

Le revenu selon l'origine sociale

Arnaud Lefranc, Nicolas Pistolesi et Alain Trannoy*

Les descendants de cadres sont-ils avantagés par rapport aux descendants d'ouvriers en termes de niveau de vie ? Cette question amène à analyser les distributions de revenus offertes aux ménages selon la catégorie socioprofessionnelle du père du chef de famille. Ces distributions de revenus sont assimilées à des loteries dont le rendement et le risque font l'objet d'estimations à partir des cinq vagues des enquêtes *Budget de Famille* réalisées entre 1979 et 2000. La comparaison de ces loteries permet d'évaluer le degré d'inégalité des chances.

L'inégalité des chances de revenu n'a pas disparu au cours des deux dernières décennies et cette persistance provient en grande partie des écarts de revenus espérés. Les écarts de risque inhérent à chaque loterie sont en effet de faible ampleur. En 2000, un descendant de cadre peut espérer bénéficier d'un niveau de vie de 50 % supérieur à celui d'un descendant d'ouvrier. L'écart a diminué de 20 points en vingt ans. En revanche, les descendants de non-salariés, et en particulier ceux des agriculteurs, ont de meilleures perspectives de revenu qu'auparavant. Au total, la hiérarchie des revenus selon l'origine sociale a peu changé mais s'est resserrée. Les résultats plaident donc pour une réduction du degré de l'inégalité des chances. Cette évolution du revenu espéré est décomposée entre mobilité sociale et évolution du revenu par catégorie socioprofessionnelle : l'amélioration des perspectives des descendants d'indépendants et d'agriculteurs provient surtout d'une rémunération accrue des métiers auxquels ils se destinent. À l'opposé, l'érosion de l'avantage des descendants de cadres traduit leur difficulté croissante à rester dans leur groupe social d'origine, en dépit de l'augmentation de la proportion de cadres dans la population. L'inégalité des chances provenant de la CSP du père ne contribue que pour une part assez faible à l'inégalité globale. Cette part a tendance à diminuer au cours du temps.

* Arnaud Lefranc appartient au Thema, Université de Cergy-Pontoise et Idep, Nicolas Pistolesi au Thema, Université de Cergy-Pontoise, et Alain Trannoy à l'EHESS, Greqam-Idep.
Les noms et dates entre parenthèses renvoient à la bibliographie en fin d'article.

L'adage « Tel père, tel fils ? », repris dans le titre d'un célèbre ouvrage (Thélot, 1982), a été passé au crible de très nombreuses études empiriques. Ainsi, les travaux de Boudon (1974), Erikson et Goldthorpe (1992), Goux et Maurin (1997) ou encore Vallet (2001) ont permis d'éclairer la dynamique récente des inégalités sociales à l'œuvre dans la société française. Toutefois, en se focalisant sur la transmission du statut social, mesuré par la catégorie socio-professionnelle (CSP) individuelle, ces travaux ne permettent pas de comprendre l'influence exercée par l'origine sociale sur une variable clé pour l'analyse économique : le revenu individuel. Cet article s'efforce de combler cette lacune, en analysant le rôle du milieu d'origine dans la formation des inégalités de revenu, en France, au cours des 25 dernières années. Pour cela on procède à la comparaison des distributions de revenus en fonction de l'origine sociale. Ainsi peut-on savoir si la répartition du revenu satisfait ou non à un principe d'égalité des chances.

Isoler la part du milieu d'origine dans la formation du revenu

S'il s'inscrit dans le prolongement des travaux sociologiques sur la mobilité sociale, cet article rejoint aussi les analyses économiques actuelles des inégalités de revenu. De nombreux auteurs ont cherché à mesurer l'étendue des inégalités de revenu ou de conditions de vie et à cerner le rôle joué par le fonctionnement du marché du travail et par la politique fiscale dans leurs évolutions récentes (Atkinson *et al.*, 2001). En examinant l'influence de l'origine sociale dans la formation du revenu, cet article propose une approche complémentaire de la formation des inégalités. Cette approche répond, par ailleurs, aux recommandations de certains philosophes de la responsabilité, tels que Dworkin (1981), Arneson (1989), Cohen (1989) ou encore Roemer (1998). Ces auteurs ont en effet affirmé le manque de fondement éthique d'une analyse de l'inégalité des revenus qui ignorerait le rôle de la responsabilité individuelle. Or, de ce point de vue, la formation du revenu met en jeu des déterminants de nature diverse. Certains facteurs tels que les préférences et l'effort individuels relèvent plutôt de la responsabilité individuelle. D'autres facteurs, regroupés sous le terme de circonstances, n'en relèvent pas. C'est le cas par exemple de l'origine familiale, du talent intrinsèque des individus, ou encore du facteur chance. La philosophie de la responsabilité distingue ces différents facteurs. Elle consi-

