG} = \left[\mathbf{G}^*_{xx}{}^{intra} + \theta' \mathbf{G}^*_{xx}{}^{inter} \right]^{-1} \left[\mathbf{G}^*_{xy}{}^{intra} + \theta' \mathbf{G}^*_{xy}{}^{inter} \right], \quad (2.176)$$

avec $\theta' = \sqrt{\phi}$. Ainsi, tout comme les MCG, l'estimateur Aïtken Gini peut être considéré comme une combinaison optimale des estimateurs intra et inter-groupes.

2.5 Conclusion

L'objectif premier de ce chapitre était de revenir sur les généralités de la régression Gini. La variance est la mesure de dispersion la plus populaire. Toutefois, elle suppose l'hypothèse de symétrie des distributions marginales. Ce qui peut apparaître restrictif dans la mesure où l'analyse économique considère souvent des relations non symétriques. L'utilisation de la différence moyenne du Gini permet de surmonter cet obstacle. Il semble qu'il existe plus d'une douzaine de représentations alternatives du GMD (Yitzhaki, 1998). L'intérêt essentiel que présente la multiplicité de ces définitions est de montrer l'avantage d'utiliser la GMD. Chaque définition ouvre la voie à de nombreux développements et interprétations. Par exemple, la définition basée sur la covariance permet de définir la co-Gini, la G-corrélation, l'ANOGL, la régression Gini, etc.

La GMD partage plusieurs propriétés avec la variance et la principale différence entre ces deux mesures est la norme utilisée : la variance se base sur la distance euclidienne tandis que la GMD utilise la "city block" distance (norme ℓ_1). En outre, le Gini est à la fois sensible à la distribution des revenus et à la distribution des rangs c'est-à-dire à la stratification. C'est principalement pour cette raison qu'on dit que la décomposition du Gini est plus complète que celle de la variance. Par conséquent, l'ANOGL serait plus appropriée que l'ANOVA lorsque la stratification a des implications sur l'interprétation des résultats. Par ailleurs, comme nous l'avons vu, la co-Gini n'est pas symétrique. C'est cela qui explique

d'ailleurs le fait qu'il existe deux types de régression Gini. La première est basée sur la minimisation du GMD des erreurs. Comme la régression par quantile, l'estimateur obtenu n'a pas d'expression explicite et il ne peut être obtenu que numériquement. La seconde approche consiste à construire l'estimateur en question de sorte qu'il soit moins sensible aux outliers. Les estimateurs obtenus dans ces deux approches sont généralement différents et ils sont équivalents uniquement lorsque le modèle est linéaire. Ainsi, nous pouvons tester l'hypothèse de linéarité en comparant les estimateurs de ces deux approches de la régression Gini. En outre, cette dernière est moins sensible que les MCO aux transformations monotones des données et aux points aberrants. Ainsi, elle est plus robuste et peut être vue comme une méthode par variable instrumentale (Ka and Mussard, 2015). Par ailleurs, contrairement aux MCO qui imposent trois conditions (conditions de Grenander) pour garantir le bon comportement de la matrice des moments, la régression Gini n'impose que deux conditions. Ces dernières que nous avons introduit dans Ka and Mussard (2016) sont moins restrictives.

Le second objectif de ce chapitre, qui est le principal d'ailleurs, était d'étendre cette réflexion sur la régression Gini à l'économétrie des données de panels. La principale difficulté de cette tâche était de décomposer la variabilité totale au sens de Gini en variabilité intragroupe et en intergroupe sans l'indice de chevauchement. Nous avons montré qu'il était possible de réaliser une telle décomposition. Ainsi, nous avons pu définir les estimateurs intragroupes et intergroupes dans le cadre du Gini. Comme tout estimateur basé sur le Gini, nos estimateurs sont des \mathbb{U} -statistiques et donc convergent vers la loi normale. Nous avons également proposé un test de l'existence des effets individuels et plusieurs autres développements théoriques.

Pour terminer, notons que ces dernières décennies, plusieurs travaux ont tenté de proposer des estimateurs du modèle à effets fixe moins sensible que les MCO aux points aberrants. Nous pouvons citer entre autres la M-estimation, la S-estimation (Verardi and Wagner, 2011). La particularité de notre estimateur est qu'il est en plus d'être moins sensible que les MCO aux points extrêmes, il n'impose pas de forme fonctionnelle particulière. Cette propriété est particulièrement importante dans le domaine de la croissance pro-pauvre où plusieurs travaux ont montré la sensibilité des résultats d'estimations aux formes fonctionnelles retenues.

Chapitre 3

Estimation de la croissance pro-pauvre

Ce chapitre revient sur l'approche économétrique de la croissance pro-pauvre. Son objectif est double. Tout d'abord, nous revenons sur les limites de l'approche par les indices. A l'heure actuelle, on doit admettre que les résultats obtenus par cette dernière approche ne peuvent pas être considérés comme réellement convaincants. En effet, comme nous allons le voir dans ce chapitre, cette approche peut conduire parfois à des résultats contradictoires. En outre, beaucoup de chercheurs s'indignent du caractère partiel des indices de croissance pro-pauvre et du fait qu'ils ne nous permettent pas de conclure dans certaines situations. Ensuite, nous nous intéresserons à l'approche économétrique. Cette dernière, bien qu'elle représente une alternative très intéressante, souffre de certaines difficultés. Les travaux empiriques sur la croissance pro-pauvre sont généralement sujets à deux types de biais que sont le biais de sélection et le biais d'endogenéité. Ces derniers s'expliquent principalement par les limitations inhérentes aux données. En effet, malgré les améliorations considérables de ces dernières années, les données sur la croissance pro-pauvre continuent de poser de sérieux problèmes aux chercheurs. Ces derniers déplorent le nombre élevé d'observations manquantes, les erreurs de mesures et la présence de points aberrants. Par ailleurs, dans la plupart des travaux empiriques les échantillons sont composés d'une soixantaine de pays parmi lesquels les pays développés sont surreprésentés et les pays pauvres surtout ceux d'Afrique sub-saharienne presque inexistantes. Ceci conduit à des échantillons très hétérogènes et à une sélection drastique. A cela s'ajoute, le fait

que la dimension temporelle des données est extrêmement limitée. Ce qui fait que la variance inter-individuelle est très supérieure à la variance inter-temporelle (Cogneau et al., 2002; White and Anderson, 2001; Ghura et al., 2002, entre autres).

Par ailleurs, il serait nécessaire de tester la robustesse des conclusions en analysant les canaux de transmission et de voir aussi si les résultats restent valables dans d'autres échantillons et également avec des combinaisons différentes de variables explicatives. C'est pourquoi, nous avons construit deux échantillons en données de panel. L'avantage des données de panels est qu'elles permettent de tenir compte de l'hétérogénéité entre les pays. Mais également, la double dimension des données de panel et leurs grands nombres permettent d'obtenir plus de variabilité, plus de degré de liberté et moins de colinéarité.

La première section revient sur les limites de l'approche par les indices. Elle propose une étude comparative des mesures de la croissance pro-pauvre. Plus précisément, nous allons étudier le caractère inclusif de la croissance Thaïlandaise en utilisant plusieurs indices : CIC, CPP, IPCC et TECP. Le cas de la Thaïlande est particulièrement intéressant. Ce pays a connu une croissance économique régulière durant les années 80, ce qui a conduit à réduire très sensiblement l'incidence de la pauvreté. Mais, en même temps les inégalités de revenu ont été très fortes. D'ailleurs, elle a été longtemps considérée comme l'un des pays les plus inégalitaires en terme de distribution de revenu. Malheureusement, ces inégalités de revenu pourraient limiter l'impact de la croissance sur la réduction de la pauvreté. En effet, la théorie de la croissance pro-pauvre soutient que la lutte contre les inégalités de revenu est un vecteur important de lutte contre la pauvreté. Ainsi, nous pourrions légitimement nous demander en présence d'inégalité de revenu aussi importante dans quelle mesure la croissance Thaïlandaise est-elle pro-pauvre ? Plusieurs études se sont intéressées à cette question. La particularité de celle-ci est qu'elle utilise un ensemble d'indices de croissance pro-pauvre.

La seconde section s'intéresse aux conséquences de la limitation des données, aux choix de spécification et de la méthode d'estimation. La sensibilité des estimateurs par effets fixes standards aux outliers, aux erreurs de mesures et aux formes fonctionnelles fait que nous privilégierons l'utilisation de notre estimateur intragroupe de la régression Gini. Ce dernier est plus robuste que l'estimateur intra basé sur les MCO Ka and Mussard (2016). Par ailleurs, pour tenir compte du biais de sélection, nous avons construit deux échantillons en données de panels.

Un premier échantillon qui couvre la période 2003-2007 et comportant 69 pays dont 37 sont des pays développés ou en transition et 32 sont des pays en voie de développement. Et un deuxième échantillon à partir du premier en considérant uniquement les pays en voie de développement et en transition. Ce traitement nous permet d'avoir un second échantillon constitué de données homogènes.

3.1 Limites de l'approche par les indices de croissance pro-pauvre

A l'heure actuelle, il n'existe pas de mesure de la croissance pro-pauvre faisant l'unanimité. L'une des principales limites de ces indices est qu'ils se focalisent uniquement sur la dimension monétaire de la pauvreté, ce qui peut paraître très restrictif bien que nous reconnaissons le lien entre la dimension monétaire et non monétaire de la pauvreté. En outre, comme nous allons le voir à travers l'exemple ci-dessous, ces indices présentent des limites plus sérieuses. Tout d'abord, parfois ils ne nous permettent pas de conclure. De plus, il faut également faire attention sur le fait que ces derniers peuvent aboutir à des conclusions contradictoires. Cela s'explique par le fait que chaque indice de croissance pro-pauvre se rattache à une définition donnée : globale, absolue et relative. Malheureusement, les travaux empiriques ne prennent généralement pas en compte ces limites. Afin, de mieux illustrer ces insuffisances, nous allons dans cette section utiliser un ensemble d'indices pour étudier le caractère inclusif de la croissance Thaïlandaise des années 90.

L'exemple de la Thaïlande est particulièrement intéressant. Ce pays a connu durant les années 80 et jusqu'au milieu des années 90 une forte croissance économique mais également une forte augmentation des inégalités de revenu. En effet, durant cette période le taux de croissance annuel moyen a été de 9.5%. Ces performances économiques s'expliquent notamment par une globalisation commerciale et un ensemble de réformes économiques : stabilité macroéconomique, une déréglementation commerciale, une plus grande diversification de ses activités industrielles (textile, montage électronique, l'agroalimentaire etc). Mais également à un développement fulgurant du secteur des services (tourisme notamment). Cette croissance régulière des années 80 a permis de réduire très significativement l'incidence de la pauvreté (Krongkaew, 1996; Kakwani and Krongkaew, 2000). En

effet, cette dernière est passée de 27% au début des années 90 à environ 16% en 1999 (Table 3.1).

TABLE 3.1 – Incidence de la pauvreté en Thaïlande 1990-2000

| Années | 1990 | 1992 | 1994 | 1996 | 1998 | 2000 |
|------------------------------|------------------|------|------|------|------|------|
| Incidence de la pauvreté | 27,2 | 23,2 | 16,3 | 11,4 | 12,9 | 15,9 |
| Nombre des pauvres (million) | 15,3 | 13,5 | 9,7 | 6,8 | 7,9 | |
| Source | Krongkeaw (2000) | | | | | |

Cette croissance économique soutenue a été la principale explication de la réduction de l'incidence de la pauvreté. Toutefois, comme nous l'avons souligné dans le premier chapitre, une erreur courante est de croire qu'une croissance économique forte conduit nécessairement à une croissance pro-pauvre. En effet, cette dernière dépend également du niveau de la distribution des revenus et pour la Thaïlande les inégalités de revenu ont été très importantes. Ainsi, il nous paraît essentiel de se demander si cette forte performance économique a pleinement bénéficié aux pauvres ? Cette section se donne pour objectif de répondre à cette question. Notons que plusieurs travaux (Kakwani et al., 2003; Son, 2007) se sont déjà intéressés à cette question. La particularité de cette étude est qu'elle propose d'utiliser un ensemble de mesures de la croissance pro-pauvre : la courbe d'incidence de la croissance (CIC) de Ravallion and Chen (2003), la courbe de croissance pro-pauvre (CPP) de Son (2004), le taux de croissance équivalent à la pauvreté (TCEP) de Kakwani and Son (2002) et l'indice de croissance pro-pauvre (ICPP) de Kakwani et al. (2000). Nous utiliserons les dépenses de consommations déflatées de l'indice harmonisé des prix comme indicateur de bien-être. Les données utilisées dans ce travail proviennent du Socio-Economic Surveys (SES).

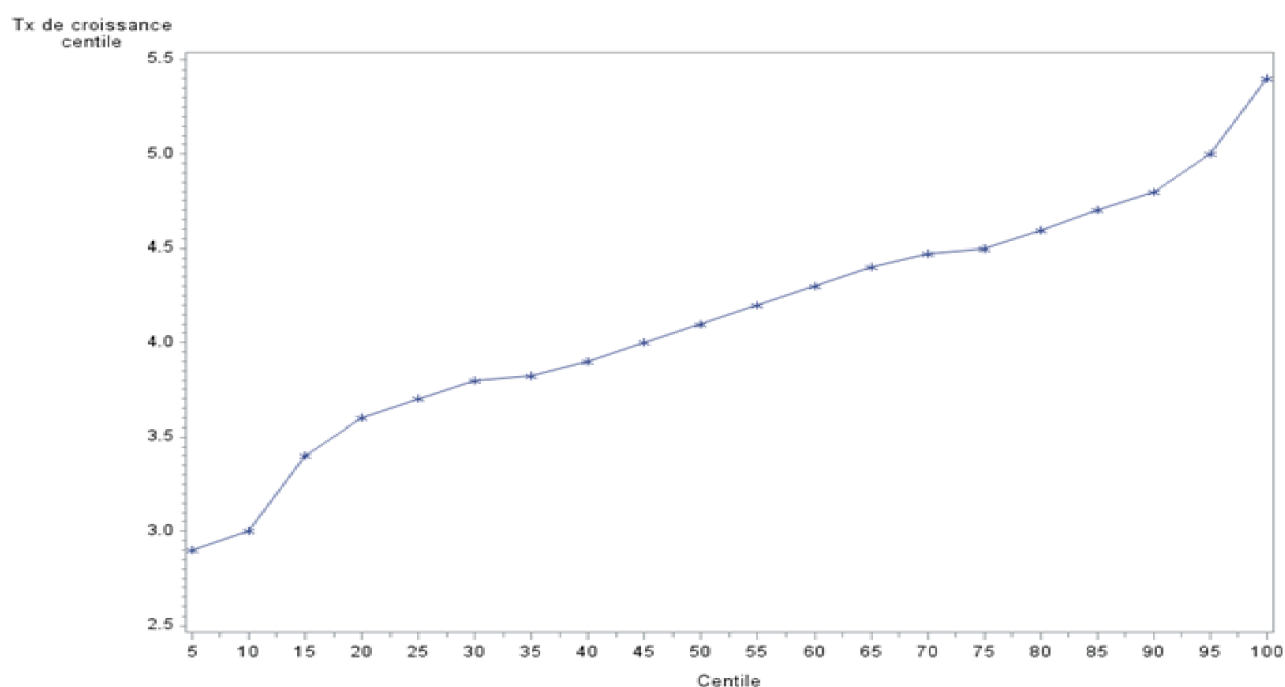
3.1.1 Croissance pro-pauvre en Thaïlande avec la CIC.

