

# Une réduction de l'inégalité des chances dans l'obtention du revenu en France ? \*

Arnaud LEFRANC<sup>†</sup>    Nicolas PISTOLESI<sup>‡</sup>    Alain TRANNOY<sup>§</sup>

Janvier 2004

## Résumé

Nous appréhendons l'égalité des chances de revenu comme une situation où il n'est pas possible de classer les distributions conditionnelles de revenu du fils par rapport à la CSP du père selon les critères de dominance stochastique d'ordre 1 ou 2. Les données utilisées proviennent des 5 enquêtes BDF réalisées entre 1979 à 2000. La robustesse statistique des conclusions est appréciée au travers d'instruments d'inférence statistique. L'inégalité des chances de revenu primaire entre fils de catégories de salariés régresse, les fils de cadres perdant du terrain par rapport à toutes les autres catégories cependant que les fils de non-salariés améliorent leurs positions. L'espérance de revenu disponible d'un fils de cadre n'est que de 50% supérieure à celle d'un fils d'ouvrier. La politique redistributive ne corrige qu'assez modestement l'inégalité des chances. En première approximation, l'origine socioprofessionnelle joue d'une façon multiplicative : le revenu d'un individu est décrit par le produit d'une variable aléatoire - distribuée indépendamment de la CSP du père - et d'un coefficient qui dépend de l'origine familiale et qui majore au plus de 50% le niveau de revenu.

## 1 Introduction

L'analyse empirique des inégalités s'est longtemps focalisée sur la mesure des inégalités inter-individuelles de revenus ou de conditions de vie et sur l'étude de leurs évolutions dans le temps. Cette façon de procéder, compréhensible au regard des données généralement disponibles, n'est cependant pas totalement en phase avec les recommandations de certains philosophes de la responsabilité, tels que Dworkin (1981), Arneson (1989), Cohen (1989) ou encore Roemer (1998). A divers titres, ces auteurs ont notamment souligné le manque

---

\*Cet article a été réalisé dans le cadre d'une étude financée par le Commissariat au Plan. Nous remercions Gérard Forgeot de l'INSEE pour la mise à disposition des données de la vague 2000 de l'enquête BdF, ainsi que François Bourguignon, Russell Davidson et Jean-Yves Duclos pour leurs remarques sur une version précédente de ce texte. Nous restons évidemment seuls responsables des éventuelles erreurs ou omissions.

<sup>†</sup>THEMA, Université de Cergy-Pontoise 33 bd du Port 95011 Cergy-Cedex. Email : arnaud.lefranc@eco.u-cergy.fr

<sup>‡</sup>THEMA, Université de Cergy-Pontoise. Email : nicolas.pistolesi@eco.u-cergy.fr

<sup>§</sup>EHESS, GREQAM-IDEP. Email : alain.trannoy@ehess.cnrs-mrs.fr

de pertinence éthique d'une analyse qui ne s'intéresserait qu'à l'inégalité des résultats et des situations individuelles, en ignorant le rôle de la responsabilité individuelle dans l'hétérogénéité des situations observées. Or, de ce point de vue, force est de constater que la formation du revenu individuel met en jeu un vaste ensemble de déterminants de nature très diverse : d'une part, les préférences et l'effort individuels contribuent à l'émergence des inégalités de revenu entre individus ; d'autre part, différents facteurs, que nous regroupons sous le terme de circonstances, et qui incluent notamment l'origine familiale, le talent intrinsèque des individus, ou encore le facteur chance, expliquent aussi pour partie les disparités de revenu. Au regard de la philosophie de la responsabilité, ces différents facteurs doivent alors être distingués, dans la mesure où les premiers relèvent de l'exercice de la responsabilité individuelle, alors que les seconds n'en dépendent pas. Les inégalités de revenu provenant des premiers facteurs peuvent en effet être jugées équitables, car elles sont le produit de la liberté de choix (au sens large) des individus. Au contraire, les inégalités de résultats dues à des différences de circonstances ne sont pas éthiquement acceptables.

Cet éclairage permet alors d'apporter des fondements éthiques à l'adoption d'une politique de redistribution ayant pour but de compenser les inégalités dues aux circonstances, et d'assurer ainsi l'égalité des chances dans la formation du revenu. En pratique, la mise en œuvre d'une telle politique soulève cependant plusieurs problèmes. Tout d'abord, comme pour toute politique redistributive, le degré de compensation souhaitable devra tenir compte de contraintes d'efficacité économique (Roemer *et alii* (2003)). D'autre part, la mise en œuvre d'une politique d'égalité des chances pose d'importants problèmes d'information liés à l'observabilité des déterminants du revenu. Ainsi, les contributions du facteur chance ou du (manque de) talent intrinsèque ne sont pas observables : compenser ces circonstances à l'aide d'une politique redistributive d'égalité des chances peut alors s'avérer difficile.

De même, préalablement à toute action publique redistributive, l'étude empirique de l'égalité des chances à laquelle ce travail est consacré pose des problèmes informationnels de nature similaire. Ceux-ci sont mêmes accrus par la pauvreté de la description du milieu social d'origine des individus dans les données d'enquête. Seuls sont en général connus la

catégorie sociale de leurs parents, parfois leur niveau de diplôme, et leur lieu de naissance. Aussi, établir une distinction entre les éléments dont les individus sont responsables, et ceux qui échappent à leur contrôle n'est pas une chose aisée. Dans cet article, nous examinons dans quelle mesure l'origine familiale repérée par la catégorie socio-professionnelle du père <sup>1</sup> exerce un effet significatif sur les niveaux de vie obtenus par les enfants. Notre étude se concentre sur le devenir des fils <sup>2</sup> mais la même méthodologie pourrait s'appliquer à celui des filles. Dans cette perspective, toute différence de revenu associée à des différences d'origine sociale sera alors interprétée comme un symptôme d'inégalité des chances.

Cette assimilation ne va évidemment pas de soi. Il est en effet vraisemblable que l'origine familiale est empiriquement associée à d'autres déterminants du revenu individuel. En ne tenant pas compte explicitement de ces autres facteurs, on attribuera donc une partie des inégalités de revenu à des différences d'origine familiale, alors même qu'elles relèvent d'autres facteurs. Du point de vue de la mesure de l'inégalité des chances, cette ambiguïté peut apparaître plus ou moins gênante. Si les facteurs omis et corrélés à l'origine familiale ne relèvent pas, non plus, de l'exercice de la responsabilité individuelle, l'ambiguïté est sans conséquence : du point de vue de la théorie de l'équité-responsabilité, la présence d'une relation entre origine familiale et revenu proviendra bien d'une inégalité des chances. Si au contraire ces facteurs omis relèvent de l'exercice de la responsabilité individuelle, alors notre analyse conduirait à une évaluation biaisée de l'inégalité des chances. Tel serait par exemple le cas si les individus provenant d'un milieu "favorisé" exercent aussi un niveau d'effort individuel plus élevé. Notons que cette ambiguïté ferait alors écho à une difficulté de la théorie de l'équité-responsabilité dans un cadre déterministe : Fleurbaey (1995) et Bossert (1995) ont en effet montré qu'il n'était pas possible à la fois de compenser l'effet des facteurs dont les individus ne sont pas responsables et de respecter les conséquences de préférences ou d'effort différents, à moins que la fonction qui conduit des facteurs au revenu ne soit additivement séparable dans les deux groupes de facteurs. Nous ne cherchons toutefois pas à résoudre cette difficulté dans cet article et l'analyse empirique entreprise ici aura donc, du point de vue de l'étude de l'inégalité des chances un domaine de

---

<sup>1</sup>Pour des raisons de taille d'échantillon, nous ne prenons pas en compte d'autres caractéristiques du milieu familial d'origine. Pour une prise en compte du lieu de naissance en plus de l'origine sociale, on se reportera à Goux et Maurin (2003).

<sup>2</sup>Voir Thélot (1982) pour une des premières études sur l'influence du père sur les destinées du fils.

validité circonscrit par l'hypothèse maintenue que les préférences et l'effort individuels sont indépendants du milieu familial<sup>3</sup>. Ces difficultés méthodologiques se rencontrent également lors des analyses d'égalité des chances scolaires, même si elles sont rarement évoquées<sup>4</sup>.

Une fois posées les limites de l'exercice, il convient de se doter d'un instrument statistique permettant de mesurer l'incidence de l'origine familiale sur le niveau de revenu individuel. Une première option proposée par Van de Gaer (1993) consiste à comparer la moyenne des distributions individuelles de revenu, conditionnelles à différentes origines familiales : on conclurait alors à l'inégalité des chances en cas de différences dans les moyennes conditionnelles. Ce critère, fondé sur la seule comparaison des tendances centrales de distributions conditionnelles, conduirait cependant à ignorer des phénomènes intéressants du point de vue de l'inégalité des chances. Par exemple, il se pourrait que les moyennes conditionnelles de deux groupes soient identiques, alors même que les rapports des valeurs du premier décile conditionnel et du dernier décile conditionnel pointent dans des directions opposées. La traduction économique d'une telle observation serait alors que l'un des deux milieux sociaux serait plus favorable pour les enfants qui échouent relativement (ceux du premier décile), alors que la conclusion inverse prévaudrait pour les enfants qui réussissent particulièrement bien (ceux du dernier décile). On voit alors que l'égalité des moyennes conditionnelles ne suffit pas à garantir l'égalité des chances. L'analyse de l'égalité des chances implique au contraire de reconnaître que le milieu d'origine exerce potentiellement un effet sur l'ensemble de la distribution de revenu offerte aux individus. Dans ce cas, le fait de naître dans un milieu familial et social donné peut être assimilé au tirage d'une loterie de revenu particulière, certaines loteries étant plus risquées ou plus performantes que d'autres.

Notre étude de l'égalité des chances repose donc sur la comparaison de ces loteries, à l'aide des instruments de dominance stochastique, en prêtant une attention particulière aux problèmes d'inférence statistique. Nous retenons la démarche suivante : si deux loteries de revenu conditionnelles à des milieux d'origine distincts ne peuvent être classées à l'aide

---

<sup>3</sup>Notons toutefois que si on privilégie le respect du principe de compensation sur le principe de responsabilité, la correction de l'effet d'une corrélation positive entre milieu social d'origine et effort serait alors exigée.

<sup>4</sup>Pour un exemple d'étude de l'égalité des chances devant l'enseignement voir Goux et Maurin (1995).

des instruments usuels de dominance stochastique<sup>5</sup>, alors nous dirons qu'il y a égalité des chances entre ces deux milieux<sup>6</sup>. Cette méthodologie nous permet d'une part de proposer une définition de l'égalité des chances qui s'appuie sur les instruments les plus solides de la mesure de risque. Cette définition présente d'autre part l'avantage d'ouvrir naturellement vers les instruments classiques de l'analyse des inégalités comme la courbe de Lorenz ou la courbe de Lorenz généralisée. Toutefois, il faut être conscient que cette définition, contrairement à celle adoptée par Roemer (1998) et (2003), conduit à un classement en général incomplet des distributions. C'est le prix à payer pour ne pas faire dépendre le classement de l'acceptation d'axiomes plus spécifiques.

L'analyse des relations de dominance à partir de données d'enquête suppose de recourir à des techniques d'inférence statistique particulières qui relèvent d'une littérature théorique récente et en rapide évolution. En la matière, il est possible de mettre en œuvre des tests reposant sur l'estimation paramétrique de distributions de revenu<sup>7</sup> mais la robustesse des conclusions obtenues est alors fortement limitée par le choix de telle ou telle forme paramétrique particulière dans la modélisation de la fonction de distribution. Notre méthodologie statistique repose quant à elle sur la mise en œuvre de tests non-paramétriques de dominance stochastique. Elle prend appui sur des résultats récents dus à Davidson et Duclos (2000) qui unifient les procédures de tests de dominance (pour n'importe quel ordre de dominance) et permettent, par l'absence d'hypothèse paramétrique particulière, d'aboutir à des conclusions statistiques plus solides.

Les études portant sur l'inégalité de résultats retiennent comme indicateur de réussite soit le revenu du ménage, soit le salaire de l'individu. Le même dilemme se présente pour une mesure de l'égalité des chances. Dans cet article, nous évaluons l'égalité des chances en termes de niveau de vie, le niveau de vie étant apprécié par un revenu par unité de consommation. L'alternative, qui aurait été d'étudier l'inégalité des chances de salaire, mesure évidemment de plus près la réussite de l'individu sur le marché du travail, alors

---

<sup>5</sup>Ces instruments supposent que le décideur obéit aux axiomes de choix dans un univers risqué posés par Von Neuman et Morgenstern. Ces axiomes sont critiqués pour ne pas traduire correctement le comportement en risque mais leur valeur normative est rarement discutée. Dans la mesure où notre définition de l'égalité des chances veut s'appuyer sur des fondements normatifs, nous sommes fondés à les mobiliser.

<sup>6</sup>Une autre façon de procéder aurait consisté à pratiquer des régressions quantiles mais un tel exercice suppose de disposer d'échantillons suffisamment fournis pour un certain nombre de déciles. Malheureusement, nos données ne sont pas suffisamment riches pour retenir une telle méthodologie.

<sup>7</sup>Voir notamment McDonald et Xu (1995).

que le revenu du ménage enregistre également la réussite du conjoint. Le revenu du ménage permet donc de mesurer la réussite économique globale de l'individu qu'elle passe par le marché du travail ou par le choix du conjoint. Comme nous nous focalisons dans cette étude sur la filiation père-fils, le revenu dépend en moyenne un peu plus du travail du fils que du travail du conjoint. La seconde option présente en outre l'inconvénient de restreindre le champ aux salariés. Or les données font état d'une progression relative de la situation des non-salariés depuis 20 ans, un phénomène largement passé inaperçu et qu'il n'est pas souhaitable d'occulter. S'intéresser à l'égalité des chances de revenu se justifie également par l'importance que revêt le revenu dans l'appréciation du bien-être dans nos sociétés. Toutefois, la complexité de la formation du revenu, qui telle une boîte noire mêle l'effet de l'éducation, de la qualification, du chômage et de la constitution des couples, pour ne citer que quelques uns des facteurs principaux, rend objectivement l'analyse beaucoup plus délicate que si nous nous étions restreints à l'analyse de l'égalité des chances de salaires. Le but ultime serait d'arriver à une analyse de décomposition de l'inégalité des chances, permettant d'apprécier l'impact de chaque facteur (voir pour une étude sur le Brésil, Bourguignon, Ferreira et Menendez (2003)). Notre objectif est ici plus modeste et il s'agit simplement d'essayer d'établir un constat qui peut stimuler le désir de recourir à des analyses de décomposition dans des études ultérieures.

Nous mettons l'accent sur deux étapes essentielles de la formation du revenu disponible. D'une part, nous examinons l'égalité des chances dans la formation du revenu primaire. D'autre part, nous étudions dans quelle mesure l'impôt sur le revenu et la politique de transferts redistributifs ont été de nature à corriger l'inégalité des chances dans la formation du revenu primaire. Notre analyse exploite les cinq vagues disponibles des enquêtes Budget des Familles (BdF) de l'INSEE, de 1979 à 2000. Ces enquêtes présentent l'avantage de fournir une information particulièrement détaillée sur l'ensemble des revenus des ménages (revenus du travail, revenus du patrimoine et du capital, revenus de transferts). Nous nous attachons principalement aux comparaisons en coupe de distributions de revenu conditionnelles à l'origine familiale, repérée par la catégorie socio-professionnelle (CSP) du père. Les comparaisons dans le temps de distributions conditionnelles à une CSP d'origine donnée, par exemple les distributions conditionnelles de fils de cadres sur les cinq

vagues, doivent quant à elles être effectuées avec prudence, car la composition du groupe de conditionnement est susceptible de se modifier au cours du temps.

L'article est organisé comme suit. La première section procède tout d'abord à une définition formelle, en termes de dominance stochastique, du concept d'égalité des chances et décrit la méthodologie d'inférence statistique mise en œuvre. La deuxième section présente les données utilisées dans l'analyse empirique. Les principaux résultats sont présentés dans la troisième section.

## 2 Égalité des chances : définition et inférence

Nous proposons d'abord une définition de l'égalité des chances inspirée de Roemer (1998) et (2003) avant d'indiquer la procédure d'inférence statistique mise en œuvre dans le reste de l'article. La définition s'écarte des propositions existantes (Roemer (1998), Van de Gaer (1993), Goux et Maurin (2003)<sup>8</sup>) en ce qu'elle ne cherche pas à établir un classement complet des distributions de revenu.

### 2.1 Une définition à l'aide des concepts de dominance stochastique

Nous nous intéressons à la distribution du revenu individuel, que nous notons  $x$ , une variable aléatoire continue sur un support positif. Parmi les aléas affectant  $x$ , figure le milieu d'origine mesuré, par exemple par la CSP du père. Le milieu d'origine est représenté par une variable aléatoire discrète<sup>9</sup> notée  $s$ ,  $s \in S = \{1, \dots, \bar{s}\}$ . La fonction de répartition conditionnelle de  $x$  sachant  $s$  est notée  $F(x | s)$ . Une "distribution de chances" l'année  $t$  est donc repérée par la donnée de  $\bar{s}$  fonctions de répartition conditionnelles, soit  $\{F_t(x | s); s = 1, \dots, \bar{s}\}$ .

Nous rappelons ici les définitions usuelles de la dominance stochastique à l'ordre 1 et à l'ordre 2.

**Définition** Soient  $F$  et  $G$  deux fonctions de répartition définies sur un support positif.

---

<sup>8</sup>Voir Ooghe *et alii* (2003) pour une axiomatisation des solutions proposées par Roemer et Van de Gaer.

<sup>9</sup>Poser  $s$  multidimensionnelle n'engendre pas de difficulté supplémentaire.

La distribution  $F$  domine stochastiquement à l'ordre 1 la distribution  $G$  si :

$$F(x) \succeq_{DS_1} G(x) \Leftrightarrow F(x) \leq G(x), \forall x \in \mathbb{R}_+$$

La dominance stricte requiert, de surcroît, qu'il existe au moins une valeur de  $x$  telle que  $F(x) < G(x)$ . En situation de choix en univers risqué, la distribution dominante est préférée à la distribution dominée par tout individu dont la fonction d'utilité est croissante en  $x$ . Ce critère est un critère partiel qui ne permet pas de décider dans tous les cas de figure. Dans le cas où les fonctions de répartition se coupent, on peut recourir au critère de dominance stochastique d'ordre 2.