dère ainsi que les inégalités de revenu résultant des premiers facteurs sont équitables, car elles sont la conséquence de l'exercice par les individus de leur liberté de choix (au sens large). Il n'en est pas de même des inégalités résultant des différences de circonstances. La philosophie de la responsabilité préconise l'égalité des chances : les circonstances dont les individus ne sont pas responsables ne doivent pas leur conférer d'avantage ou de désavantage systématique dans l'obtention du revenu.

Vers une définition de l'égalité des chances adaptée à l'approche empirique

Une fois posé le principe d'égalité des chances, plusieurs difficultés apparaissent, dans le cadre d'une étude empirique. Tout d'abord, il convient de définir l'ensemble des circonstances individuelles pertinentes. Parmi celles-ci, les circonstances tenant au milieu familial d'origine peuvent, sans aucun doute, être qualifiées d'exogènes au regard de la formation du revenu. En pratique les enquêtes disponibles n'offrent qu'une description très limitée de ce milieu. Pour cette raison, cet article se limite à l'étude de l'effet de la catégorie socio-professionnelle du père (1) sur les niveaux de vie obtenus par les enfants.

Il convient par ailleurs de se doter d'une définition de l'égalité des chances qui soit adaptée à une approche empirique. Une première option, proposée par Van de Gaer (1993), consiste à comparer les revenus moyens des individus en fonction de leur milieu d'origine. L'égalité des chances est alors définie par l'égalité de ces revenus moyens conditionnellement à l'origine sociale. Ce critère conduit cependant à ignorer des phénomènes intéressants du point de vue de l'inégalité des chances. Par exemple, il se pourrait que les revenus moyens des fils d'agriculteurs et des fils d'ouvriers soient identiques, alors même que le premier décile de la distribution des fils d'ouvriers est plus élevé que le premier décile de la distribution des fils d'agriculteurs, l'inverse prévalant pour le dixième décile. La traduction économique d'une telle observation serait alors qu'avoir un père ouvrier serait plus favorable qu'avoir un père agriculteur pour les enfants qui échouent relativement (ceux du premier décile), alors que la conclusion inverse

1. Pour des raisons de taille d'échantillon, on ne prend pas en compte, à une exception près, 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).

prévaudrait pour les enfants qui réussissent particulièrement bien (ceux du dernier décile). L'égalité des revenus moyens conditionnellement à l'origine sociale ne suffit donc pas dans ce cas, à garantir l'égalité des chances.

Il convient donc d'élaborer un critère d'égalité des chances qui tienne compte de l'effet qu'exerce le milieu d'origine sur l'ensemble de la distribution de revenu offerte aux individus. Dans cet article, l'égalité des chances est définie par les choix que feraient les individus s'il leur était possible de choisir leur milieu d'origine.

Chacun de ces milieux est assimilé à une loterie offrant aux individus qui en sont issus une distribution de revenu possible : le fait de naître de parents ouvriers apporte certaines perspectives (aléatoires) de revenu ; le fait de naître de parents cadres apporte des perspectives différentes. La destinée, en termes de revenu, est inscrite dans ces distributions. Les fréquences associées à chaque distribution donnent les chances d'atteindre tel ou tel revenu.