Le choix de cet indicateur est principalement guidé par le fait que la CIC (courbe d'incidence de la croissance) est la mesure la plus utilisée dans les travaux sur la croissance pro-pauvre. Cela s'explique certainement par sa simplicité de mise en œuvre mais également par le fait qu'elle peut être utilisée pour n'importe quel indicateur de bien être : monétaire et non monétaire. Par exemple

Grosse et al. (2008) l'utilise dans la dimension non monétaire de la pauvreté. Elle ne nécessite pas non plus de définir un seuil de pauvreté, ce qui constitue un avantage considérable. Selon cette méthode la croissance est pro-pauvre si, pour tout centile p , la courbe se situe au-dessus de l'axe des abscisses. On dit alors qu'il y a dominance stochastique de premier ordre de la distribution à la période t sur la distribution à la période $t - 1$. Toutefois, si la CIC change de signe ou plus précisément s'il n'y a pas de dominance stochastique de premier ordre on ne pourra pas conclure avec cet indice.

Pour construire la courbe d'incidence de la croissance (CIC) pro-pauvre, nous avons calculé les taux de croissance des centiles des dépenses de consommation sur la période 1990-2000.

FIGURE 3.1 – Courbe d'incidence de la croissance

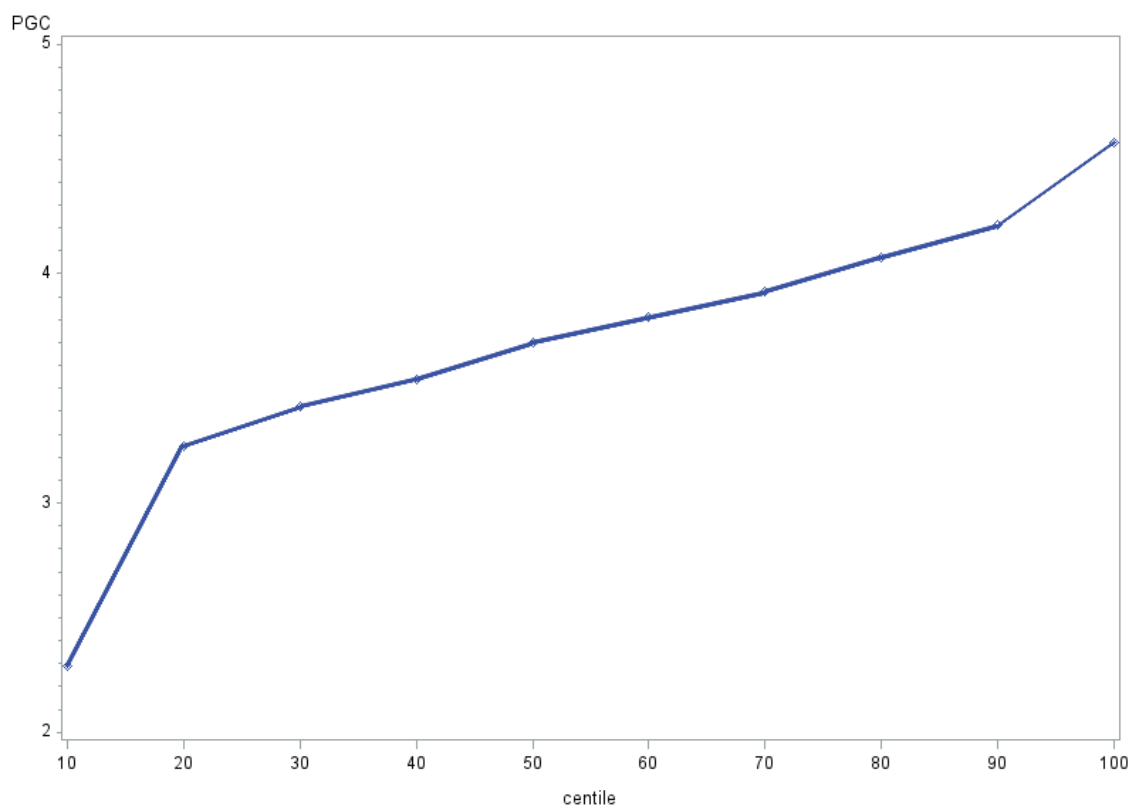


La croissance est pro-pauvre lorsque la courbe est au-dessus de l'axe des abscisses. Ainsi, globalement nous pouvons conclure que la croissance a été favorable aux pauvres sur la période 1990-2000. Cela signifie que la croissance a été favorable à toutes les couches de la population, notamment aux deuxième et dernier déciles, sur cette période.

3.1.2 Croissance pro-pauvre en Thaïlande avec la CPP.

Une approche alternative pour étudier la croissance pro-pauvre en Thaïlande est possible en utilisant la CPP (courbe de croissance de la pauvreté) de Son (2004). Et contrairement à la CIC, la courbe de Son (2004) est issue à partir des conditions de dominance stochastique de deuxième ordre¹. D'après cette courbe la croissance est pro-pauvre au sens absolu lorsque cette courbe se situe au-dessus de l'axe des abscisses. Ainsi, nous pouvons conclure que la pauvreté a décliné durant la période 1998-2000 en Thaïlande avec la même conclusion que l'approche par la CIC à savoir une amélioration des conditions de vie des premiers et derniers déciles.

FIGURE 3.2 – La courbe de croissance de la pauvreté



1. Rappelons que l'ordre 1 de la dominance stochastique est relatif au critère de Pareto donc non compatible avec les principes redistributifs. L'ordre 2 permet en revanche des transferts de type Pigou-Dalton.

3.1.3 Croissance pro-pauvre en Thaïlande avec le TCEP.

Une autre mesure de la croissance pro-pauvre est le TCEP (Taux de croissance équivalent à la pauvreté). Ce dernier tel que défini par Kakwani and Son (2002) correspond au taux de croissance qui engendre une baisse de la pauvreté dans la même proportion que celle obtenue avec le taux de croissance du revenu moyen (TCRM), lorsque l'on suppose les inégalités de revenu inchangées, c'est-à-dire lorsque chaque individu reçoit la même proportion des bénéfices de la croissance. Selon cette mesure la croissance sera considérée comme étant pro-pauvre (pro-riche) au sens relatif lorsque le TCEP est supérieur (inférieur) au taux de croissance du revenu ou des dépenses de consommation moyennes et pro-pauvre au sens absolu si le TCEP est positif sur toute la période.

TABLE 3.2 – Taux de Croissance Équivalent à la Pauvreté en Thaïlande 1990-2000

| Années | Taux de croissance | Taux de Croissance Équivalent à la Pauvreté | | |
|-----------|---------------------|---|---------|----------|
| | | TCEP | Absolue | Relative |
| 1990-1992 | 7.49 | 4.27 | Oui | Non |
| 1992-1994 | 7.65 | 8.82 | Oui | Oui |
| 1994-1996 | 5.75 | 7.34 | Oui | Oui |
| 1996-1998 | -1.00 | -2.67 | Non | Non |
| 1998-2000 | -0.85 | -2.32 | Non | Non |
| Source | nos propres calculs | | | |

Dans le tableau ci-dessus, nous pouvons constater que (comme chez Kakwani et al. (2004)) durant la période 1990-92 le taux de croissance par tête est de 7,49 ; alors que le TCEP est de 4,27. Cela signifie que 3,22% du taux de croissance est perdu parce que les pauvres n'ont pas pleinement bénéficié de la croissance. Ainsi, la croissance n'a pas été pro-pauvre sur cette période et sur la période 1992-1993 ; la croissance a été favorable aux plus pauvres. Cela s'explique par l'effet combiné de la croissance économique et de la baisse des inégalités de revenu. Et très globalement, sur toute la période, la croissance économique thaïlandaise n'a pas été pro-pauvre. Elle a entraîné une hausse des dépenses des ménages riches plus que proportionnelle à celles des ménages pauvres de 1990 à 2000².

2. Sur la période 1988-2000, Kakwani et al. (2004) concluent que la croissance économique Thaïlandaise n'a pas été pro-pauvre.

3.1.4 Croissance pro-pauvre au Thaïlande avec l'ICPP.

Une dernière mesure utilisée pour évaluer le caractère pro-pauvre de la croissance Thaïlande est l'ICPP (Indice de croissance pro-pauvre). Ce dernier introduit par (Kakwani et al., 2000) correspond au rapport entre la réduction totale de la pauvreté et la réduction de la pauvreté obtenue en supposant les inégalités de revenu inchangées. Ainsi, avec cette mesure, la croissance sera pro-pauvre au sens relatif lorsque l'indice est supérieur à l'unité et pro-pauvre au sens absolu lorsque qu'il est positif. Signalons toutefois, que cet indice ne respecte pas l'axiome de monotonie. Cela signifie que l'indice n'augmente pas de valeur lorsque le taux de réduction de la pauvreté est plus élevé.

TABLE 3.3 – Indice de croissance pro-pauvre en Thaïlande 1990-2000

| Années | Taux de croissance | ICPP | | |
|-----------|---------------------|------|---------|----------|
| | | ICPP | Absolue | Relative |
| 1990-1992 | 7.49 | 0.57 | Oui | Non |
| 1992-1994 | 7.65 | 1.18 | Oui | Oui |
| 1994-1996 | 5.75 | 1.29 | Oui | Oui |
| 1996-1998 | -1.00 | 2.67 | Oui | Oui |
| 1998-2000 | -0.85 | 1.35 | Oui | Oui |
| Source | nos propres calculs | | | |

Il peut être surprenant de constater que la croissance a été pro-pauvre avec la CIC et la CCP et non pro-pauvre avec le TCEP et ICPP. D'où vient cette apparente contradiction ? En réalité, ce type de résultat est très fréquent. Il s'explique principalement par le fait que ces indices se basent sur des définitions différentes de la croissance pro-pauvre : absolue (baisse du taux de pauvreté), relative (baisse des inégalités) et globale (baisse de la pauvreté et des inégalités). Comme la réalité le montre, une baisse du taux de pauvreté peut s'accompagner dans certaines situations d'une aggravation des inégalités de revenu. C'est le cas de la Thaïlande et de la Chine durant ces dernières décennies. En effet, la croissance rapide Thaïlandaise des années 80 et des années 90 a permis de réduire la pauvreté absolue mais s'est accompagnée d'une aggravation des inégalités (et donc d'une pauvreté relative, voir Table 3.4 ci-dessous). D'après Mazumdar and Son (2002), la faible mobilité de la main d'œuvre du secteur agricole au secteur industriel est la principale

responsable de cette explosion des inégalités de revenu durant cette période de forte croissance économique. D'ailleurs, cette faible mobilité de la main d'œuvre a expliqué pourquoi la crise économique de 1997 n'a pas eu d'impact significatif sur la distribution des revenus en Thaïlande (voir Table 3.4). Le premier quintile a faiblement évolué durant la crise de 1996 (4.3 en 1996, 4.3 en 1998 et 4.2 en 2000) et les écarts entre les différents quintiles sont restés stables.

TABLE 3.4 – L'évolution des inégalités en Thaïlande 1990-2000

| Années | Indice de Gini | Quintile | | | | |
|--------|------------------|----------|-----|------|------|------|
| | | 1 | 2 | 3 | 4 | 5 |
| 1990 | 0.481 | 4.4 | 7.6 | 11.7 | 19.4 | 54.9 |
| 1992 | 0.499 | 4.1 | 7.2 | 11.2 | 18.8 | 58.7 |
| 1994 | 0.486 | 4.1 | 7.4 | 11.7 | 19.6 | 57.1 |
| 1996 | 0.477 | 4.3 | 7.6 | 11.9 | 19.8 | 56.5 |
| 1998 | 0.486 | 4.3 | 7.8 | 11.9 | 19.7 | 56.3 |
| 2000 | 0.509 | 4.2 | 7.7 | 12.1 | 19.8 | 56.2 |
| Source | Krongkeaw (2000) | | | | | |

Globalement, cette étude a le mérite de montrer les limites des indices de la croissance pro-pauvre mais également de l'importance de prendre en compte la distribution du revenu dans les politiques pro-pauvres. L'incidence de la pauvreté a diminué de façon continue durant cette période malheureusement l'ampleur aurait été plus importante en l'absence d'inégalités de revenu. Ainsi, bien qu'elle revête une grande importance, la croissance n'explique pas l'ensemble de l'évolution de la réduction de l'incidence de la pauvreté. D'où l'importance de prendre en compte l'inégalité de revenu dans toute politique de croissance pro-pauvre.³

3. La forte croissance chinoise, des années 2000, a été la principale cause du rapide baisse de la pauvreté. En effet, en 2001 l'indice de croissance pro-pauvre chinoise était environ de 17% contre 64% au début des années 80 Ravallion (2009). Pourtant d'après Kakwani cette croissance n'a pas été favorable aux pauvres mais, au contraire, elle leur a été défavorable puisqu'elle s'est accompagnée d'une augmentation des inégalités.

3.2 Analyses économétriques de la croissance pro-pauvre

Cette section s'intéresse respectivement à l'impact de la croissance, des inégalités de revenu et des mécanismes de redistribution sur la réduction de la pauvreté. Son originalité est de tenter de mesurer l'impact de ces différents éléments sur les quantiles inférieurs. Ainsi, elle permet de faire des recommandations plus ciblées. En effet, pour élaborer une politique de croissance pro-pauvre, il serait indispensable d'identifier les mécanismes permettant d'accroître le revenu des quantiles inférieurs. En outre, cette technique présente plusieurs avantages : elle permet d'évaluer plus facilement le caractère pro-pauvre de la croissance et les estimations obtenues sont insensibles au seuil de pauvreté choisi. Ce dernier point est particulièrement important. Comme nous l'avons souligné précédemment estimer une ligne de pauvreté pour l'ensemble des pays de notre échantillon est un exercice assez délicat.

Nous nous intéresserons également aux conséquences des choix de spécification et de la méthode d'estimation. Par ailleurs, au-delà de la technique d'estimation utilisée, il est important de ne pas perdre de vue les problèmes liés à la sélection de l'échantillon. La littérature empirique récente est largement revenue sur les conséquences d'un tel problème. Par exemple, Ghura et al. (2002) montrent la fragilité des résultats de Dollar and Kraay (2002). Bien que leurs travaux ne soient pas facilement comparables à ceux de Dollar and Kraay (2002) du fait que les échantillons sont différents, ils montrent tout de même leurs limites et la nécessité d'utiliser des méthodes d'estimation plus robustes. Sur ce point la régression Gini semble être une méthode d'estimation particulièrement adaptée aux travaux sur la croissance pro-pauvre. Comme nous l'avons vu dans le précédent chapitre, ce type de régression présente plusieurs avantages : une moindre sensibilité que les méthodes traditionnelles aux erreurs de mesures et aux points aberrants. Mais aussi, elle est plus robuste car les estimateurs obtenus par cette approche sont moins sensibles à la sélection de l'échantillon et plus généralement aux erreurs de spécification.

Cette section conclut aux limites des mécanismes de redistribution et à la nécessité de prendre en compte les inégalités de revenu dans toutes politiques pro-pauvres. Et elle est organisée en quatre sous-sections. La première présente notre

modèle. La seconde et la troisième présentent respectivement les données et les méthodes d'estimation et enfin la dernière nos différents résultats.