**Définition** La distribution  $F$  domine stochastiquement à l'ordre 2 la distribution  $G$  si :

$$F(x) \succeq_{DS_2} G(x) \Leftrightarrow \int_0^x F(y)dy \leq \int_0^x G(y)dy, \forall x \in \mathbb{R}_+.$$

La distribution dominante est préférée à la distribution dominée par tout individu averse au risque et ayant une utilité marginale du revenu positive (sa fonction d'utilité est croissante et concave en  $x$ ). Il a été établi par Shorrocks (1983) que le test de dominance d'ordre 2 est équivalent à la comparaison des courbes de Lorenz généralisées soit plus précisément

$$F(x) \succeq_{DS_2} G(x) \Leftrightarrow GL_F(p) \geq GL_G(p), \forall p \in [0, 1],$$

en notant  $GL_F(p)$ , l'ordonnée de la courbe de Lorenz généralisée en  $p$  pour la distribution  $F$ .

En considérant la situation hypothétique d'un individu à qui serait offerte la liberté de choisir son milieu d'origine, la meilleure chose que peut faire un individu au début de sa vie est de comparer les distributions de revenu conditionnelles au milieu d'origine. S'il n'a pas d'aversion au risque, il utilisera le test de dominance stochastique d'ordre 1. Ce test détermine un quasi-ordre sur  $S$ . On peut alors définir l'ensemble des milieux d'origine



non-dominés pour la relation binaire DS1 l'année  $t$ . Il est formellement donné par :

$$P_{1t} = \{s \in S \mid \nexists s' \in S \text{ avec } F_t(x \mid s') \succeq_{DS_1} F_t(x \mid s)\}.$$

L'individu neutre au risque choisira donc l'un des milieux d'origine  $s$  qui appartient à cet ensemble. Si cet ensemble est trop grand et si l'individu est averse au risque, il peut poursuivre la comparaison en utilisant un test de dominance stochastique d'ordre 2. On obtient alors l'ensemble  $P_{2t}$  qui est défini d'une manière analogue à  $P_{1t}$  avec  $P_{2t} \subseteq P_{1t}$ .

Une définition de l'égalité des chances découle naturellement de l'obtention de ces deux ensembles de milieux sociaux non dominés. On dira que l'égalité des chances est réalisée pour l'année  $t$  si pour l'ensemble des milieux d'origine  $S$ , il n'existe pas de relation de dominance, à l'ordre 1 ou à l'ordre 2, entre les différentes distributions de revenu conditionnel  $F_t(x \mid s)$ . Cette définition de l'égalité des chances recouvre en particulier le cas d'égalité des distributions conditionnelles, que l'on peut qualifier d'égalité des chances au sens fort. On se doute toutefois que cette égalité des chances au sens fort risque d'être rarement satisfaite en pratique, dans la mesure où elle impose des restrictions très particulières sur les distributions conditionnelles. En outre, se restreindre à cette seule notion d'égalité forte ne permettrait pas de distinguer, lorsque deux distributions conditionnelles ne sont pas égales, les cas où l'on ne peut pas établir de classement des distributions et ceux où un classement peut être opéré, alors même que ces deux situations diffèrent profondément, du point de vue de l'inégalité des chances. L'un des avantages de notre critère d'égalité des chances, défini comme absence de dominance<sup>10</sup>, est, au contraire, de permettre de distinguer ces situations.

Notre définition de l'égalité des chances sera donc donnée, formellement, par :

**Définition** L'égalité des chances est réunie si  $P_{2t} \equiv S$ .

Restreindre la définition de l'égalité des chances à l'absence de dominance stochastique d'ordre 1 aurait abouti à une définition encore plus "faible", puisque si la dominance

<sup>10</sup>Il est à noter que notre critère n'implique pas l'égalité des moyennes conditionnelles (critère de Van de Gaer). Roemer pose que les disparités de réussite pour une CS d'origine donnée sont uniquement dues aux différences d'effort, une interprétation dont notre analyse est absente. Roemer demande que les déciles des distributions conditionnelles soient égaux, une exigence que ne respecte pas non plus le critère adopté dans son papier. Le critère d'égalité des chances au sens fort respecterait lui les conditions de Van de Gaer et de Roemer.

stochastique d'ordre 2 est incapable d'établir une hiérarchie entre les milieux sociaux, il en est de même avec la dominance d'ordre 1. Il est loisible de poursuivre l'exercice avec la dominance stochastique aux ordres supérieurs à 2 mais la pertinence de l'exercice est cependant plus limitée, dans la mesure où, du point de vue de l'analyse des choix en univers risqué, les dominances stochastiques d'ordre 3 ou 4 ne sont pertinentes que dans des contextes bien particuliers.

Notons pour finir que comparer des distributions conditionnelles centrées à leur moyenne peut avoir également un sens si l'on souhaite comparer les loteries, du seul point de vue du risque, et indépendamment de leurs rendements respectifs. On utilisera alors le critère des courbes de Lorenz pour effectuer cette comparaison. D'une manière équivalente, le résultat de cette comparaison nous renseigne sur le caractère plus ou moins inégalitaire des distributions conditionnelles. En notant  $L_t^s(p)$  l'ordonnée de la courbe de Lorenz en  $p$ , pour le milieu d'origine  $s$  et l'année  $t$ , nous dirons que la loterie associée au groupe  $s$  est moins risquée que celle du groupe  $s'$  si :

$$L_t^s(p) \geq L_t^{s'}(p), \forall p \in [0, 1].$$

## 2.2 Inférence statistique

La mise en œuvre empirique de ces critères d'égalité des chances, à partir de données d'enquête suppose évidemment de se doter d'instruments d'inférence statistique adaptés. Comme nous travaillons sur des échantillons qui ne sont pas de très grande taille, la question de la robustesse statistique de nos résultats est évidemment cruciale. Les procédures mises en œuvre sont nécessairement un peu complexes et aucune procédure "standard" ne s'est encore complètement imposée. Il faut ajouter que notre problème est un peu particulier, dans la mesure où nous cherchons à tester l'absence de relation de dominance, alors qu'usuellement, l'intérêt se focalise sur la recherche d'une relation de dominance. Le choix de l'hypothèse nulle testée est alors susceptible d'affecter la nature des conclusions obtenues. Il convient donc d'explicitier les conséquences possibles de notre choix d'hypothèse nulle avant de détailler la méthode de test retenue dans la suite de cet article.

### 2.2.1 Le choix de l'hypothèse nulle

Soit deux distributions conditionnelles établies sur la base d'un échantillon. Intéressons nous par exemple au test de dominance d'ordre 2, c'est à dire, aux aires sous la fonction de répartition pour tous les revenus possibles. Trois cas de figure sont envisageables. L'aire comprise sous l'une des distributions est toujours plus grande que l'aire comprise sous l'autre distribution ; l'aire est tantôt plus grande, tantôt plus petite ; les aires sont rigoureusement égales. Si ces distributions étaient établies à partir de données exhaustives sur l'ensemble de la population, nous conclurions à l'égalité des chances dans les deux derniers cas de figure et à l'inégalité dans le premier cas. En présence de données d'échantillonnage, il faut être plus prudent et tester la significativité statistique des éventuelles relations de dominance observées. En outre ce test ne porte pas sur un seul coefficient estimé mais sur un vecteur de coefficients (des différences entre les aires sous les deux distributions calculées pour différents niveaux de revenu) dont on teste l'appartenance à une zone particulière (par exemple : toutes les différences sont positives).

Comme pour tout test, le choix de l'hypothèse nulle testée (ici, l'hypothèse nulle de dominance ou celle de non-dominance) n'est pas sans conséquence sur le résultat du test. Le choix de l'hypothèse nulle de dominance (respectivement de non-dominance) maximise la probabilité de ne pas rejeter la dominance (resp. non-dominance) et donc minimise la probabilité de ne pas rejeter la non-dominance (resp. dominance).

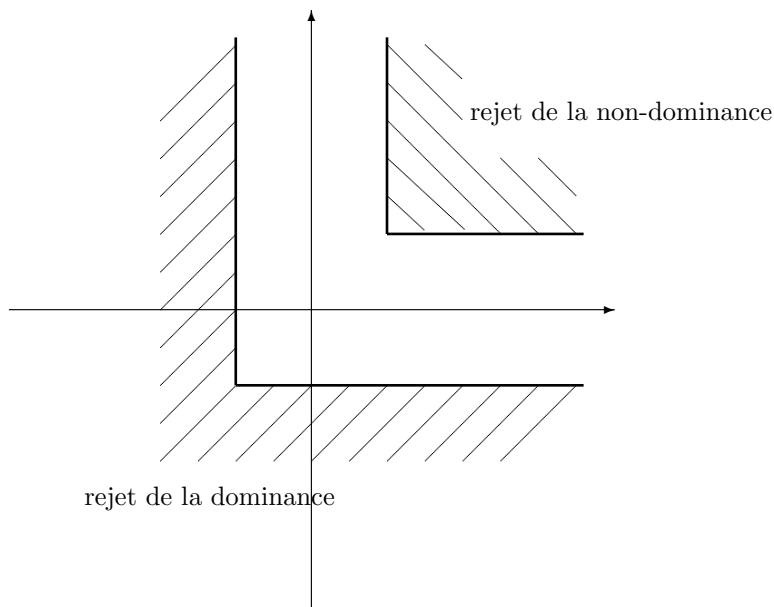
La figure 1 illustre cette différence dans le cas où la comparaison porte sur un vecteur à deux dimensions : le calcul de l'aire sous la fonction de répartition a été effectuée en deux points du support. Sans incertitude due à l'échantillonnage, nous concluons à la dominance si le vecteur appartient à l'orthant positif. Dans les autres cas de figure nous concluons à la dominance.

Si nous retenons comme hypothèse nulle, la dominance, la région d'acceptation est plus grande que l'orthant positif : nous acceptons <sup>11</sup> de conclure à la dominance en considérant que des valeurs extérieures à l'orthant positif mais cependant proches de celui-ci peuvent résulter des particularités de l'échantillon tiré et ne pas refléter les caractéristiques de la

---

<sup>11</sup>On dira plus rigoureusement que la région de non-rejet de l'hypothèse nulle de dominance est plus grande que la région de rejet de l'hypothèse nulle de non-dominance.

FIG. 1: zones de rejet des hypothèses nulles de dominance et de non-dominance



population tout entière. La région où nous concluons à l'absence de dominance s'en trouve réduite.

Si nous retenons maintenant la non-dominance comme hypothèse nulle, nous acceptons de prendre le risque de ne pas la rejeter même si les coordonnées du vecteur de différences sont toutes deux légèrement positives. Dans ces conditions, la région dans laquelle le test conclut à la dominance est plus petite que l'orthant positif ou inversement la région où nous acceptons la non-dominance s'en trouve augmentée.

L'aire comprise entre les deux zones de rejet représente la zone de conflit suivant le choix de l'hypothèse nulle.

Appliqué à notre objet, cette discussion indique que le fait de retenir la non-dominance comme hypothèse nulle<sup>12</sup> conduira à conclure plus souvent en faveur de l'égalité des chances, tandis que le choix de l'hypothèse nulle de dominance conduira à conclure plus souvent à l'inégalité des chances. Dans notre cas, nous choisirons de retenir la dominance comme hypothèse nulle. Ceci implique que lorsque nous concluons, sur la base de ce test à l'égalité des chances entre deux distributions, nous sommes certains que cette conclusion

<sup>12</sup>Howes (1993) considère ce cas de figure.

aurait également été obtenue en renversant le test.

### 2.2.2 Principe général des tests : une comparaison des montants de pauvreté

Les relations de dominance théoriques définies précédemment trouvent une expression alternative, particulièrement utile dans les analyses empiriques, si on utilise la notion de montant de pauvreté. Formellement l'indicateur de montant de pauvreté à l'ordre  $g$ , pour un seuil de revenu  $z$  est défini par :

$$D^g(z) = \int_0^z (z - x_i)^{g-1} dF(x).$$

Par exemple pour  $g = 1$ ,  $D^1(z)$  est simplement la proportion d'individus dont le revenu est inférieur au montant  $z$  ; pour  $g = 2$ ,  $D^2(z)$  représente le montant de pauvreté moyen, c'est à dire le montant qu'il faut donner en moyenne aux individus pour qu'ils soient au dessus du seuil de pauvreté  $z$ .

Il est connu (Foster et Shorrocks (1988) par exemple) que si la distribution A domine à l'ordre 2 la distribution B, alors le montant de pauvreté à l'ordre 2 pour la distribution B est toujours supérieur au montant de pauvreté pour la distribution A. Plus généralement, pour deux populations A et B, il existe une relation d'équivalence qui lie dominance stochastique de la distribution  $F_A$  sur la distribution  $F_B$ , pour tout ordre de dominance et montants de pauvreté de chacune des distribution. On peut en effet établir que pour tout  $g \in \mathbb{N}_+$  :

$$F_A(x) \succeq_{DS_g} F_B(x) \Leftrightarrow D_A^g(z) < D_B^g(z), \forall z \in \mathbb{R}_+.$$

où  $\succeq_{DS_g}$  désigne la dominance stochastique à l'ordre  $g$ .

Le principe du test statistique de dominance proposé dans de nombreux travaux (voir par exemple, Davidson et Duclos (2000)) consiste à tester cette inégalité des montants de pauvreté pour un nombre fini  $k$  de seuils de pauvreté estimés. Pour un seuil de pauvreté  $z$  donné, un estimateur sans biais et asymptotiquement normal du montant de pauvreté

$D^g(z)$  est donné par<sup>13</sup> :

$$\hat{D}^g(z) = \frac{1}{N(g-1)!} \sum_{i=1}^N (z - x_i)^{g-1} I(x_i \leq z),$$

où  $N$  est la taille de l'échantillon,  $x_i$  le revenu du  $i^{\text{ème}}$  individu et  $I(\cdot)$  est une fonction indicatrice. Pour un ensemble de seuils de pauvreté donné, on notera  $\hat{\mathbf{D}}^g$  le vecteur de montants de pauvreté estimé et  $\Sigma$  sa matrice de variance-covariance asymptotique dont un estimateur est donné dans Davidson et Duclos<sup>14</sup>.

On peut réexprimer l'hypothèse de dominance à l'ordre  $g$  sous forme vectorielle. Soit  $\delta$  le vecteur des différences<sup>15</sup> des montants de pauvreté calculées pour  $k$  seuils de pauvreté dans la population. On veut tester  $\mathbf{H}_0 : \delta \in \mathbb{R}_+^k$  contre  $\mathbf{H}_1 : \delta \notin \mathbb{R}_+^k$ . La procédure de test repose sur le vecteur  $\hat{\delta} = (\hat{\mathbf{D}}_B^g - \hat{\mathbf{D}}_A^g)$  dont la matrice de variance-covariance asymptotique est donnée par  $\frac{\Sigma_A}{N_A} + \frac{\Sigma_B}{N_B}$  sous l'hypothèse d'indépendance des distributions  $A$  et  $B$ <sup>16</sup>.

Comme on le voit, l'hypothèse nulle est donc définie par un ensemble de  $k$  contraintes. Deux voies permettent de procéder au test de cette hypothèse. La première consiste à tester séparément chacune des  $k$  contraintes. On teste alors l'intersection de  $k$  sous-hypothèses nulles définies en chacun des points où l'on compare les distributions. Ce type de test, aussi appelé test "union-intersection" est par exemple utilisé par Bishop *et alii* (1992). On peut toutefois montrer que ce type de test est relativement peu puissant. Comme le soulignent Davidson et Duclos, ceci tient en particulier au fait que la structure de covariance du vecteur  $\hat{\delta}$  n'est pas exploitée dans la formulation de ce type de test. La seconde voie, que nous retiendrons ici, consiste à tester globalement l'ensemble des contraintes à l'aide d'un test de Wald. Ce test agrège l'ensemble des  $k$  contraintes et tient compte explicitement de la structure de variance-covariance des montants de pauvreté estimés pour différents seuils de revenu. Le principe général de ce test consiste à examiner la "distance" du vecteur estimé  $\hat{\delta}$  aux deux ensembles définissant respectivement l'hypothèse nulle et l'hypothèse

<sup>13</sup>La formule a été adaptée à l'utilisation de variables de pondérations des données.

<sup>14</sup>Voir Davidson et Duclos (2000), théorème 1. p.1441.

<sup>15</sup>Le montant de pauvreté de la distribution B diminué du montant de pauvreté de la distribution A.

<sup>16</sup>Il est aussi possible de calculer la matrice de variance covariance de  $\hat{\delta}$  si les deux ensembles d'observations résultent de tirages *i.i.d* de  $N$  couples de niveaux de revenus dans une loi jointe. La comparaison de distributions avant et après impôt peut constituer un exemple de cette deuxième possibilité.

alternative. Bien que les tests globaux de Wald, que nous utiliserons, soient plus complexes à mettre en œuvre en pratique que les tests union-intersection, leur plus grande puissance, démontrée notamment dans Dardanoni et Forcina (1999), plaide largement en faveur de leur utilisation.

Enfin, les tests de Wald peuvent être utilisés pour tester à la fois l'égalité de deux distributions et l'existence d'une relation de dominance. Notre examen de l'égalité des chances combine ces deux types de tests, dont nous précisons maintenant le détail.

### 2.2.3 Les tests d'égalité des distributions

L'égalité de deux distributions traduit une égalité des chances au sens fort, telle que nous l'avons discutée ci-dessus. Le test de Wald de l'égalité de deux distributions est relativement aisé à mettre en œuvre. Dans ce cas, l'hypothèse nulle est donnée par  $\mathbf{H}_0 : \boldsymbol{\delta} = \mathbf{0}$ . On peut alors montrer (voir notamment Beach et Davidson (1983) et Davidson et Duclos (2000)) que sous l'hypothèse nulle, le vecteur des différences de montant de pauvreté estimé est asymptotiquement normal et on a :

$$\hat{\boldsymbol{\delta}} \sim \mathcal{N}\left(0, \frac{\boldsymbol{\Sigma}_A}{N_A} + \frac{\boldsymbol{\Sigma}_B}{N_B}\right)$$

On en déduit donc la loi asymptotique de la statistique  $T_1$  : sous l'hypothèse nulle d'égalité, on a :

$$T_1 = \hat{\boldsymbol{\delta}}' \left( \frac{\boldsymbol{\Sigma}_A}{N_A} + \frac{\boldsymbol{\Sigma}_B}{N_B} \right)^{-1} \hat{\boldsymbol{\delta}} \sim \chi_k^2.$$

On peut donc tester l'égalité des deux distributions  $A$  et  $B$  à l'aide d'un test du  $\chi^2$  sur la statistique  $T_1$ .