Choisir un milieu d'origine revient donc à choisir une loterie particulière. Le choix entre ces loteries – par définition aléatoires – s'apparente à un choix en univers risqué. On dira que l'inégalité des chances prévaut dès lors qu'un individu préférera les loteries offertes par certains milieux d'origine à celles offertes par d'autres milieux. On est alors amené à comparer les distributions de revenus conditionnellement à l'origine sociale, à l'aide des instruments de dominance stochastique issus de la théorie de la mesure du risque. Cette définition présente l'avantage de conduire naturellement aux instruments habituels de la mesure des inégalités que sont la courbe de Lorenz ou la courbe de Lorenz généralisée. Toutefois, contrairement à la définition adoptée par Roemer (1998 et 2003), elle conduit, en général, à un classement incomplet des distributions de revenu. C'est le prix à payer pour ne pas faire dépendre le classement d'axiomes plus spécifiques.

L'enquête *Budget de Famille* : une mesure appropriée du revenu

Dans cet article, l'égalité des chances est évaluée en termes de niveau de vie, mesuré par le revenu par unité de consommation. Deux étapes importantes de la formation du revenu disponible se trouvent analysées : la formation du revenu primaire, et l'effet du système fiscal français (impôt sur le revenu et transferts redistributifs) sur la correction de l'inégalité des

chances dans la formation du revenu primaire. On exploite à cet effet les cinq vagues des enquêtes *Budget de Famille* (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).

L'analyse des revenus des ménages se justifie par l'importance que cette variable revêt dans l'appréciation du bien-être. Toutefois, la formation du revenu représente un processus complexe, qui mêle notamment l'effet de l'éducation, de la qualification, du chômage et de la constitution des couples. Il serait évidemment souhaitable de pouvoir mesurer la contribution propre de chacun de ces facteurs à l'inégalité des chances (2). Ceci constitue cependant une entreprise d'autant plus difficile que cet article ne se restreint pas à l'analyse des inégalités de salaire.

La comparaison de distributions de revenus construites à partir de données d'enquête nécessite de s'interroger sur la significativité statistique des résultats obtenus. On a recours pour cela à des techniques d'inférence statistique adaptées à l'analyse des relations de dominance stochastique. Ces techniques sont récentes et en rapide évolution. Elles s'appuient sur les procédures de tests non paramétriques développées par Davidson et Duclos (2000) qui permettent, par l'absence d'hypothèse paramétrique particulière, d'aboutir à des conclusions statistiques plus robustes (3).

Comparer les distributions de revenu selon l'origine sociale

L'idée essentielle est la suivante : naître dans un milieu social, repéré ici par la CSP du père, revient à tirer un billet de loterie. Les prix attachés à ces billets de loterie sont constitués des distributions de revenu des descendants d'un milieu social. Certains individus ont plus de chance : ils naîtront dans un milieu social plus porteur qui leur donnera accès à un billet de loterie plus favorable.

2. Voir, dans le cas du Brésil, Bourguignon, Ferreira et Menendez (2003).

3. 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, la taille de l'échantillon ne permet pas de retenir une telle méthodologie.

Cette formalisation peut s'expliquer au moyen des notations suivantes. Le milieu d'origine est représenté par une variable aléatoire discrète (4) notée s , $s \in S = \{1, \dots, \bar{s}\}$. Le revenu des descendants est la variable aléatoire X continue sur un support positif, et on note x ses réalisations. La fonction de répartition de X sachant s , $F(x|s)$, donne la distribution de revenu des descendants d'origine sociale s . La génération des descendants fait donc face à une « distribution de chances » repérée par la donnée de s distributions de revenu selon l'origine sociale, soit $\{F(x|s); s = 1, \dots, \bar{s}\}$.

Les fonctions de répartition du revenu des descendants pour trois milieux d'origine sont données, à titre d'exemple, pour deux pays : la Suède et les États-Unis (5) (cf. graphique I). Chaque milieu est

défini à partir du niveau d'éducation du père de la personne de référence (sans éducation, niveau d'éducation secondaire, niveau d'éducation supérieur). Pour chaque pays figurent deux distributions : celle du revenu primaire et celle du revenu disponible (6).