3.2.1 Spécification de notre modèle

Dans l'objectif de modéliser la croissance pro-pauvre, nous avons régressé les taux de croissance des différents quantiles sur un ensemble de variables comme la croissance et les inégalités de revenu. L'originalité de cette spécification est multiple. Tout d'abord, contrairement aux approches traditionnelles, ce type de spécification nous évite de définir un seuil de pauvreté. En effet, estimer une ligne de pauvreté pour l'ensemble des pays de notre échantillon est un exercice risqué : les informations disponibles ne sont pas souvent suffisantes pour définir des seuils de pauvreté rigoureusement fondés sur la technique des coûts et des besoins de base. En outre, la plupart des travaux empiriques sur la croissance pro-pauvre tente d'estimer l'élasticité du taux de pauvreté. Bien évidemment, nous reconnaissons qu'il est important d'estimer cette élasticité mais malheureusement elle n'est pas suffisamment informative. En effet, elle ne nous dit rien sur la profondeur de la pauvreté. Enfin, cette spécification répond mieux à la question de savoir si les pauvres ont bénéficié de la croissance ?

Plus concrètement, l'équation retenue suppose une relation linéaire entre le taux de croissance de chaque quantile (Q_i) et la croissance globale (Txc). Elle fait intervenir outre l'indice de Gini, un ensemble de variables explicatives parmi lesquelles figurent les taux de scolarisations primaire (Eduprim) et secondaire (Educon), le taux d'inflation (Txinf), le log du Pib agricole $\log(\text{Piba})$, les dépenses publiques de santé (DP de santé), un terme d'interaction $[(1 - \text{Gini}) * \text{Txc}]$ et la part des dépenses publiques (DP du Gov). Elle se présente de la manière suivante :

$$Q_i = \beta_0 + \beta_1 \text{Txc} + \beta_2 \text{Gini} + \beta_3 \text{Eduprim} + \beta_4 \text{Educon} + \beta_5 \text{DP de santé} \\ + \beta_6 (1 - \text{Gini}) * \text{Txc} + \beta_7 \log(\text{Piba}) + \beta_8 (\text{DP du Gov}) + \beta_9 \text{Txinf} + \epsilon \quad (3.1)$$

Le terme d'interaction, $[(1 - \text{Gini}) * \text{Txc}]$, correspond à l'impact de la croissance sur la réduction de la pauvreté si la distribution du revenu est constante.

Plusieurs procédures d'estimation sont envisagées dans ce document. Tout

d'abord, nous allons estimer ce modèle par les approches traditionnelles (l'estimateur par effet fixe et aléatoire). Cependant, la sensibilité des estimateurs par MCO aux outliers, aux erreurs de mesures et aux formes fonctionnelles fait que cette technique reste peu adaptée à ce type de données. Le traitement qui consiste à supprimer les observations aberrantes bien qu'intéressant n'est pas pleinement satisfaisant car il conduit à supprimer beaucoup d'observations et donc à perdre de l'information ce qui n'est pas sans conséquence sur les résultats. Pour toutes ces raisons, nous avons préféré la régression Gini. L'un des intérêts que présente cette dernière méthode réside dans la possibilité de corriger à la fois ces trois limites : points aberrants, erreurs de mesure et moins de sensibilité aux formes fonctionnelles.

3.2.2 Les données

Dans ce document les inégalités de revenu sont mesurées par l'indice de Gini. Ce dernier, provient de la nouvelle version de la "World Income Inequality Database : WIID⁴ 3c. Celle-ci constitue une amélioration incontestable par rapport à la base de données de Deininger and Squire (1996). D'abord, plus récente, elle permet de disposer de beaucoup plus d'observations (7175 observations contre 2634 pour la base de Deininger and Squire (1996) et 5314 par rapport à la première version du WIID). En outre, la procédure conduisant à la sélection des données y est beaucoup moins contrainte et permet ainsi un recoupement plus fiable avec les autres bases de données (la Luxembourg Income Study notamment). Une autre particularité de cette dernière version est qu'elle contient, en plus des indices de Gini, les déciles, les quantiles et les centiles P_5 et P_{100} . Cependant, à l'image des autres bases de données sur les inégalités, les observations non renseignées sont extrêmement nombreuses dans la WIID3c.

Les variables PIB agricole par tête et taux de croissance proviennent de la base de données du "United States Department of Agriculture" (USDA). Ces données ont été compilées par "Economic Research Service (ERS)". Cette base de données contient le PIB et le taux de croissance de plus de 194 pays et couvre la période 1969 – 2015. Par ailleurs, toutes ces observations sont corrigées de l'inflation et ont été rapportées au dollar américain de 2015.

4. UNU-WIDER, 'World Income Inequality Database (WIID3c)', September 2015, <https://www.wider.unu.edu/project/wiid-world-income-inequality-database>

Les variables éducation secondaire et primaire proviennent de la base de données de Barro and Lee (2013). Les trois autres variables d'analyse proviennent des bases de données de la Banque Mondiale : le taux d'inflation, les dépenses publiques de santé et la part des dépenses publiques.

Sur un plan purement économétrique, nous devons admettre que les résultats empiriques obtenus ne sont pas réellement satisfaisants. En effet, il serait nécessaire de tester la robustesse des conclusions en analysant si elles restent valables dans d'autres échantillons et également avec des combinaisons différentes de variables explicatives. C'est pourquoi, nous avons construit deux échantillons en données panels dont les statistiques descriptives sont présentées dans la table 3.5 ci-dessous.

TABLE 3.5 – Statistiques descriptives

| | Echantillon 1 | Echantillon 2 |
|---------------------------------|---------------|---------------|
| A. Nombre de période | 2003-2007 | 2003-2007 |
| B. Nombre de pays | 69 | 32 |
| C. Nombre d'observations | 345 | 170 |
| D. Moyennes | | |
| Indice de Gini | 38.08 | 34.53 |
| Q1 | 6.53 | 7.248 |
| Q2 | 11.126 | 12.120 |
| Q3 | 15.327 | 16.238 |
| Q4 | 21.423 | 22.00 |
| Q5 | 45.436 | 42.18 |
| Eduprim | 84.112 | 99.33 |
| Edusecon | 102.563 | 101.683 |
| Taux d'inflation | 6.00 | 4.51 |
| DP de santé | 7.01 | 7.78 |
| taux de croissance (Txc) | 5.506 | 5.14 |
| log(PIBagr) | 4.51 | 4.58 |
| DP du Gov | 5.88 | 4.22 |
| [1-Gini]*Txc | -208.59 | -179.301 |

Rappelons que la double dimension des données de panel et leurs grands nombres permettent d'obtenir plus de variabilité, plus de degré de liberté, moins de colinéarité et permettent également de contrôler l'hétérogénéité des comportements. Mais surtout cette double dimension permet de tenir compte de l'influence de caractéristiques non observables des individus sur leur comportement, *cf.* Baltagi (2008).⁵ Par ailleurs, comme nous l'avons souligné dans l'introduction de ce document, modéliser la croissance pro-pauvre ne consiste pas simplement à régresser le taux de croissance et les inégalités sur les quantiles inférieurs ou sur l'incidence de la croissance. En effet, la logique requiert que l'on considère uniquement les épisodes de croissance (donc taux de croissance positif). Ainsi, dans nos échantillons nous avons considéré les pays qui ont des taux de croissance positifs.

Un premier échantillon qui couvre la période 2003 – 2007 et comportant 69 pays dont 37 sont des pays développés ou en transition et 32 sont des pays en voie de développement a été construit. Pour ces cinq périodes d'analyse, ce panel est cylindré et comporte 345 observations. Toutefois, au-delà des tests de robustesse que l'on peut effectuer, il est très probable que les résultats obtenus avec ce premier échantillon soient sujets au biais de sélection. Cette question de sélection de l'échantillon constitue selon notre point de vue un problème crucial. En effet, ce premier échantillon regroupe des pays très hétérogènes en terme d'organisation économique et sociale. Ainsi, afin de tenir compte de ce biais de sélection, nous avons construit un deuxième échantillon à partir du premier en considérant uniquement les pays en voie de développement et en transition. Ce traitement nous permet d'avoir un second échantillon constitué de données homogènes. Ce deuxième échantillon couvre la même période et est non cylindré et il comporte 170 observations.

5. En outre, un avantage décisif des données de panel par rapport aux autres types de données tient au fait qu'elles sont souvent extrêmement nombreuses ce qui conduit à des estimations dont les propriétés peuvent être assimilées aux propriétés asymptotiques des méthodes utilisées.

3.2.3 Analyse et interprétation des résultats

Concernant l'impact de la croissance sur les quantiles inférieurs, le lien apparaît positif dans le cas de l'estimateur par effets fixes basé sur les MCO (Table 3.7). Selon cette première procédure d'estimation une hausse d'un point supplémentaire du taux de croissance favoriserait une augmentation de 0,14% (20%) du revenu des 0,20% (des 40%) les plus pauvres de notre de notre échantillon. En outre, les inégalités de revenu seraient particulièrement défavorables aux plus démunis. En effet, plus les inégalités de revenu sont importantes moins la croissance profite aux pauvres. Cependant, nous devons être extrêmement prudents sur ce premier résultat obtenu. En réalité, même si la plupart des paramètres sont significatifs, l'estimation reste de très mauvaise qualité et peu robuste. C'est principalement pour ces raisons que nous considérons que l'estimation par MCO est imprécise et de faible qualité également. Cette instabilité de l'estimateur par les MCO provient de la présence de points aberrants dans notre échantillon 1. Pour s'en rendre compte nous allons tester très brièvement la présence de points aberrants.

Il existe une multitude de manière de déceler la présence de points aberrants. Les tests les plus connus sont : le Test de Dixon et le Test de Grubbs et à côté, il existe plusieurs grandeurs permettant de détecter la présence de points aberrants dans l'échantillon : la distance de Cook, Résidu studentisé, Dffits, Dfbetas, Covratio etc.

- Test de Dixon et le Test de Grubbs.

Le test de Dixon est le test le plus utilisé mais il a une très faible puissance lorsque le nombre d'observations n'est pas élevé. Le principe de ce test est très simple : il consiste à comparer les observations les plus éloignées du modèle et les observations les plus proches de l'étendue totale des résidus. Plus concrètement, on commence tout d'abord par classer les résidus du modèle par ordre croissant ($e_1 < e_2 < \dots < e_n$). Ensuite, on calcule les rapports ci-dessus.

$$\text{Si } 3 \leq n \leq 7 \quad Q1 = \frac{e_2 - e_1}{e_n - R_1} \quad \text{et} \quad Q1 = \frac{e_n - e_{n-1}}{e_n - e_1} \quad (3.2)$$

$$\text{Si } 8 \leq n \leq 12 \quad Q1 = \frac{e_2 - e_1}{e_n - e_1} \quad \text{et} \quad Q1 = \frac{e_n - e_{n-1}}{e_n - e_2} \quad (3.3)$$

$$\text{Sin} > 12 \quad Q1 = \frac{\mathbf{e}_3 - R_1}{\mathbf{e}_{n-2} - \mathbf{e}_1} \quad \text{et} \quad Q1 = \frac{\mathbf{e}_n - \mathbf{e}_{n-2}}{\mathbf{e}_n - \mathbf{e}_3} \quad (3.4)$$

Si la valeur de ces rapports est inférieure à la valeur critique (lue dans la table de Dixon), on conclut qu'il y'a des points aberrants dans notre échantillon. L'inconvénient majeur de ce test est sa faible puissance lorsque le nombre d'observation est peu élevé. Dans ce dernier cas, on privilégie le test de Grubbs.

- Le levier : il correspond à la i ème valeur de la matrice H définie (qu'on appelle habituellement la Hat matrice) ci-dessous.

$$H = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad \text{le levier est} \quad h_i = \mathbf{x}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i' \quad (3.5)$$

Lorsque

$$h_i > 2 \frac{K+1}{n}$$

alors on conclut que l'observation i est aberrante.

- Le Press (Predicted Residuals Sum of Squares) : plus elle s'éloigne de la somme des carrés des résidus, plus on suspecte l'existence de points extrêmes.

$$\text{Press} = \sum_{i=1}^n (\hat{Y}_i - \hat{Y}_i)^2 \quad (3.6)$$

- La distance de Cook : pour l'estimateur des paramètres.

Elle calcule l'écart entre l'estimateur du paramètre et ce même estimateur calculé sur le modèle privé de la i ème observation :

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_i)'(\mathbf{X}'\mathbf{X})(\hat{\beta} - \hat{\beta}_i)}{\hat{\sigma}^2(K+1)} \quad (3.7)$$

Lorsque $D_i > 1$, on conclut que l'observation i a beaucoup d'influence sur l'estimation du paramètre β .

- Dfbetas : pour l'estimateur du paramètre j .

Même principe que la distance de Cook mais sur un seul paramètre :

$$\text{Dfbetas}_{ij} = \frac{\hat{\beta}_j - \beta_{i,j}}{\hat{\sigma}_i \sqrt{(\mathbf{X}'\mathbf{X})^{-1}jj}} \quad (3.8)$$

Lorsque $|\text{Dfbetas}_{ij}| > \frac{2}{\sqrt{n}}$, on conclut que l'observation i a beaucoup d'influence sur l'estimation du paramètre β .

En appliquant ces différents tests, nous trouvons une vingtaine de points aberrants dans notre échantillon (Table3.6). Bien que ces points représentent 5% seulement de notre échantillon 1, ils peuvent compromettre la validité de nos conclusions basées sur les MCO.

TABLE 3.6 – Tests d’outliers

| Obs | Distance de Cook | DFBETAS | | | | | |
|-----|------------------|---------|----------|---------|-------------|----------------|-----------|
| | | Eduprim | Edusecon | Txinf | DP de santé | Indice de Gini | log(PIBa) |
| 25 | 2.630 | -5.0939 | 0.3027 | -0.2006 | 0.7317 | -0.4146 | 1.1881 |
| 35 | 3.130 | -2.309 | 0.504 | -0.306 | 1.031 | -0.298 | 1.952 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 237 | 1.226 | 1.101 | 0.712 | 0.502 | 2.018 | 0.0181 | 1.012 |

C’est pour cela que nous allons estimer à nouveau notre modèle en utilisant notre estimateur par effet fixe au sens de Gini. Ce qui nous amène à la Table3.8.

Les estimateurs par effets fixes MCO et Gini conduisent à des évaluations de l’impact des variations des taux de croissance et des inégalités assez sensiblement différentes. On note même une opposition de signe pour les paramètres estimés du coefficient de Gini (en couleur grise). Ces divergences ainsi observées s’expliquent principalement par la sensibilité de l’estimateur habituelle au spécification du modèle, mais également aux importantes erreurs de mesures et à la présence de points aberrants.

Ces deux résultats d’estimation qui viennent d’être présentés constituent une bonne illustration de l’avantage que présente l’utilisation de la régression Gini. Toutefois, au-delà des tests de robustesse que l’on peut effectuer, il est très probable que les résultats obtenus soient sensibles à la sélection de l’échantillon et aux outliers (Table 3.6). Ces questions concernant les outliers et la sélection de l’échantillon constituent de notre point de vue un problème à ne pas négliger. En

effet, ce premier échantillon regroupe des pays très différents en terme d'organisation économique et sociale. Et nous pouvons très légitimement nous interroger sur la pertinence d'un tel regroupement. Bien évidemment, nous pouvons utiliser des méthodes robustes à la présence d'effets fixes. Toutefois, bien que certaines différences organisationnelles et sociales peuvent être fixes sur une période donnée, elles ont souvent des effets non négligeables sur la constitution du revenu des plus pauvres dans le moyen et le long terme. Par exemple, dans les pays riches, il existe souvent un ensemble de systèmes de protection sociale qui aident assez efficacement les plus démunis à se protéger contre les effets des chocs macroéconomiques. C'est pour cela que nous avons construit un second échantillon à partir du premier en considérant uniquement les pays en voie de développement. Ce traitement nous permet d'avoir un échantillon constitué de données homogènes dont les résultats sont présentés dans la sous-section suivante.