### 2.2.4 Les tests de dominance faible

Les tests de dominance faible sont plus complexes à mettre en œuvre dans la mesure où l'ensemble correspondant à l'hypothèse nulle est défini par une contrainte d'inégalité. On cherche à tester  $\mathbf{H}_0 : \boldsymbol{\delta} \in \mathbb{R}_+^k$  contre  $\mathbf{H}_1 : \boldsymbol{\delta} \notin \mathbb{R}_+^k$ . La statistique du test de Wald en présence de contraintes de ce type a été analysée par Kodde et Palm (1986) et Wolak

(1989). Pour ce jeu d'hypothèses, on considère la statistique  $T_2$  définie par :

$$T_2 = \min_{\boldsymbol{\delta} \in \mathbb{R}_+^k} \|\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}\|$$

où  $\|x\| = x'\Sigma^{-1}x$ . Kodde et Palm (1986) démontrent que la statistique  $T_2$  suit un mélange de loi du  $\chi^2$  donné par :

$$T_2 \sim \sum_{j=0}^k w(k, k-j, \Sigma) Pr(\chi_j^2 \geq c)$$

où  $w(k, k-j, \Sigma)$  représente la probabilité que  $k-j$  des éléments de  $\boldsymbol{\delta}$  soient strictement positifs. La loi de ce mélange de  $\chi^2$  n'est pas tabulé mais des bornes inférieures et supérieures des valeurs critiques sont cependant données par Kodde et Palm. Il est par ailleurs possible, si les bornes inférieures et supérieures ne permettent pas de conclure, d'estimer les valeurs critiques de la statistique  $T_2$  par une procédure de Monte-Carlo<sup>17</sup>.

### 2.2.5 L'enchaînement des tests

Notre procédure couple des tests de dominance faible et des tests d'identité des distributions. Il est loisible de se demander s'il n'y a pas là une redondance dans la mesure où si une distribution A domine faiblement une distribution B et réciproquement, on doit évidemment conclure à l'identité des deux distributions. Cependant, il se trouve que le test d'égalité des deux distributions est un test plus puissant que le test de dominance. Donc la probabilité de ne pas rejeter l'hypothèse d'identité des distributions alors que c'est une hypothèse fautive est plus importante avec la conjonction des deux tests de dominance qu'avec le test d'identité. Si on attache une importance spécifique à l'égalité des distributions, il faut donc réaliser séparément le test d'égalité.

L'enchaînement des tests est le suivant. Dans un premier temps, on teste l'hypothèse de dominance faible de la distribution A par la distribution B et de la distribution B par la distribution A, pour la dominance stochastique d'ordre 1. Si une des hypothèses de dominance est acceptée, on conclut à l'inégalité des chances. Si les deux hypothèses de

<sup>17</sup>Il suffit de tirer 10.000 vecteurs normaux multivariés d'espérance 0 et de matrice de variance covariance  $\Sigma$ , puis de calculer ensuite la proportion de vecteurs qui ont  $j$  éléments positifs (pour  $j \in (0, k)$ ), proportion qui est un estimateur du poids  $w(k, j, \Sigma)$ .



dominance sont acceptées ou si la dominance est rejetée dans les deux cas, on teste alors l'égalité des distributions à l'aide du test de Wald simple. Si l'égalité des deux distributions est acceptée, on conclut à l'égalité des chances (au sens fort). Si l'égalité des deux distributions est rejetée, on teste alors la dominance de  $A$  par  $B$  et de  $B$  par  $A$ , pour la dominance stochastique d'ordre 2. Si une seule des deux hypothèses de dominance est acceptée, on conclut à l'inégalité des chances. Sinon, on conclut à l'égalité des chances. La même procédure est appliquée pour les distributions centrées.

Pour finaliser la procédure, il reste encore à choisir les points du support où l'on calcule les éléments du vecteur  $\delta$  qui intervient dans la statistique de test. Nous avons retenu les déciles et le 19<sup>ème</sup> vingtile des enquêtes Revenus Fiscaux correspondant aux années des enquêtes de BdF<sup>18</sup>. La théorie statistique est en effet établie en supposant que les points du support ne sont pas tirés aléatoirement et donc l'utilisation de quantiles empiriques est déconseillée.

### 3 Données

Les données exploitées dans cette étude proviennent des enquêtes Budget des Familles (BdF) de l'INSEE. Ces enquêtes ont été réalisées tous les cinq ans, depuis 1979, auprès d'échantillons d'environ 12 000 ménages, représentatifs de la population des ménages ordinaires français. Nous utilisons ici les cinq vagues disponibles : 1979, 1984, 1989, 1994 et 2000. Pour chaque ménage enquêté, les données fournissent une information détaillée sur les ressources perçues et les dépenses effectuées et permettent d'identifier le milieu social d'origine du ménage. Il est alors possible d'estimer et de comparer les distributions de revenu conditionnelles à l'origine sociale.

#### 3.1 Champ

##### **Origine Sociale : les Pères**

Dans les enquêtes BdF, le chef de ménage et son conjoint sont interrogés sur la catégorie

---

<sup>18</sup>Plus exactement, les déciles de la distribution du revenu disponible par ménage par unité de consommation.

socio-professionnelle de leurs deux parents<sup>19</sup>. A partir de ces quatre variables, il est donc, en principe, possible d'établir une nomenclature détaillée des milieux sociaux d'origine. Ceci aboutirait cependant à un découpage trop fin de l'échantillon des ménages enquêtés et conduirait à des estimations peu précises des distributions de revenu conditionnelles à l'origine sociale. De ce fait, nous ne retenons ici, pour caractériser l'origine sociale du ménage, que l'information afférente à la catégorie socio-professionnelle du père du chef de ménage. La CSP de la mère est moins exogène que la CSP du père par rapport à la présence d'enfants. Ceci nous amène à distinguer six groupes d'origine sociale distincts : descendants de pères agriculteurs ; descendants d'artisans commerçants et chefs d'entreprise ; descendants de cadres ; de professions intermédiaires ; d'employés ; d'ouvriers<sup>20</sup>.

### **Descendance**

Afin de réduire les possibles effets de composition par âge au sein de chaque groupe d'origine sociale, l'échantillon étudié est restreint aux ménages dont le chef est âgé de 30 à 50 ans à la date de l'enquête. Dans le cadre de cette étude, nous nous limitons par ailleurs aux ménages dont le chef est de sexe masculin. Les ménages dont le chef se déclare retraité ou étudiant, ou fils de retraité ou d'étudiant, sont écartés de l'analyse. En appliquant ces critères de sélection, on aboutit à un échantillon de 3347 ménages en 2000, 3993 en 1994, 3152 en 1989, 3958 en 1984 et 3914 en 1979.

### **Deux mesures du niveau de vie du ménage**

L'information disponible permet une évaluation exhaustive des ressources perçues par les ménages<sup>21</sup>. Ceci représente un avantage majeur des enquêtes BdF, par rapport aux autres bases de données existantes. Schématiquement, les ménages enquêtés sont interrogés sur trois grandes sources de revenu : les revenus d'activité (y compris revenus d'activité indépendante), les revenus du patrimoine et les revenus de transferts (nets d'impôts sur le revenu). Le détail des sources de revenu déclarés à l'enquête est présenté en annexe.

---

<sup>19</sup>Pour toutes les enquêtes, sauf 1979, il s'agit de la catégorie socio-professionnelle des parents lorsque l'individu enquêté avait 16 ans. En 1979, il s'agit de la dernière catégorie socio-professionnelle d'active des parents de l'individu.

<sup>20</sup>Les catégories socioprofessionnelles déclarées à l'enquête 1979 ont été recodées pour tenir compte du changement dans la nomenclature PCS intervenu en 1982.

<sup>21</sup>Nous avons également cherché à construire une variable individualisée de revenu, mais cette construction soulève des difficultés quant à la répartition entre les membres du ménage de ressources communes comme les revenus du patrimoine ou les revenus de transfert. La prise en compte d'effets d'échelle et la mise en commun des ressources au niveau du ménage rend par ailleurs l'exercice très difficile et moins satisfaisant.

L'ensemble des revenus a été déflaté par l'indice des prix à la consommation de l'INSEE (base 100 en 2000).

Nous distinguons, dans cette étude, deux mesures du revenu du ménage. La première correspond au revenu primaire du ménage, avant impôts et transferts redistributifs. Elle inclut les revenus d'activité et les revenus du patrimoine du ménage, ainsi que les revenus de transferts assurantiels (chômage, retraite)<sup>22</sup>. La seconde est le revenu disponible des ménages, égal au revenu net d'impôts et de transferts redistributifs. Dans les deux cas, nous tenons compte de la taille du ménage en appliquant l'échelle d'équivalence de l'OCDE, qui revient à déflater le revenu du ménage par la racine carrée du nombre d'individus du ménage<sup>23</sup>. Finalement, il faut noter que nous avons utilisé les variables de pondérations fournies dans la base. Ces variables permettent aux données définies au niveau des ménages d'être représentatives de la population totale des ménages ordinaires français. Nous avons donc pondéré toutes nos observations par le poids que l'INSEE leur a attribué.

### 3.2 Traitement de l'échantillon

L'analyse de ces échantillons fait apparaître, pour les premières vagues d'enquête, un taux de non-réponse important aux questions sur les revenus du ménage. Ces cas de non-réponse peuvent correspondre soit à une absence complète de réponse à une ou plusieurs questions concernant les revenus (ou les impôts payés), soit à une déclaration en tranche des revenus perçus par le ménage. En outre, la non-déclaration partielle ou complète des revenus ne peut être assimilée, du point de vue de notre étude, à un phénomène aléatoire de non-réponse. Tout d'abord, comme le montre le tableau 14, le taux de non-réponse varie fortement en fonction de la catégorie socio-professionnelle du chef de ménage : il est en particulier très élevé, dans les premières vagues, pour les ménages dont le chef est agriculteur ou artisan, et est beaucoup plus faible pour les ménages dont le chef est salarié. Ne considérer que les déclarations de revenu non-manquantes, dans le calcul des distributions de revenu conditionnelles, conduirait donc à une évaluation biaisée des perspectives de mobilité sociale offertes aux ménages : par exemple, on sous-estimerait implici-

<sup>22</sup>Nous suivons ici la distinction maintenant bien établie dans Bourguignon (1998).

<sup>23</sup>C'est l'échelle d'équivalence de l'OCDE, d'autres échelles sont possibles, notamment celle de l'INSEE ou l'échelle d'Oxford.

tement la proportion de fils d'agriculteurs devenant eux-mêmes agriculteurs. Par ailleurs, au sein de chaque groupe d'origine sociale, les cas de non-réponse ne semblent pas, non plus, aléatoires : même à origine donnée, la probabilité de non-réponse est corrélée aux caractéristiques observables du ménage. En outre, si on prédit le revenu des ménages non-répondants à partir de ces caractéristiques observables, il ressort que le revenu moyen prédit est en général plus faible que le revenu moyen calculé sur les ménages répondants, notamment pour les ménages d'agriculteurs et d'artisans. En excluant les cas de non-réponses du champ de notre étude, on aboutirait donc, là encore, à une estimation biaisée des distributions de revenu.

Pour ces raisons, en cas de non-réponse totale ou partielle, nous avons procédé à l'imputation du revenu du ménage par la méthode dite du résidu simulé. Le principe de cette procédure est le suivant. On estime tout d'abord, pour les ménages déclarant un revenu (en niveau ou en tranche), une équation de revenu, où le revenu du ménage est régressé sur une série de caractéristiques observables (âge, sexe, catégorie professionnelle, dernier diplôme obtenu, consommation, nationalité, structure familiale, zone géographique d'habitation. . .)<sup>24</sup>. On prédit ensuite le revenu des ménages ne déclarant pas de revenu à partir de leurs caractéristiques observables. Enfin, on procède au tirage aléatoire d'un résidu de l'équation de salaire, qui ajouté au salaire prédit à partir des caractéristiques observables donne la valeur imputée du revenu du ménage. Cette procédure est utilisée pour imputer les revenus avant et après impôts et transferts des ménages non-répondants.

Enfin, nous avons procédé à une opération assez fouillée de vérification des données de revenu d'exploitation agricole (voir annexe 4). En effet, comme les résultats vont l'attester, un changement important semble s'être produit durant cette période qui a eu pour conséquence de bousculer le classement des CS. Nous voulions nous assurer que les changements dont nous allons faire état n'étaient pas le fruit d'une illusion statistique.

---

<sup>24</sup>Le détail des équations d'imputation est présenté dans l'annexe 3.

## 4 Résultats

Pour étudier l'évolution de l'inégalité des chances depuis 1979, nous procédons tout d'abord à l'analyse des relations de dominance existant entre les différentes loteries sociales, pour chacune des vagues d'enquête disponibles <sup>25</sup>. Dans un second temps, nous examinons les facteurs ayant pu modifier ou atténuer, au cours des vingt-dernières années, ces relations de dominance. De ce point de vue, deux facteurs méritent d'être isolés : d'une part, le rendement moyen et d'autre part, le risque inhérent à chaque loterie. En effet, l'existence d'une relation de dominance à l'ordre 2 peut provenir schématiquement soit d'un supplément de rendement à risque constant soit d'une diminution du risque à rendement donné.

Trois résultats principaux se dégagent alors de cette analyse. En premier lieu, on observe une profonde permanence des positions relatives des différents groupes sociaux, entre 1979 et 2000 : les situations de dominance observées en début de période demeurent ou se modifient peu. Toutefois, les écarts de rendement moyen entre les différentes loteries sociales tendent à se réduire au cours de la période étudiée et on observe donc une diminution du *degré* d'inégalité des chances, en dépit de la permanence des classements relatifs. Enfin, il apparaît que le degré de risque afférent à chaque loterie est approximativement similaire pour les différents milieux sociaux d'origine. Nous exploitons cette propriété pour cerner la part que représente l'inégalité des milieux d'origine dans l'inégalité des résultats.

### 4.1 Une réduction du degré d'inégalité des chances

L'examen des courbes de Lorenz généralisées, pour le revenu primaire, présentées dans la figure 2 permet une première lecture de l'inégalité des chances en 1979. Il fait apparaître un classement particulièrement net des distributions de revenu offertes aux individus en fonction de leur groupe social d'origine. Au sommet de la hiérarchie des milieux sociaux d'origine, on trouve les fils de cadres dont la distribution domine très largement celle des autres groupes sociaux. Ils sont suivis, d'assez loin, par les individus issus des professions intermédiaires. Les fils d'employés et d'artisans occupent une place intermédiaire et les dis-

---

<sup>25</sup>Comme cela a déjà été indiqué plus haut, les données ont été systématiquement pondérées, tous les résultats sur données non pondérées sont disponibles sur demande auprès des auteurs.

tributions de revenu offertes aux individus de ces deux groupes apparaissent relativement proches entre elles. Enfin, en bas de la hiérarchie on trouve les fils d'ouvriers et les fils d'agriculteurs, ces derniers semblant nettement dominés par l'ensemble des autres milieux sociaux.

Cette inégalité des chances ne semble, en 1979, que très peu atténuée par la politique redistributive. Si les écarts entre les courbes de Lorenz généralisées des différents milieux d'origine diminuent (figure 2), les relations de dominance observées avant impôts et transferts sociaux se maintiennent toutes après redistribution.

Ce classement est validé par les résultats des tests de dominance stochastique présentés dans les tableaux 1 qui 2 appellent d'abord des remarques d'ordre méthodologique. Les critères de dominance permettent presque toujours d'établir un classement des différentes distributions : sur l'ensemble des vagues d'enquête, la proportion des cas où on ne conclut ni à l'égalité, ni à la dominance est de 5% (7/150 exactement). Le coût d'avoir choisi des instruments robustes, comme les instruments de dominance, en termes de diminution dans notre capacité de classement, est donc faible. Cette situation est assez rare pour être notée et justifie ex-post notre choix d'instruments. Ensuite, il faut noter que les relations de dominance stricte sont toutes obtenues avec un critère de dominance d'ordre 1. Comme on le sait, ce critère, qui conduit à ce que la loterie dominante soit préférée quelque soit le goût du risque du décideur, est particulièrement robuste.

Les évolutions à l'œuvre entre 1979 et 2000, ne modifient que très partiellement les relations de dominance observées en début de période. Les résultats des tests de dominance du tableau 1 permettent de souligner trois phénomènes marquants. En premier lieu, la position dominante des fils de cadres se maintient tout au long de la période : pour chacune des années la distribution de revenu de ce groupe domine celle de tous les autres milieux d'origine. En second lieu, la hiérarchie des groupes intermédiaires semble s'atténuer. Ceci tient pour partie à l'amélioration de la position relative des fils d'artisans au cours de la période : alors que ceux-ci étaient dominés par les fils de professions intermédiaires en début de période, ce n'est plus le cas à partir de 1989. Enfin, on observe une permutation des positions relatives dans le bas de la hiérarchie des milieux sociaux d'origine : à partir de 2000, la distribution de revenu conditionnelle des fils d'agriculteurs domine celles des fils