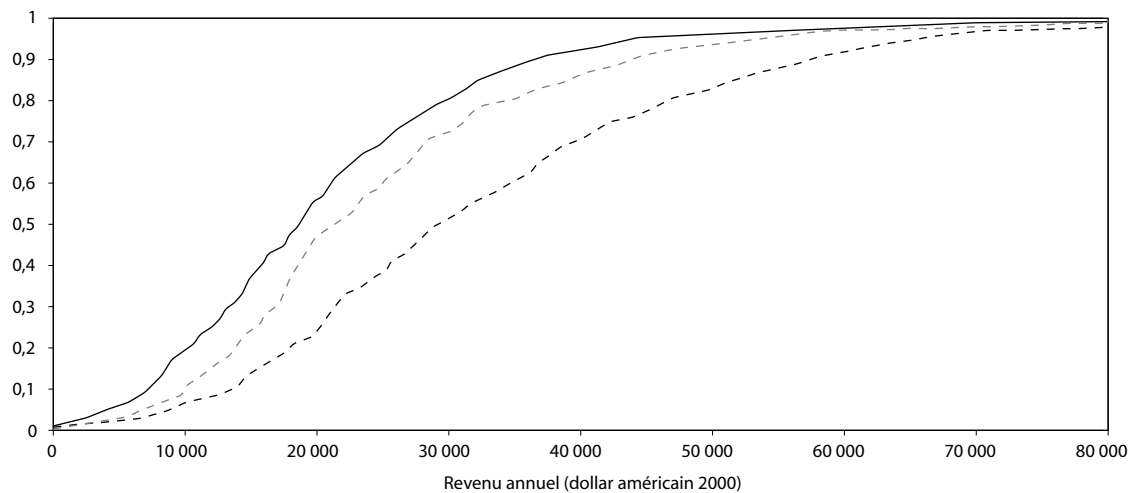
4. Poser s multidimensionnelle n'engendre pas de difficulté supplémentaire.

5. Les données suédoises sont extraites du LNU, l'enquête des conditions de vies des ménages de 1991. L'échantillon est constitué de 825 observations. Les données américaines proviennent du Panel Study of Income Dynamics (PSID) de 1991. Pour les États-Unis, l'échantillon est constitué de 1 140 ménages. Les revenus ont été convertis en dollars de 2000 grâce aux Penn World Tables. Pour tenir compte de la taille du ménage, on utilise l'échelle d'équivalence de l'OCDE.

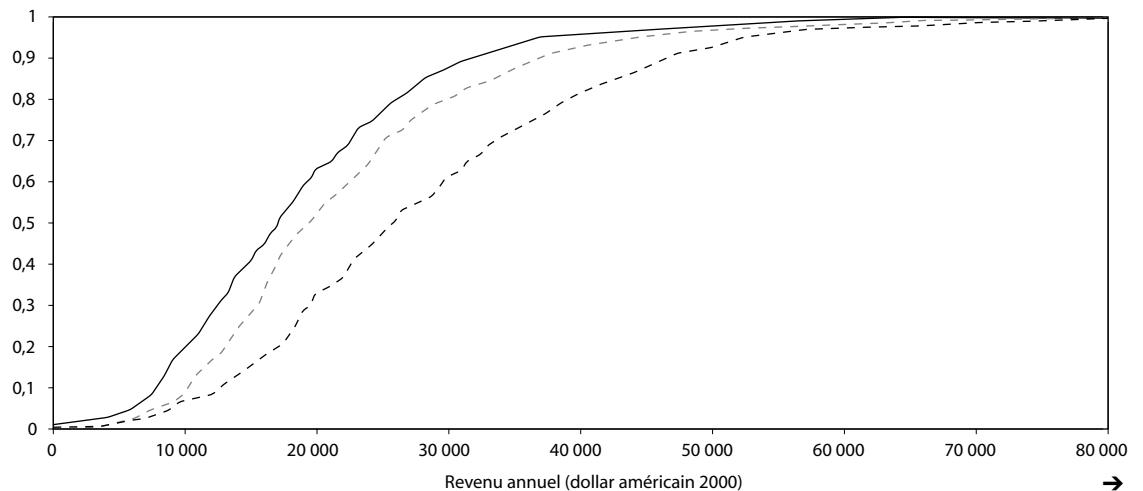
6. Le revenu primaire inclut les revenus du travail et du patrimoine ainsi que les transferts assurantiels. Le revenu disponible est égal au revenu primaire, augmenté des transferts redistributifs et diminué des impôts. Pour plus de précision sur les concepts de revenu utilisés, voir infra.