TABLE 3.7 – Estimation de notre modèle par les MCO

| $\beta =$ | Estimateur par effet fixe (MCO) | | | | |
|---------------------------------|---------------------------------|--------------------------------|--------------------------------|-------------------------------|-------------------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | -0.041 [⊙] (0.069) | -0.047 [⊙] (0.066) | -0.027 [⊙] (0.099) | 0.051 [⊙] (0.058) | 0.234 [⊙] (0.194) |
| Edu secon | 0.013 [⊙] (0.010) | 0.08 [⊙] (0.09) | 0.027 [⊙] (0.034) | 0.05 [⊙] (0.08) | 0.029 [⊙] (0.028) |
| Taux d'inflation | -0.04* (0.011) | -0.051* (0.010) | -0.045* (0.026) | -0.025** (0.015) | 0.021 [⊙] (0.031) |
| DP de santé | -0.088 [⊙] (0.072) | -0.069 [⊙] (0.067) | -0.089 [⊙] (0.002) | 0.034 [⊙] (0.06) | 0.198 [⊙] (0.199) |
| Indice de Gini 10 ⁻¹ | -0.19* (0.018) | -0.018* (0.017) | -0.037* (0.025) | -0.052* (0.015) | 0.087** (0.069) |
| Log(PIBagri) | 0.087** (0.105) | 0.021** (0.095) | 0.032** (0.087) | 0.023** (0.04) | -0.055** (0.14) |
| Taux de croissance | 0.145* (0.071) | 0.207* (0.067) | 0.226* (0.101) | 0.100* (0.059) | 0.190* (0.200) |
| (1-Gini)*Txc 10 ⁻² | 0.044* (0.002) | 0.052* (0.002) | 0.048* (0.003) | 0.0295* (0.002) | 0.821* (0.005) |
| DP du Gov 10 ⁻² | 0.15 (0.02) | 0.06 [⊙] (0.067) | 0.115 [⊙] (0.021) | 0.084 [⊙] (0.081) | -0.33 [⊙] (0.045) |
| R ² | 0.3330 | 0.3230 | 0.3907 | 0.3599 | 0.3909 |
| Durbin Watson | 0.324 | 0.042 | 0.002 | 0.017 | 0.093 |
| Breusch-Pagan | 12.25 | 9.75 | 24.12 | 14.36 | 13.15 |

- Les valeurs entre parenthèses correspondent aux écarts-types ⁶.
- Les paramètres avec un exposant [⊙] sont non significatifs.

6. Les écarts-types dans le cas du Gini sont calculés en utilisant la méthode du Jackknife.

TABLE 3.8 – Estimation de notre modèle par le Gini

| Estimates → $\beta =$ | Estimateur par effet fixe (Gini) | | | | |
|---------------------------------|----------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | -0.080 [⊙] (0.091) | -0.015 [⊙] (0.021) | 0.061 [⊙] (0.075) | 0.052 [⊙] (0.055) | 0.041 [⊙] (0.047) |
| Edu secon | 0.024* (0.006) | 0.030* (0.007) | 0.050* (0.010) | 0.04* (0.021) | 0.057** (0.045) |
| Taux d'inflation | -0.053* (0.022) | -0.075* (0.031) | -0.052* (0.024) | -0.017* (0.010) | -0.012 [⊙] (0.020) |
| DP de santé | -0.066 [⊙] (0.055) | -0.096 [⊙] (0.089) | -0.054 [⊙] (0.044) | 0.008 [⊙] (0.06) | 0.025 [⊙] (0.023) |
| Indice de Gini 10 ⁻¹ | -0.130* (0.071) | -0.135* (0.094) | -0.031* (0.012) | -0.108* (0.094) | 0.051* (0.015) |
| Log(PIBagri) | 0.016* (0.056) | 0.022** (0.013) | 0.061** (0.038) | 0.071 [⊙] (0.047) | -0.0374 [⊙] (0.029) |
| Taux de croissance | 0.12** (0.052) | 0.130* (0.045) | 0.153* (0.098) | 0.160* (0.045) | 0.140* (0.01) |
| (1-Gini)*Txc 10 ⁻² | 0.066* (0.002) | 0.036* (0.001) | 0.018* (0.002) | 0.041* (0.001) | 0.012* (0.003) |
| DP du Gov 10 ⁻² | 0.015 [⊙] (0.091) | 0.021 [⊙] (0.072) | -0.028 [⊙] (0.041) | -0.024 [⊙] (0.051) | -0.074 [⊙] (0.069) |
| GR ² | 0.301 | 0.343 | 0.351 | 0.320 | 0.405 |
| Durbin Watson | 1.78 | 1.25 | 0.99 | 1.35 | 1.56 |

- Les paramètres avec un exposant * sont significatifs à 5%.
- Les paramètres avec un exposant ** sont significatifs à 10%.

3.2.4 Estimations de notre modèle sur notre second échantillon

Les tables 3.9 et 3.10 ci-dessous présentent respectivement les résultats de nos estimations avec les MCO et le Gini sur notre second échantillon. Comparer aux résultats précédents l'amélioration de la qualité de l'estimation est particulièrement frappante. Une simple observation de ces tableaux nous permet de constater l'évolution de certains paramètres estimés et l'importance de tenir compte des spécificités des pays. Globalement, la croissance s'accompagne d'une réduction de la pauvreté. En outre, on voit bien, en comparant les résultats obtenus par la régression Gini dans les deux échantillons que l'impact de la croissance agricole sur la réduction de la pauvreté serait supérieur dans les pays à faibles revenus. Selon les estimations sur notre premier échantillon le PIB agricole n'a pas une influence significative sur le caractère pro-pauvre de la croissance. Pour rappel, dans le chapitre 1 de ce document, nous avons vu que l'agriculture joue un rôle important dans la réduction de la pauvreté. D'ailleurs, plusieurs travaux ont montré qu'à dépense égale c'est la recherche agricole qui a le plus permis de réduire la pauvreté dans les pays en voie de développement. D'après nos estimations, ce type de résultats reste vrai dans les pays pauvres caractérisés par des économies fondées généralement sur l'agriculture. C'est pour cette raison, que le PIB agricole n'a pas une incidence significative sur la réduction les quantiles inférieurs de notre premier échantillon⁷. Par ailleurs, comme le souligne à juste titre Wiggins and Higgins (2008), il est important de réduire les distorsions des marchés agricoles afin de pouvoir favoriser l'impact de la croissance agricole sur la réduction de la pauvreté.

Par ailleurs, durant ces dernières décennies, une abondante littérature est revenue sur l'importance de l'accès à l'éducation dans la lutte contre la pauvreté. En effet, l'éducation joue un rôle essentiel pour trouver un emploi et atteindre l'autonomie personnelle. Toutefois, bien que le taux de scolarisation secondaire joue favorablement à la réduction de la pauvreté, l'analyse révèle que le taux de scolarisation primaire n'a pas une influence significative sur les quantiles inférieurs. La principale explication de ce résultat est que stimuler l'accès à l'éducation primaire n'est efficace pour réduire la pauvreté qu'à condition que les populations

7. La grande majorité des pays de notre premier échantillon sont des pays industrialisés ou en transition et leurs économies ne reposent pas sur l'agriculture. Ce qui fait que la réduction de la pauvreté n'est pas clairement due à la croissance agricole

concernées puissent poursuivre leurs formations afin de bénéficier pleinement des rendements marginaux élevés associés à la formation longue.

Sans surprise l'inflation a un impact négatif sur les quantiles inférieurs. Elle érode le pouvoir d'achat des pauvres. Une hausse de dix pourcent du taux d'inflation réduit la croissance du premier quantile d'environ six points. En réalité, l'inflation peut être considérée comme une taxe régressive qui affecte davantage les plus pauvres. En effet, ces derniers sont souvent moins outillés que les autres pour se protéger contre l'inflation. Ils n'ont pas souvent accès aux produits financiers qui permettent de protéger la valeur de leur patrimoine.

Globalement, même si la croissance s'accompagne d'une amélioration du revenu des plus pauvres, ces derniers ont moins bénéficié des résultats de la croissance que les riches. En outre, ces résultats mettent en évidence l'importance de ne pas négliger la lutte contre les inégalités. L'analyse montre que ces dernières agissent négativement sur le taux de croissance des quantiles inférieurs. Malheureusement, les politiques de lutte contre les inégalités sont souvent négligées dans la lutte contre la pauvreté. D'ailleurs, les estimations ci-dessus montrent que c'est la croissance corrigée des inégalités qui a le plus d'influence sur la réduction de la pauvreté. Ce paramètre est plus important chez les quantiles inférieurs.

TABLE 3.9 – Estimation de notre modèle par les MCO

| Estimates → $\beta =$ | Estimateur par effet fixe (MCO) | | | | |
|---------------------------------|---------------------------------|--------------------------------|---------------------------------|--------------------------------|--------------------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | -0.075 [⊙] (0.081) | -0.056 [⊙] (0.083) | -0.051 [⊙] (0.027) | 0.041 [⊙] (0.055) | 0.027 [⊙] (0.0233) |
| Edu secon | 0.070* (0.012) | 0.056* (0.012) | 0.062* (0.017) | 0.020** (0.014) | 0.05** (0.026) |
| Taux d'inflation | -0.034* (0.0021) | -0.021* (0.0022) | -0.014* (0.0033) | -0.014* (0.0014) | 0.010** (0.0061) |
| DP de santé | 0.013 [⊙] (0.016) | -0.065 [⊙] (0.083) | -0.013 [⊙] (0.0158) | -0.012 [⊙] (0.068) | 0.038 [⊙] (0.029) |
| Indice de Gini 10 ⁻¹ | -0.117* (0.026) | -0.111* (0.027) | 0.037** (0.019) | 0.077** (0.043) | 0.062 [⊙] (0.052) |
| Log(PIBagri) | 0.04* (0.002) | 0.053* (0.027) | 0.048** (0.026) | 0.027** (0.015) | -0.093 [⊙] (0.056) |
| Taux de croissance | 0.145* (0.097) | 0.157* (0.099) | 0.205* (0.052) | 0.154* (0.066) | 0.183* (0.079) |
| (1-Gini)*Txc 10 ⁻² | 0.055* (0.03) | 0.041* (0.03) | 0.095* (0.04) | 0.037* (0.02) | 0.013** (0.08) |
| DP du Gov 10 ⁻² | 0.011 [⊙] (0.08) | 0.693 [⊙] (0.09) | 0.529 [⊙] (0.023) | 0.055 [⊙] (0.06) | -0.017 [⊙] (0.024) |
| R ² | 0.4627 | 0.4021 | 0.4558 | 0.3396 | 0.4722 |
| Durbin Watson | 1.02 | 0.96 | 1.024 | 0.77 | 1.24 |
| Breusch-Pagan | 8.25 | 10.15 | 14.12 | 18.76 | 15.24 |

TABLE 3.10 – Estimation de notre modèle par le Gini

| Estimates → $\beta =$ | Estimateur par effet fixe (Gini) | | | | |
|---------------------------------|----------------------------------|--------------------------------|-------------------------------|-------------------------------|--------------------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | -0.011 [⊙] (0.07) | -0.035 [⊙] (0.065) | 0.029 [⊙] (0.051) | 0.045 [⊙] (0.054) | 0.024 [⊙] (0.031) |
| Edu secon | 0.087* (0.0010) | 0.065* (0.0009) | 0.054* (0.0011) | 0.061* (0.0009) | 0.104* (0.0054) |
| Taux d'inflation | -0.089* (0.001) | -0.097* (0.0012) | -0.049* (0.003) | -0.048* (0.003) | 0.050* (0.001) |
| DP de santé | -0.011** (0.046) | -0.119** (0.071) | 0.066** (0.039) | 0.009** (0.005) | 0.017** (0.011) |
| Indice de Gini 10 ⁻¹ | -0.116* (0.009) | -0.130* (0.012) | -0.092* (0.011) | -0.129* (0.021) | 0.05** (0.025) |
| Log(PIBagri) | 0.097* (0.005) | 0.011* (0.006) | 0.010* (0.009) | 0.014* (0.004) | 0.008* (0.0022) |
| Taux de croissance | 0.110* (0.007) | 0.108* (0.0019) | 0.135* (0.0012) | 0.109* (0.009) | 0.150* (0.0019) |
| (1-Gini)*Txc 10 ⁻² | 0.039* (0.001) | 0.025* (0.001) | 0.014* (0.002) | 0.029* (0.001) | 0.04* (0.004) |
| DP du Gov 10 ⁻² | 0.013 [⊙] (0.05) | 0.014 [⊙] (0.09) | 0.038 [⊙] (0.03) | 0.019 [⊙] (0.05) | -0.086 [⊙] (0.094) |
| GR ² | 0.512 | 0.569 | 0.612 | 0.491 | 0.641 |
| Durbin Watson | 1.34 | 1.16 | 1.74 | 0.84 | 1.65 |

3.2.5 D'autres méthodes d'estimations

Afin de mieux illustrer l'avantage de la régression Gini nous allons nous concentrer sur notre second échantillon. L'objectif est de comparer les résultats obtenus avec cet échantillon à partir des deux méthodes d'estimation. Plus précisément, nous allons tout d'abord présenté dans la table 3.11, outre les estimations usuelles (MCO, inter, intra), l'estimateur MCQG et l'estimateur MCO du modèle en différence première (DP). Ce dernier représente une solution alternative à l'estimateur intra, lorsque l'on est confronté à un problème de corrélation des effets individuels. Ensuite, nous présenterons dans la table 3.12, l'équivalent de ces estimateurs dans le cadre de la régression Gini, il s'agit notamment de l'estimateur de la régression Gini, de l'inter Gini, l'intra Gini, l'Aïtken Gini et enfin l'estimateur Gini du modèle en différence première (DP). Toutefois, pour une question de présentation, nous ne présentons ici que les résultats relatifs au premier quantile. Les résultats complets sont disponibles en annexe (p173 à p175).