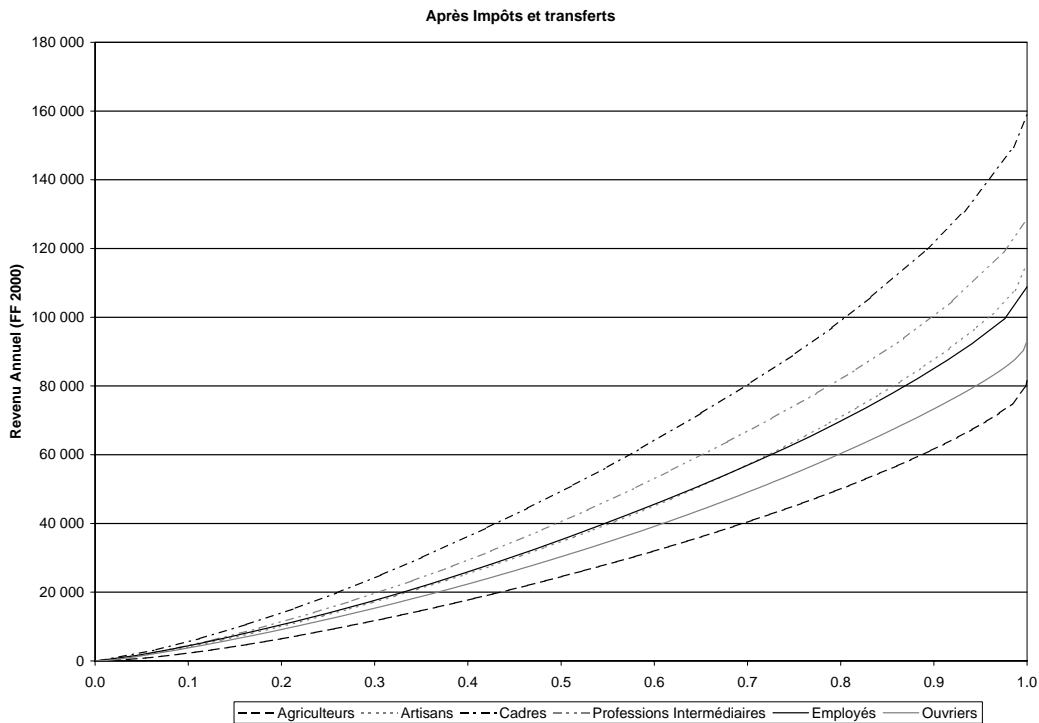
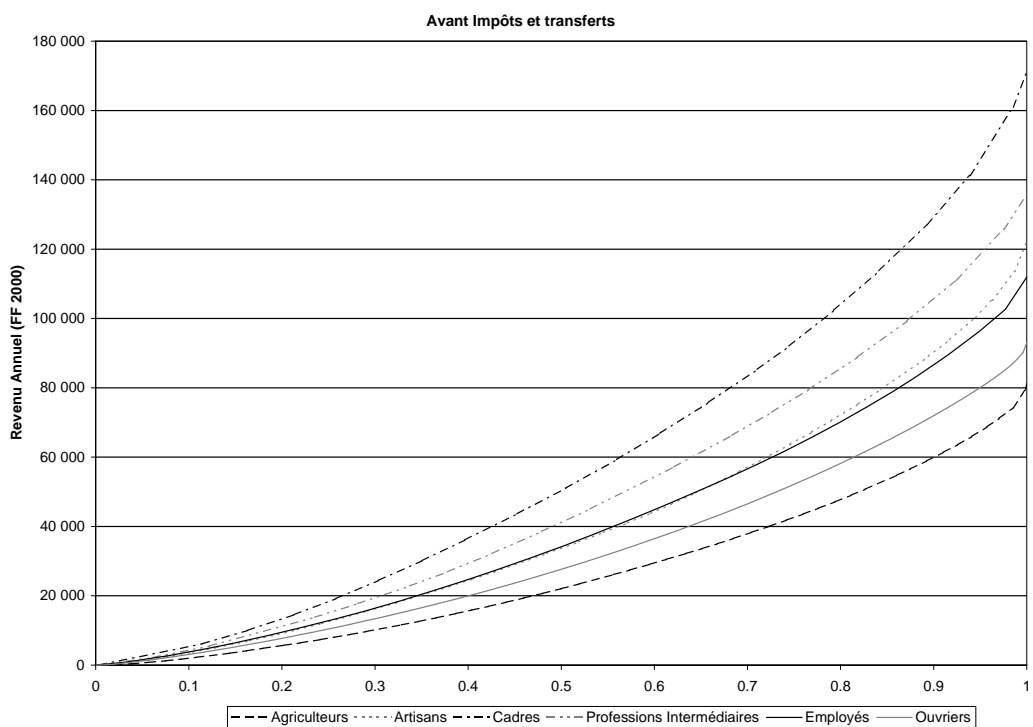


FIG. 2: Courbes de Lorenz Généralisées en 1979

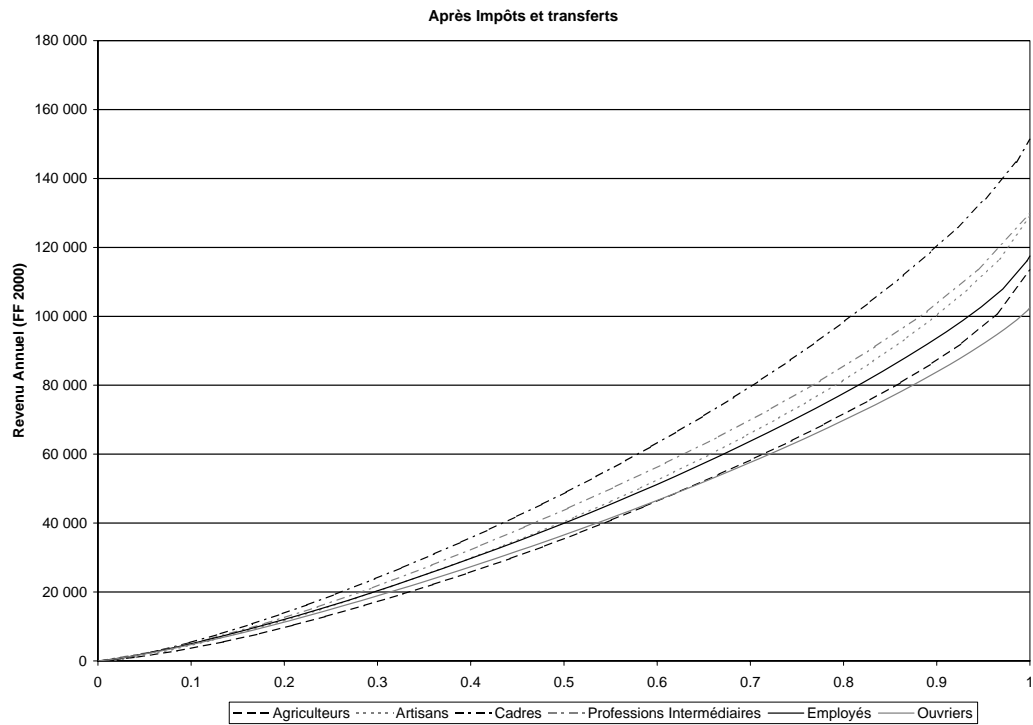
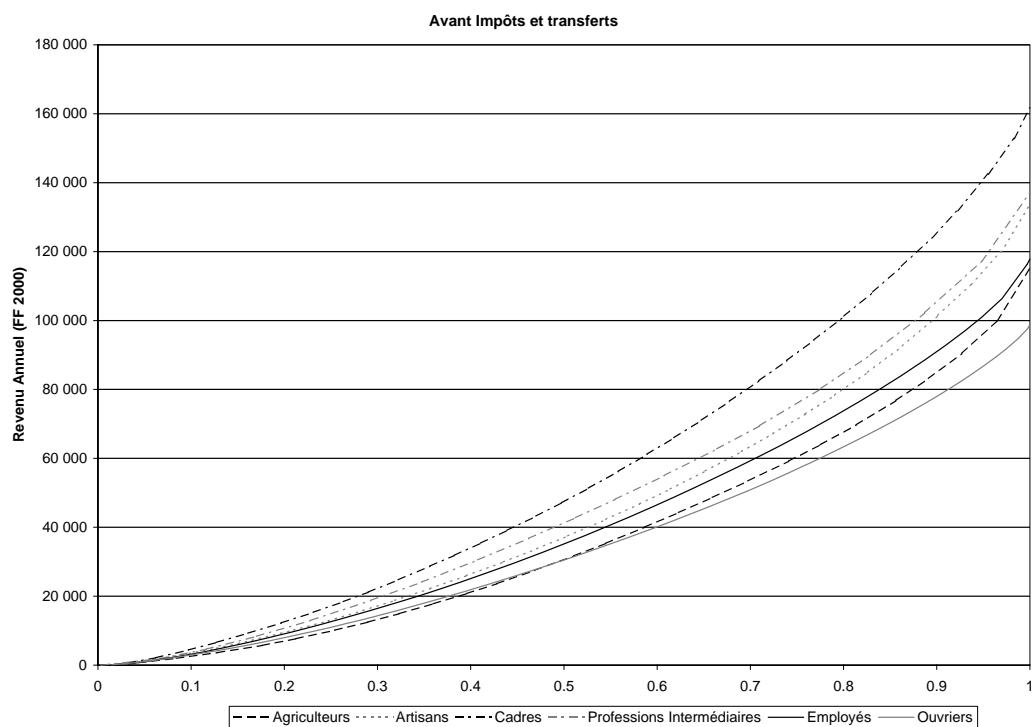


FIG. 3: Courbes de Lorenz Généralisées en 2000



<b>1979</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	< <sub>1</sub>	?	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1984</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	< <sub>1</sub>	=	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1989</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	=	=	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	=	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1994</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	?
Artisans	-	-	< <sub>1</sub>	=	> <sub>1</sub>	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	=	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>2000</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	> <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	=	> <sub>1</sub>	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-

Le tableau donne la situation de la catégorie en ligne par rapport à la colonne  
<<sub>1</sub> : La ligne est dominée par la colonne à l'ordre 1.  
><sub>1</sub> : la ligne domine la colonne à l'ordre 1.  
= : Égalité des distributions.  
? : Non comparabilité.

TAB. 1: Tests de dominance stochastique - Avant Impôts et transferts

<b>1979</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	< <sub>1</sub>	=	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1984</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1989</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>
Artisans	-	-	< <sub>1</sub>	=	?	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	?	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>1994</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	?
Artisans	-	-	< <sub>1</sub>	?	> <sub>1</sub>	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	=	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-
<b>2000</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	< <sub>1</sub>	?
Artisans	-	-	< <sub>1</sub>	=	> <sub>1</sub>	> <sub>1</sub>
Cadres	-	-	-	> <sub>1</sub>	> <sub>1</sub>	> <sub>1</sub>
Prof Interm	-	-	-	-	> <sub>1</sub>	> <sub>1</sub>
Employés	-	-	-	-	-	> <sub>1</sub>
Ouvriers	-	-	-	-	-	-

Le tableau donne la situation de la catégorie en ligne par rapport à la colonne.

<<sub>1</sub> : La ligne est dominée par la colonne à l'ordre 1.

><sub>1</sub> : la ligne domine la colonne à l'ordre 1.

= : Égalité des distributions.

? : Non comparabilité.

TAB. 2: Tests de dominance stochastique - Après Impôts et transferts

d'ouvriers. Ce dernier groupe social occupe désormais la fin du classement et est dominé par l'ensemble des autres milieux d'origine.

Ce classement n'est, là encore, et comme pour toutes les années étudiées, que très peu affecté par la politique redistributive. Le seul changement concerne la comparaison fils d'agriculteurs - fils d'ouvriers : alors que les premiers dominent les seconds en termes de revenu primaire, il y a égalité des chances en termes de revenu disponible.

On serait donc tenté de conclure à la lecture des tests de dominance stochastique que l'égalité des chances a, au cours de la période 1979-2000, peu progressé. Dans les deux premières vagues, 1979 et 1984, on était toujours capable de signer à une exception près une comparaison bilatérale entre deux classes sociales. Pour les trois autres vagues, l'égalité des chances peut tout au plus émerger de deux ou trois comparaisons, l'incertitude provenant essentiellement de la montée des deux classes de non-salariés par rapport aux classes salariés.

Au total, une hiérarchie à trois niveaux des différents milieux d'origine semble nettement persister : un groupe dominant, formé des fils de cadres, un groupe intermédiaire, de plus en plus homogène, formé des fils d'employés, d'artisans et de professions intermédiaires ; enfin, en bas de cette hiérarchie, un groupe formé des fils d'ouvriers et d'agriculteurs.

La comparaison des courbes de Lorenz généralisées, pour la période initiale (figure 2) et la période terminale (figure 3), conduit toutefois à tempérer ce constat. Si, comme l'indiquent les tests de dominance, le classement relatif des courbes évolue peu, l'écart entre les différentes courbes se réduit nettement. Sur le graphique, on constate que les courbes de Lorenz des différents groupes sociaux d'origine se rapprochent toutes de la courbe des fils de professions intermédiaires, laquelle se modifie très peu entre 1979 et 2000. Au total nous pouvons donc conclure à un changement de nature plus cardinal qu'ordinal : si les classements relatifs demeurent, le degré d'inégalité des chances diminue.

Afin d'identifier les facteurs responsables de cette tendance, nous examinons successivement l'évolution du rendement espéré (revenu moyen conditionnel au milieu d'origine) et du risque inhérent à chaque loterie.

## 4.2 Un resserrement des revenus moyens conditionnellement à l'origine sociale

Même si l'évolution des revenus moyens conditionnels à la catégorie sociale d'origine (le critère proposé par Van de Gaer (1993)) ne donne qu'une information partielle, la tendance qu'elle retrace est cependant bien en phase avec celle qui se dégage de l'examen des distributions de revenu comme le montre le tableau 3.

<b>Avant Impôts et transferts</b>					
	1979	1984	1989	1994	2000
Agriculteurs	81 096	89 282	89 403	106 392	115 034
Artisans	122 113	113 825	125 642	140 056	133 470
Cadres	170 976	173 467	176 267	165 874	161 668
Prof Inter.	135 590	141 940	132 542	133 825	136 938
Employés	111 921	119 064	128 601	124 605	117 834
Ouvriers	93 070	99 159	99 987	101 715	98 877
Revenu Moyen	104 773	110 601	113 904	123 002	120 739
<b>Après Impôts et transferts</b>					
	1979	1984	1989	1994	2000
Agriculteurs	81 487	89 214	90 800	106 329	113 361
Artisans	115 668	108 067	120 912	132 832	128 977
Cadres	158 811	157 644	160 588	152 004	151 255
Prof Inter.	128 440	132 253	126 821	126 885	129 708
Employés	108 861	114 092	123 823	120 241	117 491
Ouvriers	93 142	98 717	98 956	102 279	102 580
Revenu Moyen	102 330	107 141	110 801	118 850	119 125

En Francs constants 2000. La catégorie sociale est celle du père du ménage.

TAB. 3: Revenu moyen par origine sociale

Que l'on raisonne avant ou après impôts et transferts, un resserrement assez spectaculaire des revenus moyens conditionnels s'est opéré au cours de la période : le rapport entre le revenu moyen de la catégorie la plus favorisée (les fils de cadres) et celui de la catégorie la moins favorisée (les fils d'agriculteurs en 1979 et les fils d'ouvriers en 2000) passe, entre 1979 et 2000, de 1,95 à 1,48 en raisonnant en revenu disponible et de 2,11 à 1,63 pour le revenu primaire. L'avantage relatif des fils de cadres par rapport à la catégorie la plus désavantagée, se réduit donc de moitié. Il s'agit là d'un phénomène de grande ampleur qui interpelle en l'espace d'une génération.

Cette érosion de l'avantage des fils de cadres s'observe plus largement par rapport à tous

les autres milieux d'origine. Si l'on restreint la comparaison aux fils de salariés, on constate que le revenu disponible moyen des fils de cadres par unité de consommation a régressé sur la période étudiée de 4,8%, celui des fils de professions intermédiaires est resté pratiquement constant, celui des fils d'employés et d'ouvriers ont progressé respectivement de 7,9% et de 10,1%. La progression est donc en relation inverse avec le classement de départ des CSP d'origine. La même conclusion émerge lorsque l'on s'intéresse aux comparaisons menées en termes de revenu primaire.

Cette tendance apparaît aussi quand on compare fils de cadres et fils de non-salariés. Les différences d'ordre de grandeur sont impressionnantes : les fils d'agriculteurs améliorent leur revenu moyen de 39,1% au cours de cette période et le revenu des fils d'artisans progresse de 11,5%. En conséquence, les fils de non-salariés voient leur situation relative s'améliorer, non seulement vis à vis des fils de cadres, mais par rapport à l'ensemble des fils de salariés. Ce mouvement a contribué à resserrer les inégalités dans la période d'étude, puisque la situation de départ des fils de non-salariés était modeste : les fils d'artisans n'occupait que le troisième rang tandis que les fils d'agriculteurs fermaient la marche.

L'analyse de l'évolution du revenu des non-salariés, doit évidemment être menée avec prudence, du fait de la fragilité des données pour ce groupe, en début de période, soulignée dans la section 3.2. Il est en effet possible que les phénomènes de non-réponse conduisent, même après imputation, à sur-estimer la croissance du revenu des non-salariés, en début de période. Cependant, il faut noter que la croissance des revenus estimée dans notre échantillon pour les non-salariés est très proche de ce qu'on observe à partir de sources non-déclaratives telles que les comptes nationaux, comme le montre l'annexe 4 pour les agriculteurs. En outre, les tendances observées en longue période semblent largement imputables aux évolutions intervenues durant la décennie 1990, au cours de laquelle les données de notre enquête ont été recueillies avec plus de soin, comme en témoigne le faible pourcentage d'imputations que nous avons dû réaliser pour ces années. Au total, il est peu vraisemblable que l'érosion de la position des fils de cadres et l'ascension des fils de non-salariés soient imputables à un défaut d'enregistrement statistique. Ces deux phénomènes traduisent des évolutions importantes de l'inégalité des chances qu'il convient d'analyser.

Deux facteurs peuvent expliquer l'évolution du revenu moyen conditionnel aux différents milieux sociaux d'origine. Elle peut tout d'abord provenir d'une modification de la matrice de mobilité sociale intergénérationnelle entre catégories socio-professionnelles : par exemple, la baisse de l'avantage relatif des fils de cadres peut provenir d'une diminution de la probabilité que ces derniers deviennent eux-même cadres. Elle peut par ailleurs résulter de l'évolution des revenus moyens perçus pour chaque destination sociales possible : ainsi, même si la distribution des CSP de destination des fils d'agriculteurs reste inchangée, il est possible que le revenu moyen conditionnel de ces derniers progresse si le revenu moyen des agriculteurs progresse au cours de la période. Le premier effet provient de l'évolution de la reproduction sociale en termes de CSP, tandis que le second repose sur l'évolution de la hiérarchie des CSP de destination en termes de revenu. Nous qualifierons le premier d'effet mobilité et le second d'effet rentabilité.

<b>Avant Impôts et transferts</b>					
	1979	1984	1989	1994	2000
Agriculteurs	59 677	56 866	62 294	86 247	111 891
Artisans	107 564	93 697	128 861	139 608	127 945
Cadres	181 084	189 186	189 819	201 056	197 073
Prof Inter.	120 208	127 572	130 252	127 467	129 626
Employés	91 888	100 823	97 780	100 137	98 146
Ouvriers	77 796	80 259	81 196	83 677	84 961
Revenu Moyen	104 773	110 601	113 904	123 002	120 739

<b>Après Impôts et transferts</b>					
	1979	1984	1989	1994	2000
Agriculteurs	61 259	58 677	64 900	87 448	108 883
Artisans	99 687	87 057	118 299	128 486	119 531
Cadres	166 860	170 272	172 460	180 952	177 763
Prof Inter.	116 022	121 825	124 291	122 815	126 925
Employés	91 615	99 801	98 433	100 857	101 619
Ouvriers	80 896	84 020	85 465	88 306	92 317
Revenu Moyen	102 330	107 141	110 801	118 850	119 125

En Francs constants 2000. La catégorie sociale est celle de l'enquêté.