Graphique I
Distribution du revenu selon le milieu d'origine (États-Unis et Suède)

A - États-Unis : revenu primaire



B - États-Unis : revenu disponible

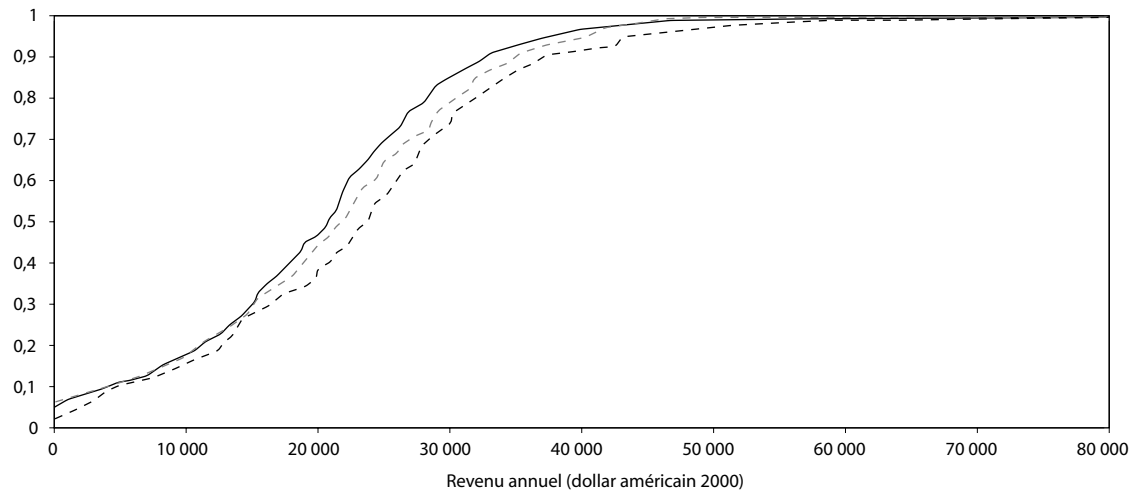


Comme on pouvait s'y attendre, avoir un père plus éduqué est associé à un niveau de revenu plus élevé, puisque la fonction de répartition est décalée sur la droite. Les différences sont beaucoup plus marquées aux États-Unis, où avoir un père qui a fait des études supérieures permet d'avoir un niveau de revenu nettement plus élevé, qu'en Suède, où l'écart des distributions selon l'origine sociale est beaucoup plus ténue. La politique de redistribution n'a pas la même efficacité dans les deux pays. Le resserrement des distributions est relativement modeste aux États-Unis, alors qu'il est très marqué en Suède.

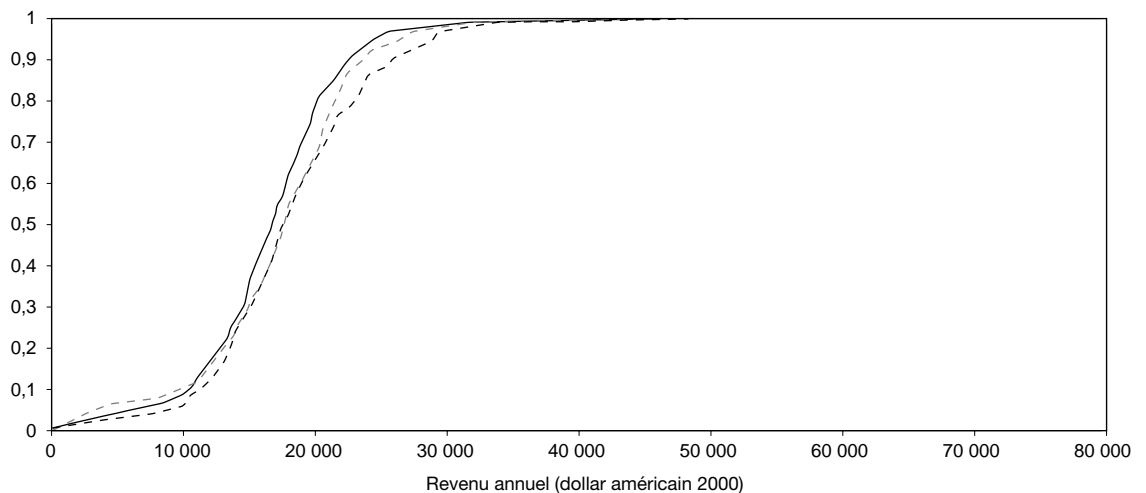
Étudier l'égalité des chances consiste à comparer de manière systématique de telles distributions de revenu conditionnelles au milieu d'origine. On utilise à cet effet les instruments développés pour analyser les choix individuels en univers risqué. On considère le choix opéré par un individu à qui serait offerte la liberté de choisir son milieu d'origine. Si l'on suppose que cet individu adopte un comportement rationnel au sens de la théorie économique, alors son choix reposera sur un critère appelé critère de dominance stochastique. On considère généralement deux critères de