TABLE 3.11 – Estimation de notre modèle sur le premier quintile MCO

| Estimates → $\beta =$ | Estimateur MCO | | | | |
|---------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------|--------------------------------|
| | Intra individuel | Inter individuel | MCO | MCQG | DP |
| Edu prim | -0.075 [⊙] (0.081) | -0.050 [⊙] (0.051) | -0.037 [⊙] (0.042) | -0.014 [⊙] (0.015) | -0.011 [⊙] (0.012) |
| Edu secon | 0.070* (0.012) | 0.048* (0.010) | 0.021* (0.011) | 0.019* (0.009) | 0.15* (0.075) |
| Taux d'inflation | -0.034* (0.0024) | -0.056* (0.035) | -0.016* (0.01004) | -0.007* (0.014) | -0.001* (0.061) |
| DP de santé 10 ⁻² | 0.013 [⊙] (0.016) | 0.035 [⊙] (0.073) | 0.011 [⊙] (0.027) | 0.012 [⊙] (0.06875) | 0.026 [⊙] (0.029) |
| Indice de Gini 10 ⁻¹ | -0.117* (0.026) | -0.065* (0.029) | -0.224* (0.012) | -0.017* (0.018) | -0.162* (0.075) |
| Log(PIBagri) | -0.04* (0.002) | -0.182* (0.153) | -0.086** (0.019) | -0.007** (0.005) | -0.063** (0.023) |
| Taux de croissance | 0.145* (0.097) | 0.181* (0.0201) | 0.098* (0.017) | 0.057* (0.036) | 0.23* (0.027) |
| (1-Gini)*Txc 10 ⁻² | 0.055* (0.03) | 0.056* (0.05) | 0.070* (0.019) | 0.097* (0.02) | 0.12** (0.08) |
| DP du Gov 10 ⁻² | 0.011 [⊙] (0.08) | 0.099 [⊙] (0.06) | 0.055 [⊙] (0.018) | 0.142 [⊙] (0.06) | 0.087 [⊙] (0.024) |
| R ² | 0.4627 | 0.887 | 0.889 | 0.856 | 0.47 |

- Les paramètres avec un exposant [⊙] sont non significatifs.
- Les paramètres avec un exposant * sont significatifs à 5%.
- Les paramètres avec un exposant ** sont significatifs à 10%.

TABLE 3.12 – Estimation de notre modèle sur le premier quintile Gini

| Estimates → $\beta =$ | Estimateur Gini | | | | |
|---------------------------------|-------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | Intra Gini | Inter Gini | Gini | Aïtken Gini | DP Gini |
| Edu prim | -0.011 [∅] (0.07) | -0.012 [∅] (0.051) | -0.011 [∅] (0.085) | -0.013 [∅] (0.058) | -0.009 [∅] (0.003) |
| Edu secon | 0.087 (0.0010) | 0.094 (0.011) | 0.091 (0.009) | 0.079 (0.009) | 0.065 (0.024) |
| Taux d'inflation | -0.089* (0.001) | -0.099* (0.003) | -0.097* (0.0012) | -0.095* (0.003) | -0.101* (0.001) |
| DP de santé 10 ⁻¹ | -0.011** (0.046) | -0.013 [∅] (0.089) | -0.120** (0.911) | -0.014 [∅] (0.042) | -0.009 [∅] (0.120) |
| Indice de Gini 10 ⁻¹ | -0.116* (0.009) | -0.13* (0.011) | -0.141* (0.012) | -0.129* (0.021) | -0.15* (0.025) |
| Log(PIBagri) | 0.097* (0.005) | 0.119* (0.009) | 0.117* (0.006) | 0.014** (0.004) | 0.179 [∅] (0.022) |
| Taux de croissance | 0.110* (0.0070) | 0.095* (0.012) | 0.086* (0.017) | 0.109* (0.009) | 0.120* (0.019) |
| (1-Gini)*Txc 10 ⁻² | 0.039* (0.001) | 0.045* (0.002) | 0.042* (0.001) | 0.044* (0.001) | 0.021* (0.004) |
| DP du Gov 10 ⁻² | 0.013 [∅] (0.05) | 0.016 [∅] (0.03) | 0.014 [∅] (0.05) | 0.012 [∅] (0.05) | 0.026 [∅] (0.014) |
| GR ² | 0.512 | 0.751 | 0.792 | 0.779 | 0.761 |

Dans un premier temps, nous observons une relative similarité entre l'estimateur total et l'estimateur inter. Ce type de résultat est très connu en économétrie. En effet, comme nous l'avons souligné dans le chapitre 2, en panel la variabilité inter-individuelle représente l'essentiel de la variabilité totale. A l'opposé, les estimations intra, MCQG et en différence première sont sensiblement différentes des précédentes. L'impact de la croissance sur le premier quantile est

nettement plus faible que celui obtenu précédemment. Pourtant, logiquement, ces estimateurs devraient être convergents et relativement proches. Bien évidemment, sous l'hypothèse que le modèle soit correctement spécifié et que toutes les hypothèses standards soient remplies. Par conséquent, nous pouvons en conclure qu'au moins le modèle n'est pas correctement spécifié. Au passage nous pouvons souligner que le R^2 du modèle intra est plus faible que les autres. L'explication principale de ce résultat est que ce modèle ne comporte pas de constante⁸ et donc l'analyse de la variance n'est plus vérifiée.

Les problèmes d'autocorrélation (ou d'hétéroscédasticité) des erreurs ne conduisent pas à des estimateurs biaisés. Nous pouvons dire que les divergences précédemment obtenues dans le cadre des MCO ne sont pas dues à la non validité de ces deux hypothèses. Elles sont probablement dues à une éventuelle endogénéité des variables explicatives ou plus généralement à une mauvaise spécification de la forme fonctionnelle retenue. Comme nous le savons un problème d'endogénéité induit la non convergence de l'estimateur MCQG alors que l'estimateur intra reste lui convergent. Ainsi, il est tout à fait possible d'envisager un test d'endogénéité à la Hausman en comparant l'estimateur MCQG et l'estimateur intra-individuelle.

$$H = (\hat{\beta}_{intra} - \hat{\beta}_{mcqg})'(\hat{V}(\hat{\beta}_{intra}) - \hat{V}(\hat{\beta}_{mcqg}))^{-1}(\hat{\beta}_{intra} - \hat{\beta}_{mcqg}) \quad (3.9)$$

Cette statistique suit asymptotiquement un khi-deux à $\dim(\beta)$ degrés de liberté. L'application de ce test, pour les équations considérées, nous conduit à rejeter systématiquement l'exogénéité de l'effet individuel.

TABLE 3.13 – Tests d'endogénéité

| | Test de Hausman | | | | |
|-------------|-----------------|--------|--------|--------|--------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| H | 72 | 56 | 81 | 69 | 78 |
| $\chi^2(9)$ | 16.919 | 16.919 | 16.919 | 16.919 | 16.919 |

Sans effectuer ce test, une simple observation de la table 3.11 pouvez nous

8. Si il n'y a pas de constante dans le modèle, le R^2 peut devenir négatif.

permettre de conclure à une éventuelle endogénéité des régresseurs. Toutefois, il faudra noter que l'écriture du modèle en intra ou en inter interdit l'estimation des coefficients des variables constantes dans la dimension temporelle. En outre, la divergence entre les estimateurs intra et MCQG peut provenir plus généralement d'une erreur de spécification. C'est pourquoi, il est important de recourir à des techniques d'estimation alternatives moins sensibles aux erreurs de spécifications. L'endogénéité est un problème classique en économétrie et la méthode des variables instrumentales demeure une solution efficace. Le problème se ramène alors à trouver des variables qui soient de bons instruments. Sur ce point, plusieurs pistes de recherche ont été envisagées dans la littérature : par exemple, Hausman and Taylor (1981) ont proposé d'utiliser les variables exprimées en écarts aux moyennes individuelles. Nous utilisons ici la proposition de Durbin (1954) qui consiste à utiliser le rang des variables explicatives comme instruments ce qui correspond à la régression Gini. L'utilisation de cette dernière permet de surmonter les problèmes d'endogénéité. C'est pour cela que nous nous intéressons à présent aux estimations obtenues dans les différentes dimensions de la variabilité en utilisant cette régression. Ceci nous amène au tableau 3.12 ci-dessus.

Contrairement aux MCO les résultats obtenus dans les différentes dimensions de la variabilité au sens de Gini nous conduisent à des estimations relativement proches. En effet, les estimations Gini, inter et intra conduisent à des évaluations de l'impact de la croissance et des inégalités sur les quantiles assez identiques, ce qui constitue un argument supplémentaire en faveur de la régression Gini. En réalité, la relative proximité des estimateurs découlant des différentes méthodes de centrage constitue en général une validation de la spécification du modèle. Ainsi, il ne faudrait pas réduire la régression Gini à une simple technique de traitement des points aberrants. Comme nous l'avons vu précédemment cette technique de régression jouit de plusieurs autres avantages : estimateur sans biais, convergent, moins sensible à la spécification du modèle et aux erreurs de mesures⁹.

9. D'ailleurs, l'insensibilité des estimations de la régression Gini aux points aberrants n'est qu'une propriété de la régression Gini. Et cette dernière découle du fait que la GMD est plus adaptée lorsque l'hypothèse de normalité n'est pas remplie et plus généralement lorsque la distribution sous-jacente des données a une queue épaisse.

3.3 Conclusion

L'objectif premier de ce chapitre était de revenir sur les limites des mesures traditionnelles de la croissance pro-pauvre. Comme nous l'avons bien vu, dans la première section de ce chapitre, l'approche par les indices présente des limites qu'on devrait prendre plus au sérieux. En effet, en plus de leur caractère partiel, ces méthodes conduisent souvent à des résultats contradictoires. En outre, dans certaines situations, elles ne nous permettent pas conclure. Au vu de cela, l'approche économétrique représente une belle alternative mais cette dernière est sujette à deux types de biais : le biais de sélection et le biais d'endogénéité. Parallèlement, elle souffre aussi de la sensibilité des résultats d'estimations à la forme fonctionnelle choisie voir Alesina and Rodrik (1994). Par ailleurs, les erreurs de mesures sur certaines variables explicatives comme l'indice de Gini constituent une autre préoccupation qu'on devrait prendre plus au sérieux. Malgré, les efforts visant à améliorer (ou élargir) les données existantes, elles restent insatisfaisantes. Les données sur la croissance pro-pauvre diffèrent entre les pays en termes de couverture géographique (couverture nationale, urbaine ou rurale), d'unités statistiques (familles, ménages ou individus) et de définition du revenu (dépenses de consommation, revenu disponible ou revenu brut). Et très globalement, ces erreurs de mesure engendrent une sous-estimation, en valeur absolue, de la vraie valeur du paramètre de la variable sujette à cette erreur de mesure. Cette sous-estimation est d'autant plus importante que la part de la variance de l'erreur de mesure dans la variance totale de la variable observée est élevée et/ou que le coefficient estimé est lui même élevé. Ainsi, il y'a une très bonne raison de privilégier la régression Gini. D'ailleurs, comme nous l'avons vu dans ce chapitre, les estimateurs qui découlent du Gini sont en général plus robustes.

Une autre conclusion importante de ce chapitre est que globalement le processus de croissance favorise la réduction de la pauvreté à condition que les inégalités de revenu soient maîtrisées. L'analyse nous conduit à affirmer également que les mécanismes de redistribution jouent un rôle assez mineur (à quelques exceptions près). Par ailleurs, d'après nos résultats, l'impact de la croissance agricole sur la réduction de la pauvreté varie en fonction du niveau de développement du pays. Dans les pays à faible revenu l'incidence de la pauvreté a tendance à être extrêmement sensible à la croissance du secteur agricole. Pour terminer, une politique de croissance favorable aux pauvres doit s'appuyer sur un ensemble de

mesures de nature à générer une croissance économique forte et soutenue. Cela passe nécessairement par une stabilité macro-économique, une faible inflation, des progrès massifs en scolarisation surtout dans le secondaire.

Conclusion Générale

Le propos de notre thèse était de modéliser la croissance pro-pauvre. Elle présente des apports théoriques et empiriques. Dans un premier temps, nous nous sommes attachés à évaluer la pertinence des arguments théoriques et empiriques avancés par la littérature sur la croissance pro-pauvre. Depuis les travaux fondateurs de Kuznets dans les années 50, la littérature économique a pendant longtemps soutenu que seule la promotion de la croissance économique suffit, à terme, pour favoriser la réduction de la pauvreté (grâce au trickle down notamment). Plus récemment, les théoriciens de la croissance pro-pauvre ont remis en cause cette vision. L'efficacité de la croissance comme vecteur de réduction de la pauvreté dépend en partie des inégalités de revenu. C'est pour cette raison que deux grandes approches se sont développées pour définir la croissance pro-pauvre : l'approche relative et l'approche absolue. Selon notre avis, la distinction entre ces deux approches est beaucoup moins claire qu'il n'y paraît. Dans la réalité, l'ampleur de la réduction de la pauvreté sous l'impulsion de la croissance dépend du niveau de départ de la distribution du revenu. De plus, l'objectif principal n'est pas tant celui d'une réduction du taux de pauvreté que celui de la promotion d'une distribution plus égalitaire. La préoccupation majeure est celle qui consiste à accroître l'influence de la croissance sur la réduction de la pauvreté.

Modéliser la croissance pro-pauvre nécessite de développer des méthodes d'estimations robustes au biais d'endogénéité et au biais de sélection. Ainsi, dans un second temps, cette thèse propose d'étendre la réflexion sur la régression Gini à l'économétrie des données de panels. Nous sommes d'abord revenus sur les fondamentaux de la régression Gini : la différence moyenne du Gini, la covariance au sens de Gini (co-Gini), la corrélation au sens de Gini (G-corrélation), le Gini généralisé et l'analyse au sens de Gini (l'ANOGini). Ensuite, nous avons montré qu'il était possible de décomposer la variabilité totale au sens de Gini en variabilité intragroupe et en intergroupe sans l'indice de chevauchement. Ainsi, nous avons pu définir les estimateurs intragroupes et intergroupes dans le cadre du Gini. Comme tout estimateur basé sur le Gini, nos estimateurs sont des \mathbb{U} -statistiques et donc convergents vers la loi normale. Les Tables 2.3, 2.4 et 2.5 illustrent de façon concrète la pertinence de nos estimateurs intragroupes et intergroupes au sens de Gini en présence de points aberrants. En outre, nos nouveaux estimateurs sont moins sensibles que les méthodes standards aux erreurs de me-

sure et aux formes fonctionnelles choisies. Cette dernière propriété est particulièrement importante dans l'étude de la croissance pro-pauvre. D'ailleurs, plusieurs travaux sont revenus sur les conséquences des choix de spécification et de la méthode d'estimation. Nous avons également proposé un test de l'existence des effets individuels.

Une autre contribution théorique de cette thèse réside dans l'introduction de nouvelles conditions afin de garantir "bon comportement" de la matrice des moments. Dans la littérature sur la régression Gini aucune hypothèse particulière n'a été posée sur cette matrice alors que nous savons, qu'à l'évidence, certaines conditions sont nécessaires. En effet, nous avons montré que lorsque les régresseurs sont comonotones il devient impossible d'appliquer la régression Gini puisque la matrice des moments n'est plus inversible. Nous avons également montré que l'absence de comonotonie implique l'hypothèse de rang plein et que la seconde condition de Grenander n'était pas indispensable dans le cas de la régression Gini. Ainsi, contrairement aux MCO qui imposent trois conditions (conditions de Grenander) pour garantir le bon comportement de la matrice des moments, la régression Gini n'impose que deux conditions.

Par ailleurs, d'après les travaux de Shelef (2014), le test de Dickey et Fuller fournit de faible puissance lorsque les données sont contaminées par des valeurs aberrantes. Afin de surmonter cette limite, elle propose un test basé sur le GMD. Nous avons montré que ce dernier test donne des estimateurs robustes et valides pour une large gamme de distributions. Toutefois, nos simulations montrent aussi que ce test Shelef (2014) a une faible puissance lorsque l'hypothèse d'absence d'autocorrélation des erreurs n'est plus satisfaite. En effet, nos résultats indiquent une diminution de la puissance de ce test au fur et à mesure que le coefficient de corrélation augmente. Afin de tenir de cette limite, nous avons introduit une extension Aïtken Gini. L'estimateur qui découle de cette dernière méthode nous permet d'appliquer le test de Shelef (2014) avec processus autocorrélé sans perte de puissance.