TAB. 4: Revenu moyen par catégorie sociale

L'étude des revenus moyens par catégorie sociale de destination (tableau 4) indique un resserrement des hiérarchies salariales au cours de la période. Le revenu moyen des cadres progresse moins vite que celui de la plupart des autres catégories socio-professionnelles

entre 1979 et 2000 et le revenu des non-salariés croit plus fortement. L'effet rentabilité explique donc, au moins partiellement, le resserrement de la distribution des revenus moyens conditionnels à l'origine sociale. On notera cependant que la réduction des disparités conditionnellement à l'origine sociale est beaucoup plus accentuée que celle que l'on peut constater sur la même période entre revenus moyens par catégorie socio-professionnelle. Ainsi, le revenu primaire moyen des cadres rapporté à celui des ouvriers est resté constant à 2,32, tandis que l'écart de revenu disponible est passé de 2,06 à 1,92. Dans le même temps, le revenu moyen primaire des fils de cadres rapporté à celui des fils d'ouvriers régresse sur la même période de 1,83 à 1,63 et le ratio des revenus disponibles moyens tombe de 1,70 à 1,47.<sup>26</sup> L'évolution des revenus par catégories socio-professionnelles ne suffit donc pas à expliquer la réduction du degré d'inégalité des chances, qui est donc en partie imputable à un effet mobilité.

L'étude de la mobilité nécessite la construction des matrices de transition (voir tableau 5 pour les matrices de 1979 et 2000, et tableau 17 pour 1984-1994). On note, au moins pour certains milieux d'origine, une augmentation de la mobilité sociale. C'est le cas notamment pour les fils de cadre dont la probabilité de devenir cadre diminue alors même que la proportion de cadres dans l'ensemble de la population a augmenté : cette probabilité passe de 62% en 1984 à 43% en 2000. Les fils d'agriculteurs connaissent quant à eux une plus grande mobilité ascendante en fin de période. Si la probabilité que ces derniers restent agriculteurs demeure stable, la probabilité qu'ils deviennent cadres ou professions intermédiaires augmente alors que la probabilité qu'ils deviennent ouvriers ou employés diminue. Pour mesurer la mobilité nette, c'est à dire pour neutraliser l'effet des changements des distributions marginales, nous utilisons l'indice de diagonalité nette (*IDN*) proposé par Forsé et Chauvel (1995) dont la formule de calcul est la suivante

$$IDN = \frac{\ln(\prod_{i < j} \frac{n_{ii}n_{jj}}{n_{ij}n_{ji}})}{k(k-1)/2 + \ln(\prod_{i < j} \frac{n_{ii}n_{jj}}{n_{ij}n_{ji}})},$$

avec  $n_{ij}$  le nombre d'individus appartenant au groupe  $j$  et issus du groupe  $i$ , et  $k$

---

<sup>26</sup>Remarquons ici, qu'en fin de période, l'écart de revenu moyen entre fils de cadres et fils d'ouvriers ne représente que la moitié de l'écart de revenu moyen entre cadres et ouvriers. Ainsi, l'inégalité des chances explique environ la moitié des inégalités de revenu moyen entre CS. Ceci ne permet toutefois pas de préciser la contribution de l'inégalité des chances à l'inégalité globale, comme nous le verrons dans la section 4.4.

<b>BdF 1979</b>	Agriculteurs	Artisans	Cadres	Prof Inter.	Employés	Ouvriers	Total
Agriculteurs	23.32	07.79	05.01	10.87	12.13	40.88	22.25
Artisans	01.03	24.61	19.62	21.57	12.40	20.78	13.90
Cadres	01.12	03.81	53.49	25.27	10.63	05.69	06.71
Prof Inter.	00.00	03.78	29.09	39.23	13.35	14.55	07.03
Employés	00.71	06.09	14.78	30.08	17.63	30.72	11.21
Ouvriers	00.56	08.94	06.49	21.24	13.15	49.63	38.90
Total	05.70	09.83	13.66	21.51	13.16	36.14	100
<b>BdF 2000</b>	Agriculteurs	Artisans	Cadres	Prof Inter.	Employés	Ouvriers	Total
Agriculteurs	24.70	05.65	10.00	16.74	07.31	35.61	11.19
Artisans	00.50	14.37	21.96	24.06	11.28	27.81	11.88
Cadres	00.63	05.21	43.10	31.03	08.83	11.20	14.91
Prof Inter.	01.31	06.97	23.36	34.75	12.69	20.92	08.26
Employés	00.95	06.18	15.43	29.01	15.50	32.93	18.17
Ouvriers	00.71	05.40	06.64	20.26	12.99	53.99	35.61
Total	03.45	06.74	17.25	24.71	11.96	35.89	100

En pourcentage, les pourcentages somment en ligne. La catégorie socioprofessionnelle du père est en ligne celle du fils en colonne. Lecture du tableau : Dans l'échantillon en 2000, 11.19% sont fils d'agriculteurs, 03.45% sont agriculteurs et 05.65% des fils d'agriculteurs deviennent artisans. Les données sont pondérées.

TAB. 5: Matrices de transition socioprofessionnelle

	1979	1984	1989	1994	2000
IDN	0.7706	0.7468	0.7299	0.7066	0.7029

Calculé sur les matrices de transition socioprofessionnelles.  
Données pondérées.

TAB. 6: Indice de Diagonalité Nette

le nombre de groupes sociaux. L'indice de diagonalité nette calculé pour les matrices de transition socioprofessionnelles indique une hausse sensible de la mobilité nette<sup>27</sup>.

Une manière de quantifier la contribution respective de l'effet mobilité et celle de l'effet rentabilité utilise une décomposition de l'évolution du revenu conditionnel moyen inspirée d'Oaxaca (1973). Le revenu moyen conditionnel s'exprime en effet sous la forme :

$$\bar{y}_s^t = \sum_k \alpha_{sk}^t \bar{y}_{sk}^t$$

où  $\bar{y}_s^t$  est le revenu moyen des individus issus de la CSP  $s$  l'année  $t$ ,  $\alpha_{sk}^t$  est la proportion d'individus issus de la CSP  $s$  qui aboutissent à la CSP  $k$  l'année  $t$  et  $\bar{y}_{sk}^t$  est le revenu

<sup>27</sup>C'est semble t-il une tendance de long terme puisque Thélot (1983) concluait déjà à une légère augmentation de la mobilité nette.



moyen des individus de CSP  $k$ , issus de la CSP  $s$  l'année  $t$  <sup>28</sup>.

La variation du revenu moyen des individus issus de la CSP  $s$  entre  $t$  et  $t'$  peut se décomposer des deux façons suivantes :

$$\Delta \bar{y}_s = \Delta \alpha_s \bar{y}_s^t + \Delta \bar{y}_s \alpha_s^{t'} \quad (1)$$

$$\Delta \bar{y}_s = \Delta \alpha_s \bar{y}_s^{t'} + \Delta \bar{y}_s \alpha_s^t \quad (2)$$

où  $\alpha_s$  et  $\bar{y}_s$  désignent respectivement les vecteurs des probabilités des différentes destinations sociales et des revenus moyens dans ces destinations, pour les individus issus de la catégorie  $s$ .

Le premier effet est un effet mobilité. Il mesure quel aurait été l'effet d'un changement de la matrice de mobilité sociale entre CSP pour une distribution de revenus  $\bar{y}_{sk}$  inchangée. Le second effet est un effet rentabilité. Il représente l'effet d'une variation des  $\bar{y}_{sk}$  sur le revenu moyen conditionnel à l'origine, pour une matrice de mobilité constante. Ces deux effets peuvent être mesurés de façon différente, selon qu'on prend pour référence la matrice de mobilité ou la distribution de revenus conditionnels de la date initiale. Les résultats de ces deux décompositions, pour le revenu moyen avant impôt et transferts sont donnés dans le tableau 7, et dans le tableau 16 en annexe pour le revenu moyen après impôts et transferts.

L'intuition dont il a été fait état plus haut pour les fils de cadres se trouve confirmée par l'analyse de décomposition. L'effet mobilité et l'effet rentabilité contribuent tous les deux à une baisse du revenu moyen de cette catégorie. Toutefois, l'effet dominant d'un point de vue quantitatif est ici l'effet mobilité. L'érosion de l'avantage relatif des fils de cadres provient donc essentiellement du fait que le statut de cadre se transmet moins facilement à la génération suivante que par le passé. Une explication assez immédiate de ce phénomène tient dans la dilution progressive du statut de cadre, liée à l'élargissement de la population des cadres dans la société française.

Dans le cas des fils d'agriculteurs, l'effet mobilité<sup>29</sup> et l'effet rentabilité contribuent

<sup>28</sup>Notons que cette décomposition fait intervenir le revenu moyen doublement conditionné par l'origine et la destination sociale, alors que les commentaires des paragraphes précédents portaient sur le revenu moyen conditionnel à la seule destination sociale.

<sup>29</sup>Ces résultats sont à mettre en rapport avec ceux de Goldthrope (1995) qui trouve une amélioration de

	Première décomposition						Seconde décomposition					
	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Ecart 1979-2000	33 938	11 357	-9 308	1 348	5 913	5 806	33 938	11 357	-9 308	1 348	5 913	5 806
Effet structure	29 642 87.34%	9 006 79.30%	-1 375 14.78%	7 827 580.65%	5 529 93.51%	7 329 126.23%	26 675 78.60%	12 907 113.65%	-786 08.45%	7 225 536.02%	6 011 101.65%	7 357 126.71%
Effet mobilité	4 296 12.66%	2 351 20.70%	-7 932 85.22%	-6 479 -480.65%	384 06.49%	-1 523 -26.23%	7 263 21.40%	-1 550 -13.65%	-8521 91.55%	-5 877 -436.02%	-98 -01.65%	-1 551 -26.71%
Écart 1979-1989	8 307	3 529	5 291	-3 049	16 680	6 917	8 307	3 529	5 291	-3 049	16 680	6 917
Effet structure	6 955 83.72%	6 241 176.86%	3 242 61.27%	885 -29.02%	14 188 85.06%	7 178 103.77%	6 342 76.34%	9 922 281.19%	971 18.36%	2 270 -74.45%	12 838 76.96%	7 073 102.25%
Effet mobilité	1 353 16.28%	-2 712 -76.86%	2 049 38.73%	-3 933 129.02%	2 492 14.94%	-260 -03.77%	1 966 23.66%	-6 394 -181.19%	4 320 81.64%	-5 318 174.45%	3 843 23.04%	-156 -02.25%
Écart 1989-2000	25 631	7 828	-14 599	4 397	-10 767	-1 111	25 631	7 828	-14 599	4 397	-10 767	-1 111
Effet structure	21 892 85.41%	4 192 53.55%	-2 136 14.63%	6 655 151.36%	-6 754 62.72%	294 -26.47%	20 629 80.49%	3 269 41.76%	-1 133 07.76%	6 692 152.22%	-8 481 78.77%	129 -11.62%
Effet mobilité	3 738 14.59%	3 636 46.45%	-12 463 85.37%	-2 258 -51.36%	-4 014 37.28%	-1 405 126.47%	5 002 19.51%	4 559 58.24%	-13 466 92.24%	-2 296 -52.22%	-2 286 21.23%	-1 240 111.62%

L'écart est un écart de revenu moyen, en francs constants 2000. La catégorie sociale est celle du père. La première décomposition correspond à l'équation (1), la seconde à l'équation (2).

TAB. 7: Décomposition de l'évolution du revenu moyen - Avant Impôts et transferts

tous les deux positivement à la progression de leur position sociale relative. Cependant, la majeure partie de cette progression s'explique ici par l'effet rentabilité. Le revenu des agriculteurs augmente très fortement au cours de la période 1979-2000 (+87,4%), l'essentiel de la hausse intervenant dans la décennie 90. Ceci bénéficie alors largement aux 25% de fils d'agriculteurs devenant eux-même agriculteurs.

De la même façon, une grande partie de la hausse du revenu des fils d'artisans découle de la progression notable du revenu des artisans (environ 20%), plus rapide que celle des catégories de salariés. L'effet mobilité est ambigu. Il est positif si l'on raisonne par rapport à la structure initiale des revenus par CSP (décomposition 1) et négatif si l'on raisonne par rapport à la structure terminale des revenus par CSP (décomposition 2). Cette sensibilité à la structure des revenus et à la table de mobilité de référence résulte d'une double évolution. Au cours de la période, les fils d'artisans accèdent plus fréquemment aux catégories salariées. Dans le même temps, le revenu moyen des artisans progresse plus rapidement et dépasse celui des principales catégories sociales salariées de destination (employés, ouvriers) ou s'en rapproche nettement (professions intermédiaires). Si on valorise l'effet de la plus grande mobilité des artisans vers le salariat, à partir de la structure de revenus du début de période, on trouve alors un effet positif alors que l'inversion des hiérarchies salariales en fin de période conduit à un effet de mobilité négatif, si on utilise la structure terminale des rémunérations moyennes.

Enfin, toutes les catégories de salariés connaissent un effet mobilité négatif ou nul. L'effet mobilité pour les catégories intermédiaires est d'ailleurs d'une ampleur assez comparable à celui constaté pour les cadres. Par contre, l'effet rentabilité est positif, traduisant le fait que le revenu des catégories modestes de salariés a continué de progresser plus vite que celui des cadres.

Il est enfin instructif de comparer le revenu moyen des individus issus d'une catégorie sociale donnée au revenu qu'ils auraient obtenu s'ils étaient restés dans leur catégorie socio-professionnelle d'origine. Les résultats de cet exercice apparaissent dans le tableau 8. L'ensemble des milieu sociaux d'origine bénéficie globalement de la mobilité sociale, à

---

la situation des fils d'agriculteurs, tout en observant une forte immobilité globale des différentes catégories socioprofessionnelles. Goux et Maurin (1995) trouvent notamment que la situation des fils d'agriculteurs s'améliore en termes de réussite scolaire.

	1979		2000	
	Avant Impôts	Après Impôts	Avant Impôts	Après Impôts
Agriculteurs	1.36	1.33	1.03	1.04
Artisans	1.14	1.16	1.04	1.08
Cadres	0.94	0.95	0.82	0.85
Prof Inter	1.13	1.11	1.06	1.02
Employés	1.22	1.19	1.20	1.16
Ouvriers	1.20	1.15	1.16	1.11

En francs constants 2000.

TAB. 8: Rapport des revenus moyens conditionnels et non conditionnels

l'exception des descendants de cadres. En fin de période, ce sont les fils d'employés qui profitent le plus, en termes de revenu, de la mobilité sociale. Pour les fils de cadres, le revenu moyen observé est plus faible que le revenu moyen qu'ils auraient perçu en restant dans leur milieu social d'origine. Par ailleurs, pour cette catégorie, les pertes de revenu moyen liées à la mobilité sociale augmentent entre 1979 et 2000. Pour les autres milieux d'origine, les gains de revenus moyen tirés de la mobilité sociale diminuent au cours de la période étudiée, du fait du resserrement des revenus moyens par catégorie socio-professionnelle de destination.

### 4.3 Des loteries sociales presque aussi risquées

Pour compléter cette analyse de l'inégalité des chances, il convient enfin d'examiner le degré de risque associé à chacune des loteries sociales observées. Ce degré de risque peut être étudié à l'aide des courbes de Lorenz. Ces courbes s'obtiennent à partir des courbes de Lorenz généralisées, en normalisant la distribution de revenu, pour un milieu social donné, par le revenu moyen des individus originaires de ce milieu social. Si la courbe de Lorenz pour une origine sociale donnée est toujours au dessus de la courbe de Lorenz d'un second milieu social, alors la loterie sociale associée au premier groupe est moins risquée que celle du second groupe. Si les deux courbes sont confondues, alors les deux loteries sociales correspondantes sont également risquées. La même conclusion prévaut au sens faible si les courbes de Lorenz se coupent.

Les tableaux 12 et 13 en annexe présentent les résultats des tests de dominance pour les courbes de Lorenz relatives. La figure 4 page 38 présente les courbes de Lorenz pour l'année

1979. Les courbes de Lorenz des différentes loteries sociales apparaissent profondément similaires. Les tests de dominance concluent à l'égalité au sens fort des distributions relatives avant impôts et transferts dans 10 comparaisons (sur 15) en 1979, dans 9 cas en 1984, dans 13 cas en 1989, 7 en 1994 et 11 en 2000. On ne conclut à la dominance que dans 1 cas en 1979, 5 cas en 1984 et 1994, et 3 cas en 2000. Le degré de risque associé aux différentes loteries sociales de revenu primaire semble donc globalement très proche, pour les différents milieux d'origine.

En première approximation, ces résultats suggèrent que la formation du revenu, conditionnellement à un milieu d'origine donné, peut être représenté par un modèle multiplicatif de la forme :

$$x_{is} = E(x | s) \times \epsilon_i \quad (3)$$

où  $x_{is}$  désigne le revenu de l'individu  $i$  dont le milieu d'origine est  $s$ ,  $E(x | s)$  est l'espérance de revenu conditionnel et  $\epsilon_i$  un terme aléatoire distribué suivant une loi qui ne dépend pas de l'origine sociale.