Graphique I (suite)

C - Suède : revenu primaire



D - Suède : revenu disponible



Niveau de formation du père de la personne de référence

— Sans formation	- - - Formation secondaire	. . . Formation supérieure
------------------	----------------------------	----------------------------

Champ : ménages dont la personne de référence est un homme de 25 à 40 ans.
 Source : États-Unis : Panel Study of Income Dynamics, 1991 ; Suède : LNU – enquête suédoise sur les conditions de vie des ménages, 1991.

dominance stochastique : la dominance à l'ordre 1 et la dominance à l'ordre 2.

Pour comprendre le critère de dominance à l'ordre 1, on considère la partie gauche du schéma ci-contre. La courbe représentant la fonction de répartition de la distribution B est toujours au dessus de celle de A. Par conséquent, quel que soit le seuil de revenu considéré, la probabilité d'obtenir un revenu supérieur ou égal à ce seuil est toujours plus faible avec B qu'avec A. Il est donc plus avantageux de choisir la distribution A car elle garantit toujours des gains plus élevés. En termes statistiques, on dira dans ce cas, que la distribution A domine stochastiquement la distribution B à l'ordre 1.

Le critère de dominance à l'ordre 1 ne permet pas de comparer toutes les distributions, par exemple lorsque les deux courbes se croisent comme sur la partie droite du schéma. Dans ce cas, la distribution A est plus avantageuse que la distribution B pour des niveaux de gain faibles. La distribution B est plus avantageuse pour des niveaux de gains élevés. Un individu qui ne souhaite pas prendre trop de risque (on dira qu'il est averse au risque) pourra préférer la distribution A qui lui donne une probabilité inférieure d'avoir un niveau de revenu faible. Le critère de dominance stochastique d'ordre 2, présenté dans l'encadré 1, généralise ce raisonnement.

Encadré 1

CHOIX DU MILIEU D'ORIGINE ET DOMINANCE STOCHASTIQUE

On considère la situation hypothétique d'un individu à qui serait offerte la liberté de choisir son milieu d'origine. S'il adopte un comportement rationnel au sens de la théorie économique standard, il choisira ce milieu d'origine de manière à maximiser l'espérance mathématique d'utilité (1) procurée par les billets de loterie sociale correspondant aux différents milieux d'origine. En désignant par u l'utilité de Von Neuman et Morgenstern associée aux différents niveaux de revenu, l'espérance d'utilité de naître dans le milieu s est donnée par

$$E_u(s) = \int_x u(x) dF(x|s)$$

Sans hypothèse sur la fonction d'utilité, la comparaison des loteries au regard de ce critère d'espérance d'utilité risque de tourner court. Il est raisonnable de supposer que la fonction d'utilité est croissante, c'est-à-dire que toute élévation du revenu entraîne un supplément d'utilité. On peut ensuite supposer que l'individu considéré fait preuve d'aversion au risque. Dans le contexte de l'utilité espérée, cette seconde hypothèse se traduit par la décroissance de l'utilité marginale en fonction du revenu, autrement dit par la concavité de la fonction d'utilité. Les critères de dominance stochastique énoncent des conditions que doivent vérifier deux loteries – ici, les distributions de revenu des descendants selon l'origine sociale – de façon à ce que l'espérance d'utilité d'une des loteries soit plus élevée que celle de l'autre pour toutes les fonctions d'utilité appartenant à une classe suffisamment large (par exemple les fonctions croissantes ou les fonctions croissantes et concaves). On dit alors que l'une des loteries domine stochastiquement l'autre. Deux critères de dominance stochastique sont très usités, celui dit à l'ordre 1 et celui à l'ordre 2 (Gollier, 2003). La définition de ces deux critères est rappelée ci-dessous.