Une autre application de l'estimateur Aïtken Gini au modèle à effets aléatoires a été envisagée. Pour ce dernier, un certain nombre d'extensions pourraient être nécessaires. Il existe plusieurs possibilités pour estimer la matrice de variance-covariance de l'erreur. Yithzaki nous suggère d'utiliser la matrice de Gini-Cog pour l'estimer. Dans cette thèse, nous avons privilégié cette dernière méthode. Toute-

fois, nous reconnaissons que d'autres options (Swamy and Arora, 1972; Amemiya, 1971, par exemple) pourraient être envisagées. IL serait ainsi important d'analyser, par des simulations par exemple, l'impact du choix d'une méthode particulière sur les estimations obtenues.

Enfin, cette thèse présente des applications empiriques qui illustrent les limites des indices de croissance pro-pauvre et l'avantage de nos estimateurs basés sur le Gini. A l'issue d'une étude comparative des indices de croissance pro-pauvre, nous avons montré que les mesures traditionnelles de croissance pro-pauvre peuvent conduire à des résultats contradictoires. Ensuite, il montre que le processus de croissance favorise la réduction de la pauvreté à condition que les inégalités de revenu soient maîtrisées. L'analyse nous conduit à affirmer également que l'impact de la croissance agricole sur la réduction de la pauvreté varie en fonction du niveau de développement du pays. Dans les pays à faible revenu l'incidence de la pauvreté a tendance à être extrêmement sensible à la croissance du secteur agricole. Pour terminer, les pouvoirs publics qui cherchent à accroître la croissance du revenu des plus pauvres doivent s'appuyer sur un ensemble de mesures de nature à générer une croissance économique forte et soutenue. Cela passe nécessairement par une stabilité macro-économique, une faible inflation, des progrès massifs en scolarisation surtout dans le secondaire.

Bibliographie

- Adelman, I. and Morris, C. T. (1973). *Economic growth and social equity in developing countries*. Stanford University Press.
- Adelman, I., Robinson, S., et al. (1989). Income distribution and development. *Handbook of development economics*, 2(949) :1003.
- Aghion, P., Caroli, E., and Garcia-Penalosa, C. (1999). Inequality and economic growth : The perspective of the new growth theories. *Journal of Economic literature*, 37(4) :1615–1660.
- Ahluwalia, M. S. (1976). Inequality, poverty and development. *Journal of development economics*, 3(4) :307–342.
- Alesina, A. and Perotti, R. (1996). Income distribution, political instability, and investment. *European economic review*, 40(6) :1203–1228.
- Alesina, A. and Rodrik, D. (1994). Distributive politics and economic growth. *The Quarterly Journal of Economics*, 104(4) :465–490.
- Amemiya, T. (1971). The estimation of the variances in a variance-components model. *International Economic Review*, pages 1–13.
- Anand, S. and Kanbur, S. R. (1993a). Inequality and development a critique. *Journal of Development economics*, 41(1) :19–43.
- Anand, S. and Kanbur, S. R. (1993b). The kuznets process and the inequality-development relationship. *Journal of development economics*, 40(1) :25–52.
- Arellano, M. and Bond, S. (1991). Some tests of specification for panel data : Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2) :277–297.

- Arvesen, J. N. (1969). Jackknifing u-statistics. *The Annals of Mathematical Statistics*, 40(6) :2076–2100.
- Atkinson, A. and Brandolini, A. (1999). Promise and pitfalls in the use of secondary data-sets : A case study of oecd income inequality. *Ronéo, Nuffield College*.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory*, 2(3) :244–263.
- Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.
- Banerjee, A. V. and Duflo, E. (2003). Inequality and growth : What can the data say? *Journal of economic growth*, 8(3) :267–299.
- Banerjee, A. V. and Newman, A. F. (1993). Occupational choice and the process of development. *Journal of political economy*, pages 274–298.
- Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of economic growth*, 5(1) :5–32.
- Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104 :184–198.
- Becker, G., Murphy, K., and Tamura, R. (1990). Economic growth, human capital and population growth. *Journal of Political Economy*, 98(5) :S12–S137.
- Benabou, R. (1996). Inequality and growth. In *NBER Macroeconomics Annual 1996, Volume 11*, pages 11–92. MIT Press.
- Bhalla, S. S. (2002). *Imagine there's no country : Poverty, inequality, and growth in the era of globalization*. Peterson Institute.
- Boccanfuso, D. and Ménard, C. (2009). La croissance pro-pauvre : un aperçu exhaustif de la «boîte à outils». *Cahier de recherche. Working Paper*, pages 09–06.
- Bourguignon, F. (2003). The growth elasticity of poverty reduction : explaining heterogeneity across countries and time periods. *Inequality and growth : Theory and policy implications*, pages 3–26.
- Bourguignon, F. (2004). Le triangle pauvreté-croissance-inégalités. *Afrique contemporaine*, (3) :29–56.

- Carcea, M. and Serfling, R. (2014). A gini autocovariance function for heavy tailed time series modeling. Technical report, Citeseer.
- Carcea, M. and Serfling, R. (2015). A gini autocovariance function for time series modelling. *Journal of Time Series Analysis*, 36(6) :817–838.
- Chataîgner, J.-M. and Rafinot, M. (2005). La croissance pro-pauvres : Définitions et politiques. *Lettre des économistes de l'AFD. Paris Cedex. 9 : 2, 9.*
- Chen, S. and Ravallion, M. (2004). Welfare impacts of china's accession to the world trade organization. *The world bank economic review*, 18(1) :29–57.
- Cling, J.-P., De Vreyer, P., Razafindrakoto, M., and Roubaud, F. (2004). La croissance ne suffit pas pour réduire la pauvreté. *Revue française d'économie*, 18(3) :137–187.
- Cogneau, D., Guénard, C., et al. (2002). Les inégalités et la croissance : une relation introuvable. *Document de travail DT*, 3.
- Cord, L. (2005). Pro-poor growth. *présentation à la session de synthèse*, 3.
- Dagum, C. (1997). Decomposition and interpretation of gini and the generalized entropy inequality measures. *Statistica*, 57(3) :295–308.
- Datt, G. and Ravallion, M. (1992). Growth and redistribution components of changes in poverty measures : A decomposition with applications to brazil and india in the 1980s. *Journal of development economics*, 38(2) :275–295.
- Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *The World Bank Economic Review*, 10(3) :565–591.
- Deininger, K. and Squire, L. (1998). New ways of looking at old issues : inequality and growth. *Journal of development economics*, 57(2) :259–287.
- Denker, M. (1985). *Asymptotic distribution theory in nonparametric statistics*. Springer.
- Denuit, M. and Dhaene, J. (2003). Simple characterizations of comonotonicity and countermonotonicity by extremal correlations. *Belgian Actuarial Bulletin*, 3 :22–27.

- Dollar, D. and Kraay, A. (2002). Growth is good for the poor. *Journal of economic growth*, 7(3) :195–225.
- Duclos, J.-Y. (2009). What is "pro-poor" ? *Social choice and welfare*, 32(1) :37–58.
- Duclos, J.-Y. and Araar, A. (2006). Poverty and equity. *Measurement, Policy and Estimation with DAD», New York et Ottawa, Springer-IDRC-CRDI.*
- Duclos, J.-Y., Makdissi, P., and Araar, A. (2014). Pro-poor indirect tax reforms, with an application to Mexico. *International Tax and Public Finance*, 21(1) :87–118.
- Durbin, J. (1954). Errors in variables. *Revue de l'institut International de Statistique*, pages 23–32.
- Durlauf, S. N. (1994). Spillovers, stratification, and inequality. *European Economic Review*, 38(3-4) :836–845.
- Easterly, W. and Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of monetary economics*, 32(3) :417–458.
- Fan, S., Huong, P. L., and Long, T. Q. (2004a). Government spending and poverty reduction in Vietnam.
- Fan, S., Zhang, X., Rao, N., et al. (2004b). Public expenditure, growth, and poverty reduction in rural Uganda. Technical report, International Food Policy Research Institute (IFPRI).
- Fishlow, A. (1996). Inequality, poverty and growth where do we stand. In *OPS Document Reproduction Series*, number 65. World Bank.
- Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth. *American economic review*, pages 869–887.
- Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica : Journal of the Econometric Society*, pages 761–766.
- Frick, J. R., Goebel, J., Schechtman, E., Wagner, G. G., and Yitzhaki, S. (2006). Using analysis of gini (anogi) for detecting whether two subsamples represent the same universe the German socio-economic panel study (soep) experience. *Sociological Methods & Research*, 34(4) :427–468.

- Galbraith, J. K. (2009). Inequality, unemployment and growth : New measures for old controversies. *The Journal of Economic Inequality*, 7(2) :189–206.
- Galor, O. and Tsiddon, D. (1997). Technological progress, mobility, and economic growth. *The American Economic Review*, pages 363–382.
- Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The review of economic studies*, 60(1) :35–52.
- Ghura, D., Leite, C., and Tsangarides, C. G. (2002). *Is growth enough ? Macroeconomic policy and poverty reduction*, volume 2. International Monetary Fund.
- Grosse, M., Harttgen, K., and Klasen, S. (2008). Measuring pro-poor growth in non-income dimensions. *World Development*, 36(6) :1021–1047.
- Günther, I., Marouani, M. A., Raffinot, M., et al. (2006). La croissance est-elle propauvres au mali. *Notes et documents*.
- Hausman, J. A. and Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica : Journal of the Econometric Society*, pages 1377–1398.
- Hibbs, D. A. (1973). *Mass political violence : A cross-national causal analysis*, volume 253. Wiley New York.
- Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *The annals of mathematical statistics*, pages 293–325.
- Hsiao, C. (2003). Analysis of panel data (vol. 34). *Econometric Society Monographs*.
- Huber, P. J. (1980). *Robust statistics*. John Wiley and Sons.
- Jaeckel, L. A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *The Annals of Mathematical Statistics*, pages 1449–1458.
- Jain, S. et al. (1975). Size distribution of income ; a compilation of data.
- Jha, S. and Srinivasan, P. (2001). Taking the pds to the poor : Directions for further reform. *Economic and Political Weekly*, pages 3779–3786.
- Jureckova, J. (1971). Nonparametric estimate of regression coefficients. *The Annals of Mathematical Statistics*, pages 1328–1338.

- Ka, N. and Mussard, S. (2015). Book review of the gini methodology : A primer on a statistical methodology. *The Journal of Economic Inequality*, 13(2) :321–324.
- Ka, N. and Mussard, S. (2016). 11 regressions : Gini estimators for fixed effects panel data. *Journal of Applied Statistics*, 43(8) :1436–1446.
- Kakwani, N. (1980). *Income inequality and poverty*. Oxford University Press.
- Kakwani, N. (1993). Poverty and economic growth with application to côte d’ivoire. *Review of Income and Wealth*, 39(2) :121–139.
- Kakwani, N. (2000). On measuring growth and inequality components of poverty with application to thailand. *Journal of Quantitative Economics*, 16(1) :67–80.
- Kakwani, N., Khandker, S., and Son, H. (2002). Poverty equivalent growth rate : with applications to korea and thailand. economic commission for africa. Technical report, mimeo.
- Kakwani, N., Khandker, S., and Son, H. (2003). Poverty equivalent growth rate : with applications to korea and thailand. economic commission for africa. Technical report, mimeo.
- Kakwani, N., Khandker, S., and Son, H. (2004). Pro-poor growth : concepts and measurement with country case studies. brasília : International poverty centre. Technical report, PNUD, 2004 (Working Paper, n. 1).
- Kakwani, N. and Krongkaew, M. (2000). Analysing poverty in thailand. *Journal of the Asia Pacific Economy*, 5(1-2) :141–160.
- Kakwani, N., Pernia, E. M., et al. (2000). What is pro-poor growth ? *Asian development review*, 18(1) :1–16.
- Kakwani, N. and Son, H. (2002). Pro-poor growth and poverty reduction : the asian experience. *The Poverty Center, Office of Executive Secretary, ESCAP, Bangkok*.
- Kakwani, N. C. (1977). Applications of lorenz curves in economic analysis. *Econometrica : Journal of the Econometric Society*, pages 719–727.
- Klasen, S. (2004). In search of the holy grail : how to achieve pro-poor growth ? *Toward pro poor policies-aid, institutions, and globalization*, pages 63–94.

- Klasen, S. (2005). Economic growth and poverty reduction : measurement and policy issues.
- Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica : journal of the Econometric Society*, pages 33–50.
- Kraay, A. (2006). When is growth pro-poor ? evidence from a panel of countries. *Journal of development economics*, 80(1) :198–227.
- Krongkaew, M. (1996). The changing urban system in a fast-growing city and economy : the case of bangkok and thailand.
- Krongkeaw, M. (2000). Changes in poverty and income inequality in thailand, 1988 to 1998. *Bangkok, session C-pa 6b, Institute of East Asian Studies, Thammasat University*.
- Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, 45(1) :1–28.
- Lecaillon, J., Paukert, F., Morrisson, C., and Germidis, D. (1984). Income distribution and economic development ; an analytical survey.
- Lehmann, E. L. (1999). *Elements of large-sample theory*. Springer Science & Business Media.
- Lenth, R. V. (1983). Some properties of u statistics. *The American Statistician*, 37(4a) :311–313.
- Levine, R. and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American economic review*, pages 942–963.
- Li, H., Squire, L., and Zou, H.-f. (1998). Explaining international and intertemporal variations in income inequality. *The Economic Journal*, 108(446) :26–43.
- Lindert, P. H. (1996). What limits social spending ? *Explorations in Economic History*, 33(1) :1–34.
- Lopez, J. H. (2004). Pro-poor growth : a review of what we know (and of what we don't). *The world bank*.