Avant Impôts et transferts

Après Impôts et transferts

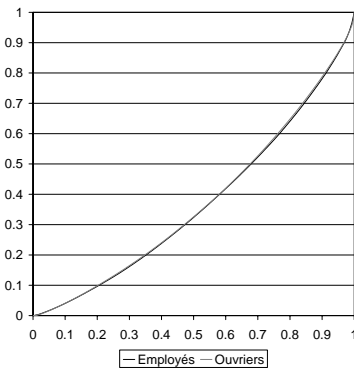
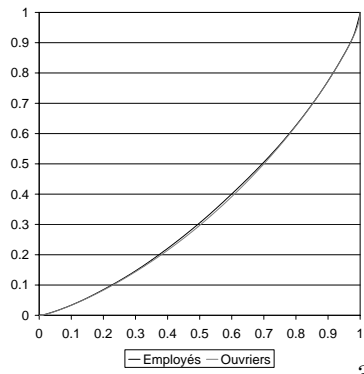
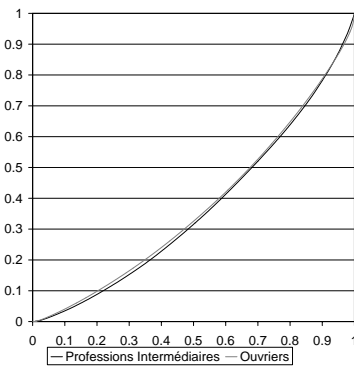
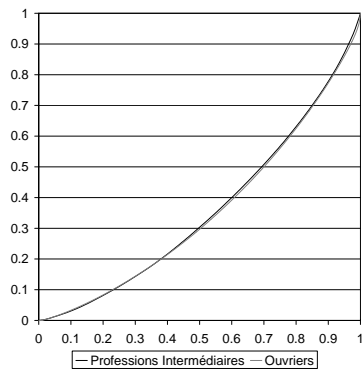
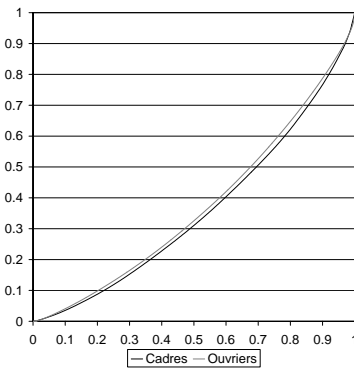
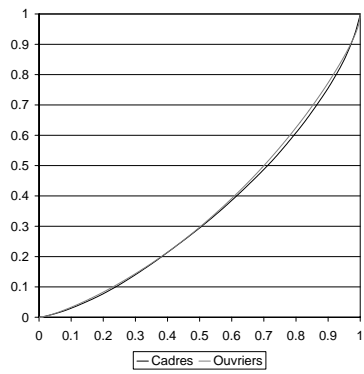
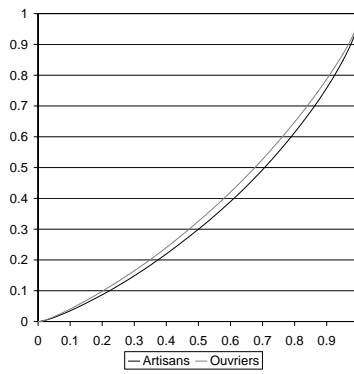
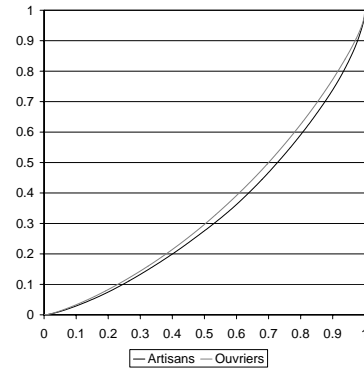
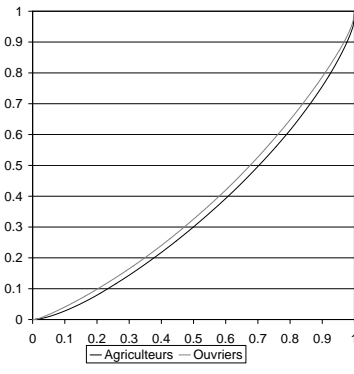
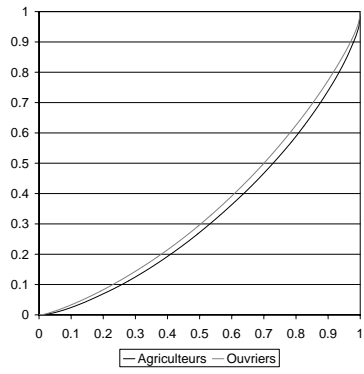
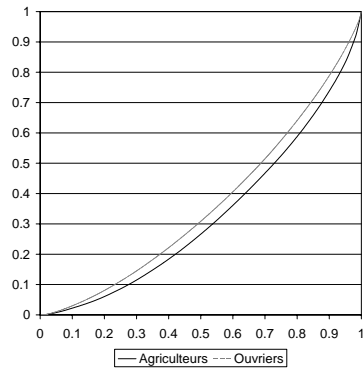


FIG. 4: Courbes de Lorenz en 1979

Avant Impôts et transferts



Après Impôts et transferts

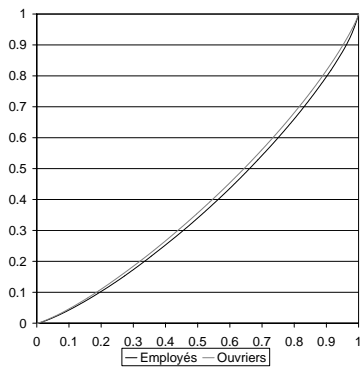
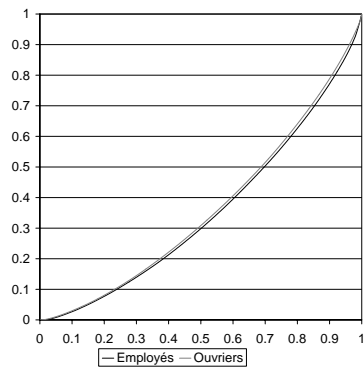
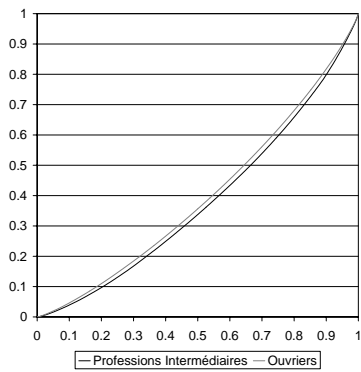
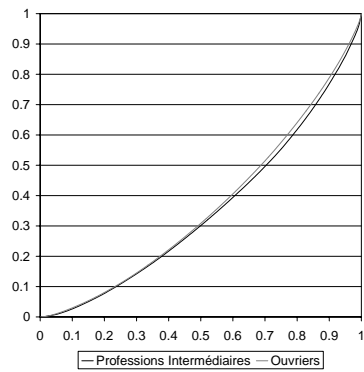
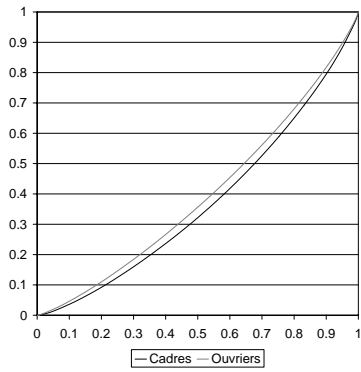
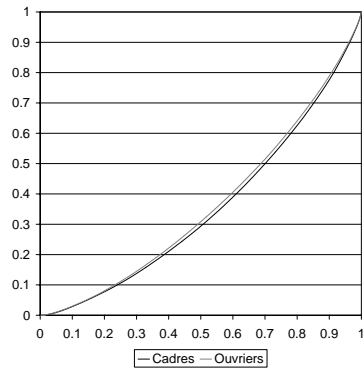
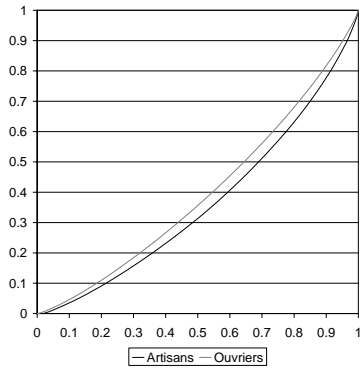
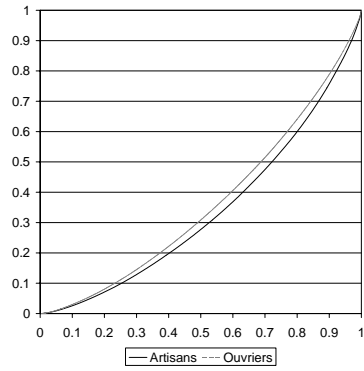
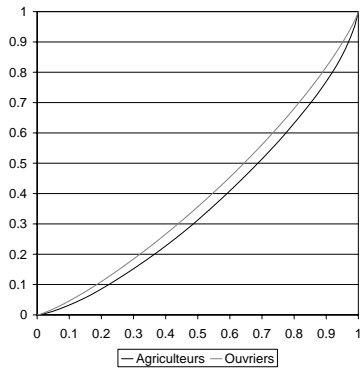


FIG. 5: Courbes de Lorenz en 2000

	Avant Impôts					Après Impôts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Écart min	0.000	0.004	0.005	0.008	0.000	0.003	0.001	0.002	0.000	0.006
Écart max	0.053	0.059	0.048	0.061	0.068	0.046	0.049	0.066	0.061	0.070

Les indices de Gini sont calculés sur les distributions conditionnelles. La différence entre les indices de Gini est calculées sur toutes les comparaisons par paires. On garde la plus grande et la plus petite différence en valeur absolue.

TAB. 9: Écart entre les courbes de Lorenz en Points de Gini

L'étude des distributions relatives de revenu après redistribution ne modifie que peu cette conclusion d'ensemble. Elle laisse toutefois apparaître quelques éléments nouveaux. Après redistribution, la courbe de Lorenz relative des ouvriers domine celle de toutes les autres catégories, à l'exception des fils d'employés. Ainsi la distribution dont le rendement est le plus bas est aussi celle dont le degré de risque est le plus faible. Au contraire, la loterie la plus performante, celle des fils de cadres, est également plus risquée. Enfin il apparaît que les loteries les plus risquées, après impôts et transferts sont celles de fils de non-salariés, ce qui semble conforme à l'intuition. C'est justement pour les fils d'agriculteurs et d'artisans que le classement en termes de risque peut contredire le classement en termes de rendement. Par exemple, en 2000, le revenu disponible moyen des fils d'agriculteurs est plus élevé de 10% que celui des fils d'ouvriers. Toutefois, comme nous avons eu l'occasion de le souligner, le test de dominance se révèle inconclusif, en raison justement du plus grand degré de risque de la loterie des fils d'agriculteurs. Le tableau 9 donne les écarts maximum et minimum entre les courbes de Lorenz exprimés en points de Gini.

Le fait que les loteries soient presque aussi risquées impliquent également que se prononcer sur l'égalité des chances en comparant les revenus moyens conditionnels, comme le propose Van de Gaer, n'induirait sur les données françaises que des erreurs de faible importance. Ce n'est pas souvent que dans une analyse statistique, la comparaison des moyennes suffit pour pouvoir porter un jugement valable sur des distributions. Cette caractéristique nous permet finalement d'offrir une comparaison simple entre inégalité des chances et inégalité des résultats.



#### 4.4 L'inégalité des chances compte t-elle pour beaucoup dans l'inégalité totale ?

L'inégalité des chances ne compte que pour une fraction de l'inégalité des résultats au sens où si nous procédions à l'expérience mentale de gommer l'inégalité des chances, l'inégalité des résultats serait réduite de 1 moins cette fraction. Nous proposons deux quantifications simples de cette fraction. La première utilise l'indice d'inégalité le plus utilisé, l'indice de Gini, la seconde la décomposition de la variance qui est l'outil favori des économètres.

Admettons de quantifier l'inégalité des revenus primaires et des revenus disponibles au moyen de l'indice de Gini et supposons qu'il soit possible d'égaliser les revenus moyens conditionnels, grâce à une transformation homothétique des revenus des fils d'une même CS d'origine. Compte tenu de l'importante similarité des courbes de Lorenz relatives mise en évidence dans la section précédente, ceci conduit, en première approximation, à égaliser les distributions conditionnelles et donc à assurer l'égalité des chances. Il est alors facile de calculer l'inégalité résiduelle, qui correspond à la courbe de Lorenz des revenus centrés à leur moyenne conditionnelle, tous individus confondus. La différence entre l'inégalité résiduelle et l'inégalité des résultats est la contribution de l'inégalité des chances à l'inégalité totale. Le tableau 10 nous livre une contribution de l'ordre de 2 points de Gini, soit 5% de l'inégalité totale. Bourguignon *et alii* (2003) sur données brésiliennes par des méthodes différentes et avec un conditionnement plus riche obtiennent une contribution de 5 à 12 points de Gini soit entre 10% et 20% de l'inégalité totale Brésilienne.

TAB. 10: Part de l'inégalité des chances dans l'inégalité globale

	Avant Impôts					Après Impôts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Gini global	0.332	0.318	0.326	0.323	0.313	0.296	0.271	0.279	0.267	0.250
Gini résiduel	0.310	0.303	0.309	0.309	0.297	0.276	0.258	0.263	0.255	0.236
Différence	0.022	0.015	0.017	0.014	0.016	0.020	0.013	0.015	0.012	0.013
Part	6.56	4.71	5.24	4.46	5.08	6.64	4.89	5.50	4.60	5.33

L'indice de Gini global est calculé sur tout l'échantillon. Le Gini résiduel est calculé en normalisant les distributions conditionnelles par leur moyenne.

La seconde façon de procéder consiste à utiliser une analyse de régression, à la fois pour tester la forme multiplicative que l'analyse de la section 4.3 suggère, et également pour mesurer la contribution de l'inégalité des chances dans l'inégalité totale. Pour cela on exprime la relation (4.3) en logarithme, et on écrit qu'en première approximation, le rôle des circonstances à la Roemer s'exprime sous une forme Galtonienne avec un coefficient  $\beta$  égal à 1 qui évoque le modèle de Gibrat où le rôle des circonstances quantifié par l'espérance de revenu conditionnelle joue le rôle du revenu d'une période passée, et un terme d'erreur  $\eta_i$  qui est un bruit blanc, soit :

$$\log x_{is} = \beta \log E(x | s) + \eta_i, \quad (4)$$

avec  $\beta = 1$

Le tableau 11 indique que l'estimation de cette relation sur tous les individus conduit à accepter l'hypothèse que  $\beta = 1$  et à rejeter l'hypothèse de normalité du résidu à 1% pour toutes les régressions. Enfin, pour savoir si les résidus sont identiquement distribués entre les différentes origines sociales, on réalise des tests de Kolmogorov-Smirnov pour chaque paires de distributions des résidus. Sur 150 tests, on rejette l'hypothèse que les résidus sont identiquement distribués 24 fois<sup>30</sup>, soit dans 16% des cas seulement. Au total cette spécification multiplicative semble assez robuste, et les différents tests menés vont plutôt dans le même sens.

On utilise alors l'hypothèse d'indépendance du résidu  $\eta_i$  pour décomposer la variance des logarithmes du revenu en la variance des revenus moyens conditionnels et la variance des résidus soit

$$\sigma^2(\log x_{is}) = \sigma^2(\log E(x | s)) + \sigma_\eta^2. \quad (5)$$

Cette écriture nous permet d'offrir une décomposition de l'inégalité des revenus entre une composante due aux circonstances et une composante résiduelle pour toutes les années, avant et après redistribution. La valeur de  $R^2$  dans la régression donne exactement la

---

<sup>30</sup>7 fois avant impôts et 17 fois après impôts.

TAB. 11: Régression du revenu sur le revenu moyen conditionnel

	Avant Impôts					Après Impôts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
$\hat{\beta}$	1.060	1.088	0.973	0.959	0.975	1.045	1.040	0.935	0.910	0.893
Écart-Type	(.046)	(.052)	(.055)	(.061)	(.060)	(.047)	(.052)	(.054)	(.056)	(.057)
Obs.	3914	3958	3152	3992	3331	3914	3958	3152	3992	3331
$R^2$	.11	.09	.08	.05	.06	.11	.09	.08	.06	.06
Test sur le coefficient $\hat{\beta}=1$										
$\hat{w}$	1.67	2.84	.22	.43	.14	.93	.59	1.41	2.49	3.46
proba. critique	.196	.091	.639	.511	.708	.334	.442	.235	.114	.063
Tests de Normalité du résidu										
Proba. critique	.010	.010	.010	.010	.010	.010	.010	.010	.010	.010

La régression est en logarithme. Pour tester la normalité du résidu, on utilise le test non paramétrique de Kolmogoro-Smirnov. L'hypothèse de normalité est rejetée à 1%. Le test  $\beta = 1$  est effectué avec une statistique de Wald,  $\hat{w} \sim \chi_1^2$ , on ne rejette pas l'hypothèse d'égalité.

contribution relative de la CSP du père à l'inégalité totale et cette part diminue de 11% à 6%. Cette seconde façon de procéder nous conduit à des résultats très similaires, à savoir une part très faible de l'inégalité des chances, telle que nous l'avons mesuré, dans l'inégalité totale.

La seule circonstance dont nous avons cherché à mesurer l'impact est la CS du père. Il est certain que bien d'autres circonstances jouent un rôle dans la réussite ou l'échec de la génération suivante et nous sous-estimons donc la part de l'inégalité des chances dans l'inégalité totale. Les chiffres dont nous avons fait état plus haut doivent donc être considérés comme des bornes minimales. Il est par exemple instructif de rapprocher les résultats obtenus de ceux que l'on obtiendrait si on appliquait la même décomposition de la variance des logarithmes au conditionnement étudié par Lefranc et Trannoy (2003) qui retiennent non pas un déterminisme social mais un simple déterminisme individuel à travers le revenu des ascendants. L'objet de leur étude est d'étudier l'impact du revenu du père sur le revenu du fils en France. La valeur estimée du coefficient lorsqu'on régresse le log du revenu du fils par rapport à celui du père est de l'ordre de 0,4. Sous l'hypothèse que l'inégalité reste constante d'une génération à une autre (la variance des log des revenus reste constante), ce qui peut être admis en première approximation pour la France, cela conduit à une contribution des inégalités de la circonstance "revenu du père" de l'ordre de

16%. Cette contribution serait en première approximation constante entre 1964 et 1993. Cette estimation, qui reste très grossière, constitue sans doute un point de référence utile pour la discussion de nos résultats qui est entreprise dans la conclusion.

## 5 Conclusion

Nous avons choisi de mesurer l'égalité des chances d'une manière non paramétrique en adoptant le point de vue de la dominance. Sur les données françaises, ce choix se révèle heureux, puisque nous avons été en mesure de conclure presque dans tous les cas. Retenir une démarche en terme d'indices, que l'on sait moins robuste, ne semble pas donc pertinent, dans cette étude.

Le rôle de l'héritage social et économique, transmis par le milieu d'origine, reste donc prégnant dans la formation des inégalités, en France depuis 1979. Même en fin de période, les différences d'origine sociale conduisent encore à des différences significatives de niveau de vie et l'égalité des chances dans l'obtention du revenu ne semble pas prévaloir, tant en ce qui concerne le revenu primaire qu'en matière de revenu disponible après impôts et transferts redistributifs<sup>31</sup>. Toutefois, les processus de mobilité sociale à l'œuvre contribuent partiellement à une redistribution des opportunités économiques, qui trouve vraisemblablement sa source dans les politiques et les choix éducatifs, dans le fonctionnement du marché du travail et dans les appariements matrimoniaux. Seule une partie de l'inégalité de revenus se transmet donc d'une génération à l'autre. De ce point de vue, les enseignements de notre analyse de l'égalité des chances dans le niveau de vie des ménages rejoignent les conclusions qu'on peut tirer d'une analyse de la corrélation intergénérationnelle des seuls revenus salariaux (voir Lefranc et Trannoy (2003)). Par ailleurs, la politique redistributive contribue elle aussi à réduire l'inégalité des chances mais dans une mesure relativement modeste : l'écart maximal de revenu moyen, conditionnellement à l'origine sociale, étant de 47% en termes de revenu disponible, contre 63% en termes de revenu primaire. Enfin, une tendance à la réduction du degré d'inégalité des chances se dessine. Ainsi, le milieu social le plus avantageux, celui des enfants de cadres, voit son avantage relatif par rapport à

---

<sup>31</sup>Cette première conclusion recoupe par exemple celle de Vallet (2001) qui constate une forte stabilité des positions sociales.

la catégorie la moins favorisée diminuer de moitié, au cours de la période 1979-2000.