Définition 1. Soient deux distributions A et B définies sur un support positif et de fonction de répartition F_A et F_B . La distribution A domine stochastiquement à l'ordre 1 la distribution B si :

$$A \geq_{DS_1} B \Leftrightarrow \forall x \in \mathbb{R}_+, F_A(x) \leq F_B(x)$$

La dominance stricte requiert, de surcroît, qu'il existe au moins une valeur de x telle que $F_A(x) < F_B(x)$. En situation de choix en univers risqué, pour un décideur adoptant le critère d'espérance d'utilité, 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 2. La distribution A domine stochastiquement à l'ordre 2 la distribution B si :

$$A \geq_{DS_2} B \Leftrightarrow \forall x \in \mathbb{R}_+, \int_0^x F_A(y) dy \leq \int_0^x F_B(y) dy$$

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

$$A \geq_{DS_2} B \Leftrightarrow \forall p \in [0, 1], GL_{F_A}(p) \geq GL_{F_B}(p)$$

en notant $GL_{F_A}(p)$ l'ordonnée de la courbe de Lorenz généralisée en p pour la distribution F_A .

1. Le décideur est censé obéir aux axiomes de choix dans un univers risqué, posés par Von Neuman et Morgenstern. On reproche à ces axiomes de ne pas traduire correctement le comportement en risque mais leur valeur normative est rarement discutée. Il est naturel de les retenir, dans la mesure où l'on se propose d'asseoir la définition de l'égalité des chances sur des fondements normatifs.

En résumé, si la distribution A domine stochastiquement la distribution B à l'ordre 1, alors indépendamment de son attitude vis-à-vis du risque (neutralité, goût du risque ou aversion), l'individu choisira toujours la distribution A. En l'absence de dominance à l'ordre 1 et si l'individu a de l'aversion au risque, alors il choisira la distribution qui domine stochastiquement à l'ordre 2. En cas d'absence de dominance à l'ordre 2, il n'est pas possible de se prononcer sur les choix de l'individu sans formuler d'hypothèses supplémentaires.

De la dominance stochastique à l'inégalité des chances

Quelle que soit son attitude vis-à-vis du risque, la personne confrontée au choix d'un milieu social, choisira un milieu qui n'est dominé stochastiquement à l'ordre 1 par aucun autre. L'ensemble des *milieux sociaux non dominés à l'ordre 1* répond à la définition suivante :

$$P_1 = \{s \in S \mid \nexists s' \in S \text{ avec } F(x|s') >_{DS_1} F(x|s)\}.$$

Dans l'exemple américain, (en revenu primaire ou en revenu disponible) le critère d'ordre 1 démontre un fort pouvoir discriminant. Il permet de classer les trois distributions sans ambiguïté (cf. graphique I). La distribution correspondant au niveau de formation le plus faible du père est dominée par les deux autres, la distribution correspondant au niveau de formation secondaire du père est dominée par la distribution correspondant au niveau de formation supérieur du père. Cette dernière est donc le seul élément non dominé au regard de la dominance stochastique d'ordre 1.