- Mazumdar, D. and Son, H. H. (2002). Vulnerable groups and the labor market in thailand : impact of the asian financial crisis in the light of thailand's growth process. *Toronto : University of Toronto, April.*
- McCulloch, N. and Baulch, B. (1999). Tracking pro-poor growth. new ways to spot the biases and benefits. *Insights, (31) :3.*
- McKean, J. W. and Hettmansperger, T. P. (1976). Tests of hypotheses based on ranks in the general linear model. *Communications in statistics-theory and methods, 5(8) :693–709.*
- Mellor, J. W. (2000). *Faster more equitable growth : The relation between growth in agriculture and poverty reduction.* Harvard Institute for International Development.
- Mookherjee, D. and Shorrocks, A. (1982). A decomposition analysis of the trend in uk income inequality. *The Economic Journal, 92(368) :886–902.*
- Moyes, P. (2007). An extended gini approach to inequality measurement. *The Journal of Economic Inequality, 5(3) :279–303.*
- Mussard, S. and Richard, P. (2012). Linking yitzhaki's and dagum's gini decompositions. *Applied Economics, 44(23) :2997–3010.*
- OCDE (2001). Rapport annuel de l'ocde.
- Olkin, I. and Yitzhaki, S. (1992). Gini regression analysis. *International Statistical Review/Revue Internationale de Statistique, pages 185–196.*
- Osmani, S. et al. (2005). Defining pro-poor growth. Technical report.
- Papanek, G. F. and Kyn, O. (1986). The effect on income distribution of development, the growth rate and economic strategy. *Journal of Development Economics, 23(1) :55–65.*
- Paukert, F. (1973). Income distribution at different levels of development : A survey of evidence. *Int'l Lab. Rev., 108 :97.*
- Perotti, R. (1992). Income distribution, politics, and growth. *The American Economic Review, 82(2) :311–316.*

- Perotti, R. (1993). Political equilibrium, income distribution, and growth. *The Review of Economic Studies*, 60(4) :755–776.
- Perotti, R. (1994). Income distribution and investment. *European Economic Review*, 38(3) :827–835.
- Perotti, R. (1996). Growth, income distribution, and democracy : what the data say. *Journal of Economic growth*, 1(2) :149–187.
- Persson, T. and Tabellini, G. (1994). Is inequality harmful for growth? *The American Economic Review*, pages 600–621.
- Pyatt, G. (1976). On the interpretation and disaggregation of gini coefficients. *The Economic Journal*, 86(342) :243–255.
- Ravallion, M. (1995). Growth and poverty : evidence for developing countries in the 1980s. *Economics letters*, 48(3) :411–417.
- Ravallion, M. (1997). Can high-inequality developing countries escape absolute poverty? *Economics letters*, 56(1) :51–57.
- Ravallion, M. (2004). Pro-poor growth : A primer. *World Bank policy research working paper*, (3242).
- Ravallion, M. (2009). Définition de la croissance favorable aux pauvres : une réponse nanak kakwani. Technical report, International Policy Centre for Inclusive Growth.
- Ravallion, M. and Chen, S. (2003). Measuring pro-poor growth. *Economics letters*, 78(1) :93–99.
- Ravallion, M. and Datt, G. (1996). How important to india's poor is the sectoral composition of economic growth? *The World Bank Economic Review*, 10(1) :1–25.
- Ray, D. (1998). *Development economics*. Princeton University Press.
- Saint Paul, G. and Verdier, T. (1996). Inequality, redistribution and growth : A challenge to the conventional political economy approach. *European Economic Review*, 40(3) :719–728.

- Schechtman, E. and Schechtman, G. (2002). On characterization of two-sample u -statistics. *Statistics & probability letters*, 58(1) :53–59.
- Schechtman, E. and Yitzhaki, S. (1987). A measure of association based on gini's mean difference. *Communications in statistics-Theory and Methods*, 16(1) :207–231.
- Serfling, R. (2010). Fitting autoregressive models via yule-walker equations allowing heavy tail innovations. *Preprint. Available at www.utdallas.edu/~serfling*.
- Serfling, R. J. (2009). *Approximation theorems of mathematical statistics*, volume 162. John Wiley & Sons.
- Shalit, H. and Yitzhaki, S. (1984). Mean-gini, portfolio theory, and the pricing of risky assets. *The journal of Finance*, 39(5) :1449–1468.
- Shelef, A. (2014). A gini-based unit root test. *Computational Statistics & Data Analysis*.
- Shelef, A. and Schechtman, E. (2011). A gini-based methodology for analyzing time series.
- Silber, J. (1989). Factor components, population subgroups and the computation of the gini index of inequality. *The Review of Economics and Statistics*, pages 107–115.
- Son, H. H. (2004). A note on pro-poor growth. *Economics Letters*, 82(3) :307–314.
- Son, H. H. (2007). Pro-poor growth : concepts and measures.
- Stark, O., Taylor, J. E., and Yitzhaki, S. (1988). Migration, remittances and inequality : A sensitivity analysis using the extended gini index. *Journal of Development Economics*, 28(3) :309–322.
- Swamy, P. and Arora, S. S. (1972). The exact finite sample properties of the estimators of coefficients in the error components regression models. *Econometrica : Journal of the Econometric Society*, pages 261–275.
- Tarozzi, A. (2005). The indian public distribution system as provider of food security : Evidence from child nutrition in andhra pradesh. *European Economic Review*, 49(5) :1305–1330.

- Tibiletti, L. and Subramanian, S. (2015). Inequality aversion and the extended gini in the light of a two-person cake-sharing problem. *Journal of Human Development and Capabilities*, 16(2) :237–244.
- Venieris, Y. P. and Gupta, D. K. (1986). Income distribution and sociopolitical instability as determinants of savings : a cross-sectional model. *The Journal of Political Economy*, pages 873–883.
- Verardi, V. and Wagner, J. (2011). Robust estimation of linear fixed effects panel data models with an application to the exporter productivity premium. *Jahrbücher für Nationalökonomie und Statistik*, 231(4) :546–557.
- Watts, H. W. (1968). *An economic definition of poverty*. Institute for Research on Poverty.
- White, H. and Anderson, E. (2001). Growth versus distribution : does the pattern of growth matter ? *Development Policy Review*, 19(3) :267–289.
- Wiggins, S. and Higgins, K. (2008). Pro-poor growth and development. *Linking economic growth and poverty reduction. Overseas Development Institute Briefing Paper*, 33 :2014–2020.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of econometrics*, 126(1) :25–51.
- Wodon, Q., Yitzhaki, S., et al. (2003). Inequality and the accounting period. *Economics Bulletin*, 4(36) :1–8.
- Wooldridge, J. M. (2015). *Introductory econometrics : A modern approach*. Nelson Education.
- Yitzhaki, S. (1990). On the sensitivity of a regression coefficient to monotonic transformations. *Econometric Theory*, 6(02) :165–169.
- Yitzhaki, S. (1998). More than a dozen alternative ways of spelling gini. *Research in economic inequality*, 8.
- Yitzhaki, S. et al. (2003). Gini's mean difference : A superior measure of variability for non-normal distributions. *Metron*, 61(2) :285–316.

- Yitzhaki, S. and Lerman, R. I. (1991). Income stratification and income inequality. *Review of income and wealth*, 37(3) :313–329.
- Yitzhaki, S. and Olkin, I. (1991). Concentration indices and concentration curves. *Lecture Notes-Monograph Series*, pages 380–392.
- Yitzhaki, S. and Schechtman, E. (2004). The gini instrumental variable, or 'the double iv estimator'. *Metron*, LXII(3) :287–313.
- Yitzhaki, S. and Schechtman, E. (2013). *The Gini methodology : A primer on a statistical methodology*, volume 272. Springer Science & Business Media.

Annexes

3.3.1 Quelques \mathbb{U} -statistiques

TABLE 3.14 – Quelques \mathbb{U} -statistiques

| Mesures | Paramètre | m | Estimateur (\mathbb{U} -statistique) |
|------------------|---|---------|--|
| L'Espérance | $\mathbb{E}(\mathbf{X})$ | $m = 1$ | $\mathbb{U}_1 = \binom{n}{1}^{-1} \sum x_i = \frac{1}{n} \sum \mathbf{x}_i$ |
| GMD | $\mathbb{E} \mathbf{x}_i - \mathbf{x}_j $ | $m = 2$ | $\mathbb{U}_3 = \frac{2}{n(n-1)} \sum_{i < j} \mathbf{x}_i - \mathbf{x}_j $ |
| Variance | $\frac{1}{2} \mathbb{E}(\mathbf{x}_i - \mathbf{x}_j)^2$ | $m = 2$ | $\mathbb{U}_2 = \frac{1}{(n-1)} (\sum_{i=1}^n \mathbf{x}_i^2 - n\mathbb{U}_1^2)$ |
| Cog | $\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{y}))$ | $m = 2$ | $\mathbb{U}_4 = 4 \binom{n}{1} \sum_{i=1}^n (2i - 1 - n) \mathbf{x}_{y(i)}$ |
| G-corrélation | $\mathbf{G}_{\mathbf{x}, \mathbf{y}} = \frac{\text{Cov}(\mathbf{x}, \mathbf{F}_{\mathbf{y}}(\mathbf{y}))}{\text{Cov}(\mathbf{x}, \mathbf{F}_{\mathbf{x}}(\mathbf{x}))}$ | $m = 2$ | $\mathbb{U}_5 = \frac{\mathbb{U}_4}{\mathbb{U}_3}$ |
| L'indice de Gini | $\frac{GMD}{2\mathbb{E}(\mathbf{X})}$ | $m = 2$ | $\mathbb{U}_6 = \frac{\mathbb{U}_3}{2\mathbb{U}_1}$ |

3.3.2 Similarités et différences entre la régression Gini et les MCO

- (\mathbf{y}, \mathbf{x}) un couple de variables aléatoires représentatif de la distribution totale d'une population et $(\mathbf{y}_m, \mathbf{x}_m)$ avec $m=1, \dots, M$, M couples de variables représentant les M sous-groupes disjoints de la population totale.
- $(\mathbf{y}_{mj}, \mathbf{x}_{mj})$ $m = 1, \dots, M$ et $j = 1, \dots, n_m$ un couple de variables aléatoires représentatif du groupe m .
- $N = \sum_{m=1}^M n_m$
- $\hat{\beta}^{sp}$ l'estimateur sémi-paramétrique obtenu dans la population totale β_m le coefficient de régression obtenu dans le groupe m ,
- \mathbf{y}_{mj} la jème observation de \mathbf{y}_m , $\mathbf{y}_.$ la moyenne de \mathbf{y} et $\mathbf{y}_m.$ la moyenne de \mathbf{y}_m
- $\mathbf{F}_{\mathbf{x}}(\mathbf{x})$ la fonction cumulative de \mathbf{x} et $\mathbf{F}_.$ la moyenne des $\mathbf{F}_{\mathbf{x}}(\mathbf{x})$.
- $p_m = \frac{n_m}{N}$

Ainsi, l'estimateur de la régression Gini peut être réécrit :

$$\begin{aligned}\hat{\beta}^{sp} &= \frac{\text{Cov}(\mathbf{y}, \mathbf{F}(\mathbf{x}))}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} = \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{mj} - \mathbf{y}_{..}) (\mathbf{F}(\mathbf{x}) - \mathbf{F}_{..}(\mathbf{x})) \right\} \\ &= \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{mj} - \mathbf{y}_{m.} + \mathbf{y}_{m.} - \mathbf{y}_{..}) (\mathbf{F}(\mathbf{x})) \right\}\end{aligned}$$

Avec $\mathbf{F}_{\mathbf{x}}(\mathbf{x})$ la fonction cumulative de \mathbf{x} ;

Par conséquent,

$$\begin{aligned}\hat{\beta}^{sp} &= \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{mj} - \mathbf{y}_{m.}) (\mathbf{F}(\mathbf{x})) + \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{m.} - \mathbf{y}_{..}) (\mathbf{F}(\mathbf{x})) \right\} \\ &= \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{mj} - \mathbf{y}_{m.}) (\mathbf{F}(\mathbf{x}) - \mathbf{F}_{m.}(\mathbf{x})) \right. \\ &\quad \left. + \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{n_m} (\mathbf{y}_{m.} - \mathbf{y}_{..}) (\mathbf{F}_{\mathbf{x}}(\mathbf{x}) - \mathbf{F}_{m.}(\mathbf{x})) \right\}\end{aligned}\tag{3.10}$$

En posant, $p_m = \frac{n_m}{N}$, l'expression précédente devient :

$$\begin{aligned}\hat{\beta}^{sp} &= \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \sum_{m=1}^M p_m \text{Cov}_m(\mathbf{y}, \mathbf{F}(\mathbf{x})) + \text{Cov}_B(\mathbf{y}_{m.}, \mathbf{F}_{m.}(\mathbf{x})) \right\} \\ &= \frac{1}{\text{Cov}(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \left\{ \sum_{m=1}^M p_m \frac{\text{Cov}_m(\mathbf{y}, \mathbf{F}(\mathbf{x}))}{\text{Cov}_m(\mathbf{x}, \mathbf{F}(\mathbf{x}))} \text{Cov}_m(\mathbf{x}, \mathbf{F}_{m.}(\mathbf{x})) \frac{\text{Cov}_m(\mathbf{x}, \mathbf{F}(\mathbf{x}))}{\text{Cov}_m(\mathbf{x}, \mathbf{F}_{m.}(\mathbf{x}))} \right. \\ &\quad \left. + \text{Cov}_B(\mathbf{y}_{m.}, \mathbf{F}_{m.}(\mathbf{x})) \right\}.\end{aligned}\tag{3.11}$$

Soient :

$$\hat{\beta}_m^{sp} = \frac{\text{Cov}_m(\mathbf{y}, \mathbf{F}(\mathbf{x}))}{\text{Cov}_m(\mathbf{x}, \mathbf{F}(\mathbf{x}))}\tag{3.12}$$

l'estimateur sémi-paramétrique dans le groupe m . Et

$$O_m = \frac{\text{Cov}_m(\mathbf{x}, \mathbf{F}(\mathbf{x}))}{\text{Cov}_m(\mathbf{x}, \mathbf{F}_{m.}(\mathbf{x}))}\tag{3.13}$$

l'interaction entre le groupe m et la population totale.

On a :

$$\hat{\beta}^{sp} = \sum_{m=1}^M V_n \beta_m^{sp} + V_b \beta_b^{sp} \quad (3.14)$$

Avec $V_n = \frac{p_m O_m GMD^m(\mathbf{x})}{GMD(\mathbf{x})}$ est la part du groupe m dans la distribution totale. La première partie 3.14 peut être interprétée comme l'intra et la seconde partie l'inter.

La décomposition de l'estimateur des MCO est :

$$\beta = \sum_{m=1}^M W_n \beta_m + W_b \beta_b. \quad (3.15)$$

3.3.3 Décomposition de l'Aitken Gini

Soient J_T une matrice carrée (T,T) composée uniquement de 1 et \mathbf{x}_n le vecteur (T,1) des observations de \mathbf{x} pour la nième individu. Ainsi,

$$\mathbf{x} - \mathbf{x}_n. = \begin{pmatrix} \mathbf{x}_{n1} - \mathbf{x}_n. \\ \mathbf{x}_{n2} - \mathbf{x}_n. \\ \vdots \\ \mathbf{x}_{nT} - \mathbf{x}_n. \end{pmatrix} = \left(I_T - \frac{J_T}{T} \right) \mathbf{x}_n$$

Il s'en suit que

$$\mathbf{X} - \mathbf{X}_n. = \left[I_N \otimes \left(I_T - \frac{J_T}{T} \right) \right] \mathbf{X}_n$$

Ainsi, $\mathbb{W}_n = \left[I_N \otimes \left(I_T - \frac{J_T}{T} \right) \right]$ permet de calculer les écarts de chaque observation aux moyennes individuelles qu'on appelle habituellement "l'opérateur intra-individuel". Par analogie, nous pouvons déduire l'opérateur inter-individuelle \mathbb{B}_n donné par l'expression suivante :

$$\mathbb{B}_n = I_{NT} - \mathbb{W}_n = I_N \otimes \frac{J_T}{T}$$

En utilisant ces deux opérateurs nous pouvons réécrire les variabilités intra et inter :

$$\mathbf{G}_{xx}^{intra} = \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{nt} - \mathbf{x}_n.) [\mathbf{R}_{nt}(\mathbf{x}) - \mathbf{R}_n.(\mathbf{x})]' = \mathbf{R}(\mathbf{x}) \mathbb{W}_n \mathbf{x}$$

et

$$\mathbf{G}_{xx}^{inter} = \sum_{n=1}^N \sum_{t=1}^T (\mathbf{x}_{n.} - \mathbf{x}_{..}) [\mathbf{R}_{nt}(\mathbf{x}) - \mathbf{R}_{..}(\mathbf{x})]' = \mathbf{R}(\mathbf{x}) \mathbb{B}_n \mathbf{x}$$