Cette évolution contraste avec les changements observés au cours des vingt dernières années dans l'inégalité des revenus individuels. En France, l'inégalité des conditions de vie est en effet restée globalement stable depuis vingt ans (voir le constat dressé dans Atkinson *et alii* (2001)). Dans d'autres pays, et notamment aux États-Unis et en Grande-Bretagne, on a assisté pendant la même période à une forte augmentation des inégalités de revenu (voir par exemple Lefranc (1997)). Dans ce contexte, le constat d'une réduction de l'inégalité des chances établi dans cet article, et qui rejoint les conclusions d'autres études (voir notamment Goux et Maurin (2003)), ne manque donc pas de surprendre et il convient de s'interroger sur le degré de robustesse de ce résultat.

Il importerait notamment de s'interroger sur la possibilité d'une dilution progressive du statut social des cadres, corrélative de la montée en puissance de cette catégorie dans la population active française. On peut en effet penser que l'augmentation de la part des cadres dans la société a conduit à une plus grande hétérogénéité des positions sociales au sein de cette catégorie et à un affaiblissement de sa position relative dans l'échelle sociale. On a par exemple assisté depuis 40 ans en France à une diminution marquée du salaire des cadres relativement au salaire des ouvriers : entre 1960 et 1988 le rapport du salaire de ces deux catégories est ainsi passé de 4 à 3. Cette hétérogénéité accrue pourrait alors expliquer l'érosion de l'avantage relatif des fils de cadres observée dans notre étude. Il importerait alors de préciser la contribution respective de la réussite scolaire, de l'insertion au marché du travail et du marché matrimonial à cet affaiblissement.

On peut d'ailleurs penser que loin de se limiter aux seuls cadres, cette hétérogénéité croissante des positions individuelles au sein d'une catégorie socio-professionnelle donnée, a touché l'ensemble des milieux sociaux. L'augmentation des inégalités salariales intra-groupes observée dans les économies développées plaide notamment en faveur de cette interprétation. Cette diversification des positions individuelles au sein des différentes CSP rendrait alors moins pertinent le conditionnement par la CSP d'origine. Cette moindre influence de la CSP d'origine sur les destinations individuelles mérite toutefois d'être interprétée avec prudence. Elle peut certes traduire une réduction de l'inégalité des chances, dès lors qu'elle témoignerait d'une redistribution des opportunités économiques au sein

de chaque CSP, d'une génération à l'autre. Il n'est toutefois pas exclu que l'hétérogénéité croissante des positions individuelles, se soit accompagnée d'une différenciation des processus de transmission sociale intergénérationnelle interne aux différents groupes sociaux. En étudiant la formation du revenu conditionnellement à un découpage relativement agrégé des positions sociales, on pourrait alors ne pas saisir l'éventuelle permanence du rôle de l'héritage socio-économique, pour un certain nombre de sous-catégories sociales. Il n'est alors pas dit qu'une appréhension plus fine du milieu socio-économique d'origine validerait le constat dressé ici d'une réduction de l'inégalité des chances. Dans ce cas, la réduction de l'inégalité des chances, conditionnellement à la CSP d'origine, témoignerait seulement de la moindre capacité des variables de CSP à appréhender les dimensions les plus pertinentes de la transmission intergénérationnelle du statut économique. L'analyse de la corrélation intergénérationnelle des niveaux de revenus plaide d'ailleurs pour cette interprétation. Elle indique que la transmission intergénérationnelle du revenu est restée relativement stable depuis 25 ans (Lefranc et Trannoy (2003)). Il importerait donc d'enrichir la prise en compte de l'origine sociale dans l'évaluation des relations de dominance, notamment par la prise en compte du niveau d'éducation des parents, du décile de revenu d'origine et par un découpage plus fin des positions sociales d'origine. Si cette voie n'a pu être empruntée dans cette étude, compte tenu de la taille relativement faible de nos échantillons, elle mériterait d'être suivie dans des recherches futures.

Enfin cet article a aussi mis en évidence un phénomène marquant : le risque des loteries sociales apparaît très similaire d'un groupe à l'autre. En première approximation, l'origine socioprofessionnelle semble donc jouer, dans la formation du revenu, d'une façon multiplicative : le revenu d'un individu est décrit par le produit d'une variable aléatoire - distribuée indépendamment de la CSP du père - et d'un coefficient qui dépend de l'origine familiale et qui majore au plus de 50% le niveau de revenu. Il conviendrait aussi, dans des travaux ultérieurs, de s'interroger sur les implications théoriques de ce fait stylisé important.

## Références

Arneson, R. (1989). Equality and equal opportunity of welfare, *Philisophical Studies* 56 : 77-93.

- Atkinson, A. B., Glaude, M., Olier, L. et Piketty, T. (2001). *Inégalités économiques*, La Documentation Française, Paris. Rapport pour le Conseil d'Analyse Economique.
- Beach, C. M. et Davidson, R. (1983). Distribution-free statistical inference with lorenz curves and income shares, *Review of Economic Studies* pp. 723–735.
- Bishop, J. A., Formby, J. P. et Thistle, P. D. (1992). Convergence of the south and the non-south income distributions, 1969-1979, *American Economic Review* **82** : 262–272.
- Bossert, W. (1995). Redistribution mechanisms based on individual characteristics, *Mathematical Social Sciences* **29(1)** : 1–17.
- Bourguignon, F. (1998). *Fiscalité et Redistribution*, La Documentation Française, Paris. Rapport pour le Conseil d'Analyse Economique.
- Bourguignon, F., Ferreira, F. H. et Menendez, M. (2003). Inequality of outcomes and inequality of opportunities in brasil, *DELTA Working Paper n° 2003-24* .
- Chopin, N. et Masse, E. (2002). Imputation de l'enquête budget de famille 2000, *Technical report*, INSEE.
- Cohen, G. (1989). On the currency of egalitarian justice, *Ethics* **99** : 906–944.
- Dardanoni, V. et Forcina, A. (1999). Inference for lorenz curve orderings, *Econometrics Journal* **2** : 49–75.
- Davidson, R. et Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* **68** : 1435–1464.
- Dworkin, R. (1981). What is equality. part 1 : Equality of welfare, *Philosophy and Public Affairs* **10** : 185–246.
- Fleurbaey, M. (1995). On fair compensation, *Theory and Decision* **36** : 277–307.
- Forsé, M. et Chauvel, L. (1995). L'évolution de l'homogamie en france. une méthode pour comparer les diagonalités de plusieurs tables, *Revue Française de Sociologie* **36(1)** : 123–143.
- Foster, J. et Shorrocks, A. F. (1988). Poverty orderings and welfare dominance, *Social Choice and Welfare* **5** : 179–198.
- Goldthorpe, J. H. (1995). Le « noyau dur » : fluidité sociale en angleterre et en france dans les années 70 et 80, *Revue Française de Sociologie* **36(1)** : 61–79.
- Goux, D. et Maurin, E. (1995). Origine sociale et destinée scolaire, *Revue Française de Sociologie* **36(1)** : 81–123.
- Goux, D. et Maurin, E. (2003). On the evaluation of equality of opportunity for income : Axioms and evidence, *CREST Working Paper* .
- Howes, S. (1993). Asymptotic properties of four fundamental curves of distributional analysis. Unpublished Paper, STICERD, London School of Economics.
- Kodde, D. A. et Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions, *Econometrica* **54** : 1243–1248.
- Lefranc, A. (1997). Evolution des marchés du travail français et américains : quelques éléments d'analyse comparative, *Revue Economique* **48(5)** : 1041–1060.
- Lefranc, A. et Trannoy, A. (2003). Intergenerational earnings mobility in france : Is france more mobile than us? Mimeo THEMA.
- Lollivier, S. (1997). Modèles univariés et modèles de durée sur données individuelles, *Technical report*, INSEE.

- McDonald, J. B. et Xu, Y. J. (1995). A generalization of the beta distribution with applications, *Journal of Econometrics* **66** : 133–152.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets, *International Economic Review* **14** : 693–709.
- Ooghe, E., Schokkaert, E. et Van de Gaer, D. (2003). Equality of opportunity versus equality of opportunity sets. Public Economics Working Paper Series, Katholieke Universiteit Leuven.
- Roemer, J. (1998). *Equality of Opportunity*, Harvard University Press, Cambridge.
- Roemer, J., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S., Lefranc, A., Marx, I., Page, M., Pommer, E., Ruiz-Castillo, J., San Segundo, M., Tranaes, T., Trannoy, A., Wagner, G. et Zubiri, I. (2003). To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?, *Journal of Public Economics* **87** : 539–565.
- Shorrocks, A. F. (1983). Ranking income distributions, *Economica* **50** : 3–17.
- Thélot, C. (1983). L'évolution de la mobilité sociale dans chaque génération, *Économie et Statistique* **161**.
- Thélot, C. (1982). *Tel père, tel fils ? Position sociale et origine familiale*, Dunod, Paris.
- Vallet, L.-A. (2001). Change in intergenerational class mobility in france from the 1970s to the 1990s and its explanation : An analysis following the casmin approach. Cahiers du LASMAS.
- Van de Gaer, D. (1993). Equality of opportunity and investment in human capital, *Unpublished Manuscript, K.U. Leuven* .
- Wolak, F. A. (1989). Testing inequality constraints in linear econometric models, *Journal of Econometrics* **41** : 205–235.



<b>1979</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	?	=	=	=	<
Artisans	-	-	=	?	?	?
Cadres	-	-	-	=	=	=
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>1984</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	?	<	<	<
Artisans	-	-	=	< <sub>1</sub>	=	=
Cadres	-	-	-	=	=	=
Prof Interm	-	-	-	-	=	>
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>1989</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	?	=	=	?
Artisans	-	-	=	=	=	=
Cadres	-	-	-	=	=	=
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>1994</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	=	<	<	=
Artisans	-	-	=	<	<	<
Cadres	-	-	-	?	=	?
Prof Interm	-	-	-	-	?	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>2000</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	<	?	<	<
Artisans	-	-	=	=	=	=
Cadres	-	-	-	=	=	=
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-

Le tableau donne la situation de la catégorie en ligne par rapport à la colonne  
> : La ligne domine la colonne.  
< : La ligne est dominée par la colonne.  
= : Égalité des courbes de Lorenz.  
? : Non comparabilité.

TAB. 12: Tests de dominance relative - Avant Impôts et transferts

## Annexes

### 5.1 Résultats complémentaires

<b>1979</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	?	=	=	<	<
Artisans	-	-	=	=	=	=
Cadres	-	-	-	=	=	?
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>1984</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	=	<	<	<
Artisans	-	-	=	?	=	=
Cadres	-	-	-	=	=	=
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>1989</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	=	=	?	<
Artisans	-	-	=	=	?	?
Cadres	-	-	-	=	?	?
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	?
Ouvriers	-	-	-	-	-	-
<b>1994</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	=	<	<	<
Artisans	-	-	=	<	<	<
Cadres	-	-	-	?	<	<
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
<b>2000</b>	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Agriculteurs	-	=	=	?	<	<
Artisans	-	-	=	=	=	<
Cadres	-	-	-	=	=	<
Prof Interm	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-

Le tableau donne la situation de la catégorie en ligne par rapport à la colonne

> : La ligne domine la colonne.

< : La ligne est dominée par la colonne.

= : Égalité des courbes de Lorenz.

? : Non comparabilité.

TAB. 13: Tests de dominance relative - Après Impôts et transferts

Année	Agriculteurs					Artisans					Cadres				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	103	104	92	143	124	275	158	160	285	243	570	527	438	763	596
Obs Imputées	115	101	58	0	0	166	127	86	42	0	95	64	51	3	0
Total	218	205	150	143	124	441	285	246	327	243	665	591	489	766	596
<b>Revenu Moyen</b>															
Avant Impôts et transferts															
Non Imputés	67 415	63 936	69 365	86 247	111 891	108 585	100 397	130 112	144 700	127 945	182 010	190 325	189 281	201 011	197 073
Imputés	52 917	49 532	51 004	-	-	105 918	85 123	126 562	106 891	-	175 406	179 829	194 478	211 262	-
Échantillon	59 677	56 866	62 294	86 247	111 891	107 564	93 687	128 861	139 608	127 945	181 084	189 186	189 819	201 056	197 073
Après Impôts et transferts															
Non imputés	70 596	67 384	73 113	87 448	108 883	99 964	93 355	121 351	132 131	119 531	167 232	171 186	171 512	181 498	177 763
Imputés	53 102	49 646	51 787	-	-	99 241	78 996	112 693	105 377	-	164 585	162 754	180 670	198 721	-
Échantillon	61 259	58 677	64 900	87 449	108 883	99 687	87 057	118 299	128 527	119 531	166 860	170 272	172 460	181 573	177 763
<b>Professions Intermédiaires</b>															
Employés															
Année	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	852	861	666	915	850	393	475	356	459	376	1067	1353	1120	1366	1158
Obs Imputées	107	61	46	1	0	45	36	25	2	0	126	91	54	14	0
Total	959	922	712	916	850	438	511	381	461	376	1193	1444	1174	1380	1158
<b>Ouvriers</b>															
Employés															
Année	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	117 396	121 267	123 840	122 899	126 925	92 921	99 577	97 496	100 994	101 619	82 336	84 322	85 418	88 937	92 317
Obs Imputées	104 894	129 541	130 473	52 118	-	78 257	102 786	111 391	62 511	-	68 114	79 464	86 386	30 837	-
Échantillon	116 022	121 825	124 291	122 805	126 925	91 615	99 801	98 433	100 791	101 619	80 896	84 020	85 465	88 306	92 317
Après Impôts et transferts															
Non Imputés	117 396	121 267	123 840	122 899	126 925	92 921	99 577	97 496	100 994	101 619	82 336	84 322	85 418	88 937	92 317
Imputés	104 894	129 541	130 473	52 118	-	78 257	102 786	111 391	62 511	-	68 114	79 464	86 386	30 837	-
Échantillon	116 022	121 825	124 291	122 805	126 925	91 615	99 801	98 433	100 791	101 619	80 896	84 020	85 465	88 306	92 317

Francs constants 2000. Échantillon d'hommes actifs de 30 à 50 ans, personnes de référence au sein de leur ménage, ayant déclaré leur catégorie sociale, ainsi que celle de leur père.

TAB. 14: Revenu Moyen en fonction de la CSP du chef de ménage

Année	Agriculteurs				Artisans				Cadres						
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	657	625	453	548	390	460	321	314	526	404	263	216	190	555	490
Obs Imputées	185	143	87	9	0	121	85	48	10	0	41	36	28	13	0
Total	842	768	540	557	390	581	406	362	536	404	304	252	218	568	490
<b>Revenu Moyen</b>															
Avant Impôts et transferts															
Non Imputés	86 525	94 081	93 201	105 959	115 034	125 871	116 556	123 367	140 289	133 470	168 990	176 441	173 881	167 651	161 668
Imputés	61 149	67 644	69 838	130 440	—	107 546	103 631	140 109	128 205	—	182 825	155 992	191 440	104 094	—
Échantillon	81 096	89 282	89 403	106 392	115 034	122 113	113 825	125 642	140 056	133 470	170 976	173 467	176 267	165 874	161 668
Après Impôts et transferts															
Non imputés	87 161	94 268	94 893	105 972	113 361	119 184	110 782	119 689	133 001	128 977	156 727	160 957	158 147	153 417	151 255
Imputés	60 640	66 431	69 715	126 155	—	102 042	97 931	128 695	124 213	—	171 246	138 180	176 109	102 877	—
Échantillon	81 487	89 214	90 800	106 329	113 361	115 668	108 067	120 912	132 832	128 977	158 811	157 644	160 588	152 044	151 255
<b>Professions Intermédiaires</b>															
Employés															
Année	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	252	332	323	307	270	382	510	352	625	603	1246	1474	1200	1370	1190
Obs Imputées	42	40	39	3	0	64	39	29	4	0	201	137	89	23	0
Total	294	372	362	310	270	446	549	381	629	603	1447	1611	1289	1393	1190
<b>Ouvriers</b>															
Employés															
Année	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Obs Non Imputées	126 824	134 111	132 801	134 285	136 938	113 152	119 961	124 456	124 667	117 834	94 204	99 617	100 049	102 654	98 877
Obs Imputées	134 155	140 524	130 452	88 544	—	103 201	107 397	177 194	115 587	—	85 430	94 322	99 240	48 593	—
Échantillon	135 590	141 940	132 542	133 825	136 938	111 921	119 064	128 601	124 605	117 834	93 070	99 159	99 987	101 715	98 877
Après Impôts et transferts															
Non Imputés	126 824	132 451	126 639	127 303	129 708	109 896	114 812	120 981	120 277	117 491	94 427	99 573	99 230	103 182	102 580
Imputés	139 043	130 610	128 287	85 802	—	101 528	104 728	157 134	115 054	—	84 480	89 676	95 608	51 266	—
Échantillon	128 440	132 253	126 821	126 885	129 708	108 861	114 092	123 823	120 241	117 491	93 142	98 717	98 956	102 279	102 580

Francs constants 2000. Échantillon d'hommes actifs de 30 à 50 ans, personnes de référence au sein de leur ménage, ayant déclaré leur catégorie sociale, ainsi que celle de leur père.

TAB. 15: Revenu Moyen en fonction de la CSP du père du chef de ménage

## 5.2 Détail des variables de revenu

### Revenus d'activité :

- revenus d'activité indépendante (Bénéfices agricoles, Bénéfices industriels et commerciaux, bénéfices non commerciaux, rémunération des gérants et associés, dividendes perçus par les dirigeants salariés de leur propre entreprise)
- salaires auto-versés (pour les indépendants)
- salaires et autres rémunérations (y compris primes, indemnités et compléments divers, piges, commissions et pourboires, jetons de présence)
- revenus d'activités secondaires ou occasionnelles

### Revenus du patrimoine :

- revenus des valeurs mobilières de placement
- loyers perçus
- dividendes sur action

### Revenu de transferts assurantiels :

- allocations chômage
- retraites
- pré-retraites.

### Revenu de transferts redistributifs :

- pensions de reversions
- minimum vieillesse
- allocations ou majorations pour tierce personne
- allocation aux adultes handicapés
- pensions d'invalidité
- pensions d'anciens combattants
- aide sociale
- allocations familiales (allocation de base, compléments familial, allocation de rentrée scolaire, allocation pour jeune enfant, aide à la garde d'enfant, allocation parentale d'éducation, allocation de parents isolé, allocation de soutien familial, allocation d'éducation spéciale)
- RMI
- allocations logement et APL.
- pensions alimentaires
- pensions d'anciens combattants.

### Impôts et prélèvements :

- impôts sur le revenu des personnes physiques
- taxe d'habitation
- taxe foncière.

### Revenus non pris en compte :

- transferts intra-familiaux

	Première décomposition						Seconde décomposition					
	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers	Agriculteurs	Artisans	Cadres	Prof Interm	Employés	Ouvriers
Écart 1979-2000	31 874	13 309	-7 556	1 268	8 630	9 439	31 873	13 308	-7 556	1 267	629	9 438
Effet structure	28 147	10 573	-1 016	7 263	8 195	10 555	26 696	13 334	-1 239	5 776	8 706	10 444
	88.31%	79.44%	13.45%	573.03%	94.96%	111.82%	83.76%	100.19%	16.39%	455.68%	100.88%	110.65%
Effet mobilité	3 727	2 736	-6 540	-5 996	435	-1 116	5 177	-25	-6 317	-4 508	-76	-1 005
	11.69%	20.56%	86.55%	-473.03%	5.04%	-11.82%	16.24%	-0.19%	83.61%	-355.68%	-0.88%	-10.65%
	<hr/>											
Écart 1979-1989	9 313	5 244	1 777	-1 620	14 962	5 814	9 313	5 244	1 777	-1 620	14 962	5 814
Effet structure	8 054	6 834	714	2 158	12 780	5 966	7 704	9 964	-1 044	2 821	11 732	5 807
	86.48%	130.31%	40.17%	-133.26%	85.42%	102.60%	82.72%	190.00%	-58.74%	-174.17%	78.41%	99.87%
Effet mobilité	1 259	-1 589	1 063	-3 778	2 182	-151	1 609	-4 720	2 821	-4 440	3 230	8
	13.52%	-30.31%	59.83%	233.26%	14.58%	-2.60%	17.28%	-90.00%	158.74%	274.17%	21.59%	0.13%
	<hr/>											
Écart 1989-2000	22 561	8 064	-9 333	2 887	-6 332	3 625	22 561	8 064	-9 333	2 887	-6 332	3 625
Effet structure	19 450	5 147	346	4 694	-2 960	4 517	18 932	4 181	1 432	4 531	-4 776	4 466
	86.21%	63.82%	-3.71%	162.58%	46.74%	124.63%	83.91%	51.84%	-15.35%	156.96%	75.43%	123.22%
Effet mobilité	3 111	2 917	-9 679	-1 807	-3 373	-893	3 629	3 884	-10 766	-1 644	-1 556	-842
	13.79%	36.18%	103.71%	-62.58%	53.26%	-24.63%	16.09%	48.16%	115.35%	-56.96%	24.57%	-23.22%
L'écart est un écart de revenu moyen, en francs constants 2000. La catégorie sociale est celle du père. La première décomposition correspond à l'équation (1). La seconde à l'équation (2).												

TAB. 16: Décomposition de l'évolution du revenu moyen- Après Impôts et transferts

<b>BdF 1984</b>	Agriculteurs	Artisans	Cadres	Prof Inter.	Employés	Ouvriers	Total
Agriculteurs	22.10	06.27	07.81	14.37	09.91	39.54	18.95
Artisans	01.75	23.79	17.29	24.27	09.38	23.52	10.50
Cadres	0.072	03.66	62.02	23.97	06.03	03.60	06.66
Prof Inter.	0.077	06.80	26.00	41.26	09.82	15.35	09.51
Employés	00.54	04.51	18.83	27.46	21.41	27.26	13.85
Ouvriers	00.70	06.25	07.34	21.01	13.97	50.71	40.53
Total	04.85	07.74	15.48	23.11	12.83	35.99	100

<b>BdF 1989</b>	Agriculteurs	Artisans	Cadres	Prof Inter.	Employés	Ouvriers	Total
Agriculteurs	23.76	06.80	06.03	14.47	09.58	39.35	16.59
Artisans	00.85	17.15	16.49	25.52	10.42	29.56	11.71
Cadres	00.40	11.03	54.73	24.35	05.13	04.35	07.15
Prof Inter.	00.49	07.82	25.44	34.68	10.48	21.08	11.50
Employés	00.74	07.82	19.02	23.94	19.49	28.99	12.12
Ouvriers	00.51	07.88	07.69	19.21	12.71	52.00	40.92
Total	04.42	09.00	15.23	21.88	11.95	37.52	100

<b>BdF 1994</b>	Agriculteurs	Artisans	Cadres	Prof Inter.	Employés	Ouvriers	Total
Agriculteurs	20.99	09.52	09.12	16.54	08.96	34.86	13.54
Artisans	01.36	16.77	24.77	23.03	10.87	23.20	13.43
Cadres	00.45	09.07	48.51	23.12	10.02	08.83	14.55
Prof Inter.	01.03	06.35	28.72	30.02	13.77	20.12	07.93
Employés	00.13	07.83	14.80	27.36	15.85	34.03	15.83
Ouvriers	00.50	05.49	08.78	20.93	11.11	53.18	34.72
Total	03.37	08.51	19.29	22.67	11.59	34.57	100

En pourcentage, les pourcentages somment en ligne. La catégorie socioprofessionnelle du père est en ligne celle du fils en colonne. Lecture du tableau : Dans l'échantillon en 1994, 13.54% sont fils d'agriculteurs, 03.37% sont agriculteurs, et 09.52% des fils d'agriculteurs deviennent artisans. Les données sont pondérées.

TAB. 17: Matrices de transition socioprofessionnelle de 1984 à 1994

- ressources exceptionnelles<sup>32</sup>
- ventes de biens durables sur le marché de l'occasion
- revenus du capital ayant un caractère transitoire : les profits de construction, les ventes de biens immobiliers, les ventes de terrain, les rentes, les avantages en nature évalués en francs. Ainsi, les revenus du capital qui sont retenus correspondent davantage à la définition économique du revenu.

NB : Tous les revenus déclarés sont nets de cotisations sociales. L'enquête 1979 a la particularité de ne pas tenir compte des ressources diverses, les prestations sociales sont restreintes au minimum : les allocations familiales, les allocations logements et le minimum vieillesse. Enfin, les revenus du patrimoine ne comprennent que les revenus des valeurs mobilières de placement.

<sup>32</sup>Héritages, gains aux jeux de hasard, dommages et intérêts, indemnités de licenciement, prime à l'amélioration de l'habitat.

### 5.3 Procédure d'imputation du revenu

Sur les différentes vagues des enquêtes Budget de Famille un nombre important de ménages ne déclarent pas leurs revenus. Afin de conserver un échantillon représentatif de la population française, nous n'avons pas éliminé ces individus et nous avons imputé une valeur de revenu dans les cas suivants :

- Lorsqu'il existe au moins un indépendant actif au sein du ménage, alors qu'aucun revenu d'activité indépendante n'est déclaré.
- Lorsque l'INSEE demande au ménage de situer son revenu moyen global mensuel dans une tranche : cette question n'est posée que si le ménage a refusé de donner les montants en clairs des différents types de revenus en 1979, 1984, et 1989. Pour 1994 et 2000, les montants en tranches sont connus pour tous les ménages : cette situation n'a donc pas été utilisée pour les deux dernières vagues de l'enquête.
- Lorsque le ménage ne figure pas dans l'ensemble des ménages déclarant un impôt sur le revenu positif ou nul.

Deux valeurs de revenu sont imputées. La première correspond au revenu avant impôts et transferts. Il s'agit des ressources globales d'activité, d'allocation chômage, ou de retraite du ménage (salaire et primes, revenu d'indépendant et salaires auto-versés, revenus pour activités secondaires, allocation chômage, retraite et préretraite). La seconde correspond au revenu disponible du ménage après impôts et transferts (les ressources globales d'activité, d'allocation chômage, ou de retraite du ménage, ainsi que les transferts sociaux moins les impôts sur le revenu). Le revenu est estimé en logarithme directement sous la forme d'un revenu par unité de consommation.

Nous avons utilisé la méthode des résidus simulés telle qu'elle est décrite et utilisée par l'INSEE (Chopin et Masse, 2002; Lollivier, 1997)). Une équation de revenu a été estimée sur l'ensemble de la population qui a déclaré son revenu en clair (sous population E) mais également sur la population dont on ne possède qu'une information en tranche (sous population T). En notant  $[t_{n_i-1}, t_{n_i}]$  la tranche de revenu déclarée par le ménage  $i$ ,  $x_i$  les caractéristiques observables de la personne de référence, et  $y_i$  la valeur de revenu déclaré, on a utilisé la vraisemblance suivante pour estimer le revenu du ménage :

$$L(\beta, \sigma) = \prod_{i \in E} \left\{ \frac{\exp[-(y_i - x_i\beta)^2 / 2\sigma^2]}{\sqrt{2\pi}\sigma} \right\} \prod_{i \in T} \left\{ \Phi\left[\frac{t_{n_i} - x_i\beta}{\sigma}\right] - \Phi\left[\frac{t_{n_i-1} - x_i\beta}{\sigma}\right] \right\}$$

On a utilisé pour chaque vague :

1. une équation avec les variables d'impôts (termes linéaire et quadratique) pour les ménages qui déclarent des impôts strictement positifs.
2. une équation sans ces termes si l'impôt est nul ou non déclaré.

Après avoir estimé l'équation de revenu, un résidu simulé est ajouté à la valeur prédite de revenu pour les ménages concernés. Pour les ménages qui ont déclarés un revenu en tranche, on impute un résidu tiré d'une loi normale, tronqué à droite et à gauche, de telle manière, que le revenu imputé appartienne à la tranche déclarée. Pour ceux qui n'ont pas déclarés de tranche, le résidu est tiré d'une loi non tronquée. A titre d'exemple nous donnons les équations d'imputation de l'année 1979.



## Equations d'imputation du revenu 1979

	Avant Impôts		Après Impôts	
	Coef.	Ecart-type	Coef.	Ecart-type
Intercept	9.1312	(0.2517)	8.2278	(0.2508)
age	-0.1079	(0.0170)	-0.0639	(0.0170)
age2	0.0024	(0.0004)	0.0014	(0.0004)
age3	-0.0000	(0.0000)	-0.0000	(0.0000)
ldep	0.1827	(0.0122)	0.2209	(0.0121)
mimpotot	0.0000	(0.0000)	0.0000	(0.0000)
mimpotot2	-0.0000	(0.0000)	-0.0000	(0.0000)
impo	0.3158	(0.0147)	0.1871	(0.0146)
diplo1	0.0344	(0.0170)	0.0187	(0.0170)
diplo2	0.0682	(0.0188)	0.0514	(0.0187)
diplo3	0.0918	(0.0217)	0.0688	(0.0216)
diplo4	0.0790	(0.0244)	0.0593	(0.0243)
diptech1	0.0369	(0.0124)	0.0337	(0.0124)
diptech2	0.0639	(0.0266)	0.0590	(0.0265)
diptech3	0.0772	(0.0347)	0.0595	(0.0346)
diptech4	0.0659	(0.0330)	0.0415	(0.0329)
diptech5	0.0507	(0.0286)	0.0246	(0.0285)
diptech6	0.0347	(0.0199)	0.0264	(0.0198)
homme	0.0629	(0.0247)	0.0465	(0.0246)
nadulte	-0.0464	(0.0070)	-0.0282	(0.0070)
indep234	0.1128	(0.0237)	0.1317	(0.0236)
aise1	0.3146	(0.0248)	0.2869	(0.0247)
aise2	0.2264	(0.0193)	0.1911	(0.0193)
aise3	0.1333	(0.0186)	0.1028	(0.0185)
zeat1	0.1306	(0.0137)	0.1204	(0.0137)
zeat2	0.0168	(0.0136)	0.0181	(0.0135)
csp2	0.0003	(0.0299)	0.0362	(0.0298)
csp3	0.1979	(0.0355)	0.2516	(0.0354)
csp4	0.1905	(0.0330)	0.2297	(0.0329)
csp5	0.1049	(0.0337)	0.1592	(0.0336)
csp6	0.0917	(0.0320)	0.1499	(0.0319)
celib	0.1063	(0.0226)	0.0699	(0.0225)
nenf12	-0.0951	(0.0115)	-0.1059	(0.0115)
Scale	0.4348	(0.0036)	0.4335	(0.0036)
Obs. en clair	6739		6739	
Obs. en tranche	742		742	

**Notes :** Age : âge de la personne de référence du méréférence. ldep : logarithme de la dépense globale du ménage. mimpotot : Montant de l'impôt sur le revenu. Mimpotot2 : Montant de l'impôt sur le revenu au carré. Diplo : Diplôme d'enseignement général le plus élevé. Diplo1 : CEP. Diplo2 : BEPC, BE, BEPS. Diplo3 : Baccalauréat. Diplo4 : Diplôme supérieur au baccalauréat. Diptech : Diplôme technique le plus élevé. Diptech1 : CAP, BEP. Diptech2 : Brevets professionnels. Diptech3 : BES, BEH, BEI. Diptech3 : Baccalauréat de technicien. Diptech4 : BTS, DUT. Diptech5 : Autres diplômes professionnels. Homme : La personne de référence est un homme. Nadulte : Nombre d'adultes dans le ménage. Indep : Présence d'un actif indépendant. Aise : Aisance financière du ménage. Zeat : Variable de zone géographique. Zeat1 : Île de France. Zeat2 : Bassin Parisien. Csp : Catégorie socio-professionnelle de la personne de référence. Csp2 : Artisans. Csp3 : Cadres. Csp4 : Profession Intermédiaires. Csp5 : Employés. Csp6 : Ouvriers. Celib : Célibataire. Nenf : Présence d'enfants dans le ménage. Scale correspond à l'écart type estimé du résidu.

### 5.4 Les revenus de l'exploitation agricole dans les enquêtes BDF sont-ils fiables ?

La forte croissance du revenu moyen des agriculteurs sur la période nous a incité à effectuer une opération de vérification, pour savoir s'il ne s'agissait pas d'une illusion

Année	1979	1984	1989	1994	Indice
1979	-	-	-	-	100
1984	-0.1074	-	-	-	89.3
1989	0.0834	0.2138	-	-	108.3
1994	0.3195	0.4784	0.2180	-	132
2000	0.5305	0.7147	0.4270	0.1580	153

Indice du revenu net d'entreprise agricole par actif non salarié.

Evolution en prix constants en pourcentage.

TAB. 18: Evolution du revenu Agricole - indice INSEE

Année	(1) Ménages d'agriculteurs	(2) Actifs agricoles non salariés	(3) Indi ne déclarant pas de ressources	(4) Individus à imputer	(5) Ménages à imputer
1979	587	1027	696	476	264
1984	623	1000	671	472	287
1989	397	597	354	219	146
1994	298	415	57	3	1
2000	254	331	47	0	0

TAB. 19: Ménages d'agriculteurs

statistique.

La croissance du revenu net d'exploitation agricole par actif non-salarié qui ressort des comptes de l'agriculture sur les sous-périodes qui nous intéressent est donnée dans le tableau 18. On assiste à une croissance très forte de ce revenu après 1989. Nous nous sommes demandé si nous obtenions, compte tenu des imputations effectuées, des résultats du même ordre de grandeur sur l'ensemble des ménages dans lesquels il existe au moins un agriculteur exploitant.

Les opérations d'imputation décrites dans l'annexe 3 portent sur le revenu primaire ou le revenu disponible et non sur le revenu net agricole. Afin de permettre une comparaison avec les données de la table ci-dessus, nous avons réalisé une opération d'imputation spécifique. La variable imputée est le revenu d'activité agricole indépendante du ménage. La procédure d'imputation utilise une estimation par moindres carrés ordinaires du revenu d'activité agricole régressé sur l'âge du chef d'exploitation, le diplôme du chef d'exploitation, le nombre de salariés et le nombre d'indépendants de l'exploitation, la ZEAT, la taille de l'exploitation, la spécialisation agricole.

Les ménages réimputés sont les ménages dans lesquels il existe au moins un agriculteur exploitant et où aucun revenu d'indépendant n'est déclaré. L'imputation ne se fait

Année	1979	1984	1989	1994	2000
imputés	41 719	44 475	44 681	27 730	-
non imputés	62 358	65 460	57 760	62 924	77 464
ensemble	52 574	55 589	53 023	62 477	77 464

En francs constants 2000. Données Pondérées.

TAB. 20: Revenus Moyens

	1979	1984	1989	1994	Revenu Moyen
1979	-	-	-	-	50 196
1984	0.1135	-	-	-	55 894
1989	0.0382	-0.0676	-	-	52 114
1994	0.2314	0.1059	0.1861	-	61 812
2000	0.5432	0.3859	0.4864	0.2532	77 464

Revenu agricole moyen par actif non salarié, après imputation au niveau du ménage. Prix constants 2000. Evolution en pourcentages.

TAB. 21: Evolution du revenu Agricole - Budget de famille

	1979	1984	1989	1994	2000
Coef.	0.9714	0.7991	0.9348	1.1167	1.0000

TAB. 22: Coefficients de calage sur les comptes de l'agriculture

qu'au niveau du ménage et non pas de l'individu car dans beaucoup de cas, les revenus ne sont ni individualisables, ni individualisés dans les déclarations : par exemple, en 79, le revenu de l'exploitation agricole n'est déclaré que par une seule personne, en général, la personne de référence du ménage, même si plusieurs membres du ménages travaillent sur l'exploitation. Par ailleurs, tous les membres du ménages de la catégorie aide familiale ont dans les données une CS égale à la CS de la personne qu'ils aident.

Il semble donc correct de ne pas réimputer les individus qui ne déclarent pas de ressources agricoles, pourvu que des ressources soient déclarées au niveau du ménage. Pour certains ménages de notre base, on trouve une déclaration de revenu nul. Il existe dans les données une case de contrôle permettant de savoir si c'est une mauvaise déclaration refus de répondre ou si la valeur traduit des pertes de l'exploitation. Dans le cas des pertes, on n'impute pas. Dans celui des mauvaises déclarations refus de répondre, on réimpute. Au total, l'é tendue des imputations est la suivante :

Les revenus moyens imputés sont en moyenne plus faible que ceux des non-imputés. Nous sommes alors en mesure de calculer une évolution du revenu agricole net sur notre échantillon et de la comparer avec celle obtenue dans les comptes nationaux.

La comparaison se révèle globalement satisfaisante. Sur l'ensemble de la période, nous ne surestimons pas la croissance du revenu agricole. Il faut par contre être beaucoup plus circonspects sur l'évolution par sous-périodes. Nous nous sommes interrogés pour savoir, s'il convenait de caler nos données microéconomiques. Sur l'ensemble re-imputés / non-reimputés, les coefficients de correction auraient été les suivants :

Hormis l'année 1984 pour laquelle les données semblent de mauvaise qualité, les coefficients de calage sont faibles et nous avons décidé de ne pas procéder à cette correction. Cette opération de vérification nous a convaincu qu'il fallait bien procéder à une réimputation, du fait que le revenu moyen des réimputés est différent de celui qui déclarent, mais qu'il fallait rester prudent à propos de tout résultat spécifique à l'année 1984. La réimputation pour le revenu des agriculteurs a donc été réalisée selon les modalités présentées dans l'annexe 3, c'est à dire en utilisant toute l'information disponible.