Ainsi, sans perte de généralité, nous pouvons réécrire la matrice Ω_G :

$$\Omega_G = \mathbb{E}(\varepsilon \mathbf{R}(\varepsilon')) = \begin{pmatrix} \mathbb{E}(\varepsilon_1 \mathbf{R}(\varepsilon'_1)) & \mathbb{E}(\varepsilon_1 \mathbf{R}(\varepsilon'_2)) & \dots & \mathbb{E}(\varepsilon_1 \mathbf{R}(\varepsilon'_N)) \\ \mathbb{E}(\varepsilon_2 \mathbf{R}(\varepsilon'_1)) & \mathbb{E}(\varepsilon_2 \mathbf{R}(\varepsilon'_2)) & \dots & \mathbb{E}(\varepsilon_2 \mathbf{R}(\varepsilon'_N)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(\varepsilon_N \mathbf{R}(\varepsilon'_1)) & \mathbb{E}(\varepsilon_N \mathbf{R}(\varepsilon'_2)) & \dots & \mathbb{E}(\varepsilon_N \mathbf{R}(\varepsilon'_N)) \end{pmatrix} =$$

$$\begin{pmatrix} ccccg_o & 0 & 0 & 0 \\ 0 & g_o & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & g_o \end{pmatrix}$$

avec

$$g_o = \begin{pmatrix} g_u^2 + g_w^2 & g_u^2 & \dots & g_u^2 \\ g_u^2 & g_u^2 + g_w^2 & \dots & g_u^2 \\ \vdots & \vdots & \ddots & \vdots \\ g_u^2 & g_u^2 & \dots & g_u^2 + g_w^2 \end{pmatrix} \quad (3.16)$$

$$\mathbb{E}(\varepsilon \mathbf{R}(\varepsilon')) = I_N \otimes g_o = g_w^2 I_{NT} + g_u^2 (I_N \otimes J_T) \quad (3.17)$$

$$= g_w^2 \left[(I_{NT} - \mathbb{B}_n) + \left(\frac{T * g_u^2}{g_w^2} \right) \mathbb{B}_n + \mathbb{B}_n \right] \quad (3.18)$$

D'où :

$$\mathbb{E}(\varepsilon \varepsilon') = g_w^2 \left[\mathbb{W}_n + \left(\frac{g_w^2 + T * g_u^2}{g_w^2} \right) \mathbb{B}_n \right] \quad (3.19)$$

$$= g_w^2 \left[\mathbb{W}_n + \frac{1}{\phi} \mathbb{B}_n \right] \quad (3.20)$$

$$= g_w^2 \Omega, \quad (3.21)$$

avec $\phi = \frac{g_w^2}{g_w^2 + T * g_u^2}$ et $g_w^2 + T * g_u^2$ et g_w^2 les valeurs propres de la matrice Ω_G .

$$\Omega_G^{-1/2} = \frac{1}{\sqrt{g_w^2}} \left(\mathbb{W}_n + \phi^{-1/2} \mathbb{B}_n \right). \quad (3.22)$$

Par conséquent,

$$\begin{aligned}
\hat{\beta}_{AG} &= \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \boldsymbol{\Omega}_G^{-1/2} \mathbf{X} \right\}^{-1} \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \boldsymbol{\Omega}_G^{-1/2} \mathbf{y} \right\} \\
&= \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) (\mathbb{W}_n + \sqrt{\phi} \mathbb{B}_n) \mathbf{X} \right\}^{-1} \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) (\mathbb{W}_n + \sqrt{\phi} \mathbb{B}_n) \mathbf{y} \right\} \\
&= \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \mathbb{W}_n \mathbf{X} + \sqrt{\phi} \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \mathbb{B}_n \mathbf{X} \right\}^{-1} \left\{ \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \mathbb{W}_n \mathbf{y} + \sqrt{\phi} \mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \mathbb{B}_n \mathbf{y} \right\}.
\end{aligned} \tag{3.23}$$

Puisque $\boldsymbol{\Omega}_G$ est une matrice diagonale définie positive, $\boldsymbol{\Omega}_G = g_o * I_{NT}$ avec g_o définie positive, nous pouvons poser $\mathbf{R}(\boldsymbol{\Omega}_G^{-1/2} \mathbf{X}) \approx \mathbf{R}(\mathbf{X})$. Il en résulte que :

$$\hat{\beta}_{AG} = \left\{ \mathbf{R}(\mathbf{X}) \mathbb{W}_n \mathbf{X} + \sqrt{\phi} \mathbf{R}(\mathbf{X}) \mathbb{B}_n \mathbf{X} \right\}^{-1} \left\{ \mathbf{R}(\mathbf{X}) \mathbb{W}_n \mathbf{y} + \sqrt{\phi} \mathbf{R}(\mathbf{X}) \mathbb{B}_n \mathbf{y} \right\} \tag{3.24}$$

Ainsi, tout comme les MCG, l'estimateur Aïtken-Gini peut être considéré comme une combinaison optimale des estimateurs intra et intergroupes :

$$\beta_{AG} = \left[\mathbf{G}_{xx}^{intra} + \theta' \mathbf{G}_{xx}^{inter} \right]^{-1} \left[\mathbf{G}_{xy}^{intra} + \theta' \mathbf{G}_{xy}^{inter} \right], \tag{3.25}$$

avec $\theta' = \sqrt{\phi}$.

3.3.4 Estimation de notre modèle sur les différents quintiles MCO

TABLE 3.15 – Estimation de notre modèle sur les différents quintiles MCO

| Estimates → $\beta =$ | Total MCO | | | | |
|-------------------------------|-------------------|-----------------|-----------------|------------------|-----------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | 0.037 (0.042) | 0.044 0.085 | 0.036 0.071 | 0.022 0.051 | -0.017 0.021 |
| Edu secon | 0.021 (0.011) | 0.094 0.047 | -0.066 0.033 | 0.041 0.021 | -0.019 0.010 |
| Taux d'inflation | -0.016 (0.010) | -0.022 0.011 | 0.081 0.021 | -0.0400 0.015 | -0.018 0.009 |
| DP de santé | 0.011 (0.027) | -0.057 0.041 | 0.056 0.045 | 0.072 0.065 | 0.086 0.079 |
| Indice de Gini | -0.224 (0.012) | -0.261 0.009 | -0.192 0.013 | -0.063 0.021 | 0.075 0.030 |
| Log(PIBa) | 0.086 (0.019) | 0.081 0.022 | 0.196 0.031 | 0.219 0.014 | -0.045 0.012 |
| Taux de croissance | 0.098 (0.017) | -0.172 0.021 | -0.228 0.014 | -0.125 0.021 | 0.079 0.018 |
| (1-Gini)*Txc 10 ⁻² | 0.070 (0.019) | -0.025 0.010 | -0.014 0.008 | 0.016 0.007 | 0.015 0.004 |
| DP du Gov 10 ⁻² | 0.055 (0.018) | -0.067 0.021 | -0.030 0.025 | 0.061 0.047 | 0.023 0.064 |
| R ² | 0.889 | 0.913 | 0.891 | 0.921 | 0.899 |

TABLE 3.16 – Estimation de notre modèle sur les différents quintiles MCO

| Estimates → $\beta =$ | Inter Individuel (MCO) | | | | |
|--------------------------|------------------------|-------------------|----------------------|-------------------|-------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | 0.05 (0.051) | 0.041 (0.071) | 0.032 (0.081) | 0.094 (0.055) | -0.018 (0.028) |
| Edu secon | 0.048 (0.010) | 0.0483 (0.026) | 0.053 (0.036) | 0.061 (0.022) | 0.031 (0.017) |
| Taux d'inflation | -0.056 (0.035) | -0.055 (0.013) | -0.061 (0.027) | -0.094 (0.019) | -0.095 (0.011) |
| DP de santé 10^{-2} | 0.035 (0.073) | 0.351 (0.07) | 0.460 (0.05) | 0.4820 (0.008) | 0.271 (0.091) |
| Indice de Gini | -0.065 (0.029) | 0.065 (0.020) | 0.143 (0.018) | 0.2816 (0.031) | 1.049 0.039 |
| Log(PIBa) 10^{-2} | 0.182 (0.053) | -0.183 (0.028) | -0.148203 (0.036) | -0.034 (0.021) | -0.484 (0.023) |
| Taux de croissance | 0.181 (0.020) | 2.426 (0.026) | 2.819 (0.017) | 3.0314 (0.018) | 2.199 (0.019) |
| (1-Gini)*Txc 10^{-2} | 0.070 (0.019) | 0.056 (0.013) | 0.066 (0.009) | 0.072 (0.014) | 0.043 (0.020) |
| DP du Gov 10^{-2} | 0.099 (0.06) | 0.01 (0.031) | 0.0159 (0.021) | 0.017 (0.04) | 0.089 (0.069) |
| R ² | 0.871 | 0.864 | 0.823 | 0.862 | 0.896 |

TABLE 3.17 – Estimation de notre modèle sur les différents quintiles Gini

| Estimates → $\beta =$ | Total (Gini) | | | | |
|-------------------------------|--------------------|------------------|-----------------|-----------------|-----------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | 0.011 (0.085) | 0.035 0.067 | 0.030 0.048 | 0.039 0.037 | -0.027 0.022 |
| Edu secon | 0.091 (0.009) | 0.067 0.001 | 0.026 0.0014 | 0.038 0.0011 | 0.019 0.0061 |
| Taux d'inflation | -0.097 (0.0012) | -0.026 0.0014 | -0.048 0.007 | -0.050 0.008 | -0.046 0.002 |
| DP de santé | 0.120 (0.911) | -0.120 0.081 | -0.065 0.041 | -0.011 0.006 | 0.046 0.013 |
| Indice de Gini | -0.130 (0.012) | 0.11 0.017 | 0.085 0.014 | 0.100 0.025 | 0.046 0.031 |
| Log(PIBa) | 0.117 (0.006) | 0.010 0.009 | 0.075 0.011 | -0.063 0.008 | 0.047 0.0027 |
| Taux de croissance | 0.086 (0.017) | 0.171 0.0025 | 0.092 0.0018 | 0.011 0.012 | 0.007 0.0024 |
| (1-Gini)*Txc 10 ⁻² | -0.040 (0.001) | 0.025 0.002 | 0.012 0.003 | 0.029 0.002 | 0.003 0.005 |
| DP du Gov 10 ⁻² | 0.014 (0.05) | 0.014 0.11 | 0.039 0.038 | -0.018 0.054 | -0.089 0.101 |
| GR ² | 0.750 | 0.812 | 0.781 | 0.834 | 0.769 |

TABLE 3.18 – Estimation de notre modèle sur les différents quintiles Gini

| Estimates → $\beta =$ | Inter Gini | | | | |
|--------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | Q1 | Q2 | Q3 | Q4 | Q5 |
| Edu prim | 0.012 (0.090) | 0.037 (0.074) | 0.032 (0.051) | -0.042 (0.042) | -0.027 (0.028) |
| Edu secon | 0.094 (0.011) | 0.068 (0.003) | 0.057 (0.002) | 0.063 (0.003) | 0.115 (0.007) |
| Taux d'inflation | -0.099 (0.003) | -0.012 (0.002) | -0.051 (0.008) | 0.049 (0.01) | 0.045 (0.004) |
| DP de santé 10^{-2} | 0.136 (0.089) | -0.123 (0.091) | 0.068 (0.051) | 0.012 (0.007) | 0.021 (0.017) |
| Indice de Gini | -0.22 (0.014) | -0.14 (0.022) | -0.089 (0.025) | -0.109 (0.031) | 0.049 (0.040) |
| Log(PIBa) 10^{-2} | 0.119 (0.009) | 0.012 (0.01) | 0.1801 (0.013) | -0.065 (0.009) | -0.501 (0.003) |
| Taux de croissance | 0.095 (0.01ç) | 0.112 (0.003) | 0.095 (0.002) | 0.013 (0.017) | 0.009 (0.003) |
| (1-Gini)*Txc 10^{-2} | 0.045 (0.002) | 0.029 (0.003) | 0.017 (0.003) | 0.032 (0.004) | 0.003 (0.006) |
| DP du Gov 10^{-2} | 0.016 (0.06) | 0.016 (0.013) | 0.041 (0.043) | -0.022 (0.061) | -0.090 (0.112) |
| GR ² | 0.749 | 0.784 | 0.801 | 0.823 | 0.811 |

Université de Montpellier

Faculté d'Économie

Résumé : Cette thèse contribue à l'approche économétrique de la croissance pro-pauvre. Elle présente des apports théoriques et empiriques. En premier lieu, elle présente les différentes définitions, indices et politiques de croissance pro-pauvre proposées dans la littérature théorique. Elle examine également les modèles théoriques et empiriques portant sur les interactions entre distribution du revenu et croissance. Elle montre que les mesures traditionnelles, en plus de leurs caractères partiels, peuvent conduire à des résultats contradictoires. Pour contourner ces limites, cette thèse privilégie l'approche alternative qui consiste à utiliser des modèles économétriques. Cette dernière approche, bien qu'elle présente l'avantage d'inclure l'ensemble des dimensions de la pauvreté, souffre de deux types de biais : le biais de sélection et le biais d'endogénéité. Ces derniers s'expliquent par les limitations inhérentes aux données : erreurs de mesures, points aberrants. En outre, les résultats obtenus avec cette approche sont sensibles aux formes fonctionnelles choisies. Ainsi, il y a de bonnes raisons d'utiliser la régression Gini. Cependant, les régressions de type Gini n'existaient qu'en coupe instantanée et en séries temporelles. Ainsi, dans un second temps, cette thèse propose d'étendre la réflexion sur la régression Gini en panel. Elle introduit les estimateurs intragroupes, intergroupes, le test d'existence de l'effet individuel et l'estimateur Aitken Gini. Enfin, cette thèse présente des applications empiriques qui illustrent de façon concrète la robustesse de nos estimateurs. Elle s'intéresse particulièrement aux conséquences de la méthode d'estimation et à la section de l'échantillon. Elle conclut que le processus de croissance favorise la réduction de la pauvreté à condition que les inégalités de revenu soient maîtrisées. Mais aussi, que l'impact de la croissance agricole sur la réduction de la pauvreté varie en fonction du niveau de développement du pays.

Abstract : This thesis contributes to the econometric approach of pro-poor growth. It presents theoretical and empirical contributions. First, it brings out the different definitions, indices and the policies of pro-poor growth proposed in the theoretical literature. It also examines the theoretical and empirical interactions between income distribution and growth. It shows that the traditional measures, in addition to their partial characteristics, can lead to contradictory results. To avoid these limits this thesis lays the emphasis on the alternative approach by using econometric models. The latter approach, although it has the advantage of including all the dimensions of poverty, suffers from two types of bias : selection bias and bias of endogeneity. These are due to the limitations of the data : measurement errors, outliers. In addition, the results obtained with this approach are sensitive to the selected functional forms. So, there are good reasons to use the Gini regression. However, the Gini regressions existed only for cross sectional data and time series. Thus, in a second time, this thesis proposes to extend the Gini regression for panel data. It introduces within and between-group estimators, the individual effect test and the Gini Aitken estimator. Finally, this thesis presents empirical applications that illustrate the robustness of our estimators. It is particularly interested in the consequences of the estimation method and the sample section. It concludes that the growth process promotes poverty reduction when income inequalities are overcome. But also, the impact of agricultural growth on poverty reduction varies depending on the country's level of development.

LAMETA, Avenue Raymond Dugrand, 34960 Montpellier