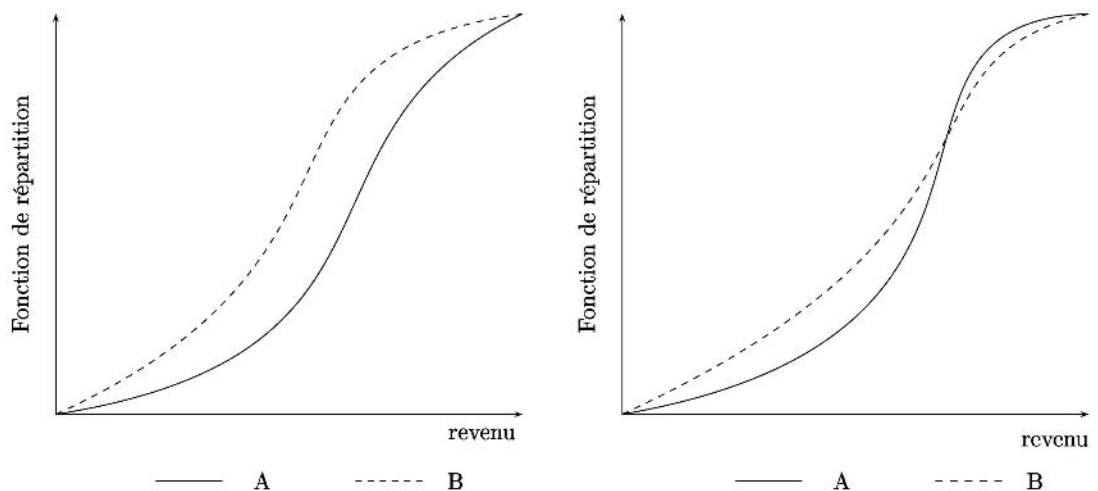
Dans le cas d'aversion au risque, le choix s'effectue en poursuivant la comparaison au moyen du critère de dominance stochastique d'ordre 2. Le milieu d'origine retenu l'est alors dans l'ensemble P_2 des *milieux sociaux d'origine non dominés à l'ordre 2* qui est défini d'une manière analogue à P_1 . Le fait que le critère d'ordre 2 soit plus discriminant que celui d'ordre 1 a pour conséquence que l'ensemble P_2 est inclus dans l'ensemble P_1 : $P_2 \subseteq P_1$.

La définition de ces deux ensembles de milieux sociaux non dominés conduit à une définition de l'égalité des chances. On dira que l'égalité des chances est réalisée si aucun milieu social n'est dominé stochastiquement par un autre à l'ordre 2. En d'autres termes, *l'égalité des chances est atteinte si $P_2 \equiv S$* .

Cette définition s'écarte des propositions existantes (Roemer, 1998 ; Van de Gaer, 1993 ; Goux et Maurin, 2003) (7) car elle ne cherche pas à établir un classement complet des distributions de revenu. Elle recouvre en particulier le cas d'égalité des distributions de revenu conditionnellement au milieu d'origine, que l'on peut qualifier d'*égalité des chances au sens fort*. On se doute toutefois que l'égalité des chances au sens fort risque d'être rarement satisfaite en pratique, même si l'exemple suédois paraît s'en approcher : les trois distributions de revenu disponible correspondant aux trois niveaux de formation du père sont dans ce cas extrêmement proches, sans être toutefois totalement confondues.

7. Voir Ooghe et al. (2003) pour une axiomatisation des solutions proposées par Roemer et Van de Gaer. Dans sa thèse, Van de Gaer (1993, p. 66) propose déjà l'utilisation des outils de dominance stochastique d'ordre 1 et d'ordre 2 pour définir l'égalité des chances.

Comparaison de deux distributions A et B



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 de ceux où un classement peut être opéré, alors même que ces deux situations diffèrent profondément, du point de vue de l'inégalité des chances. En revanche, la définition de l'égalité des chances en tant qu'absence de dominance (8) permet de distinguer de telles situations.

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 ». En effet, l'absence de toute dominance d'ordre 2 implique qu'*a fortiori* il en est de même avec la dominance d'ordre 1, puisque cette dernière est plus exigeante (9).

On peut également souhaiter comparer les loteries, du seul point de vue de leur risque, indépendamment de leurs rendements respectifs. 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. Si la courbe de Lorenz pour une origine sociale donnée est toujours au-dessus de la courbe de Lorenz d'un autre milieu, 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. Par contre, on ne peut conclure si les courbes de Lorenz se coupent. D'une manière équivalente, le résultat de cette comparaison renseigne sur le caractère plus ou moins inégalitaire des distributions de revenu au sein de chaque milieu d'origine.

L'inférence statistique repose sur des tests non paramétriques

Pour utiliser ces critères d'égalité des chances, à partir de données d'enquête, il faut se doter d'instruments d'inférence statistique adaptés. On ne peut en effet conclure à l'égalité des chances au sens fort en Suède, uniquement au vu des graphiques des fonctions de répartition empiriques obtenus à partir d'un échantillon de 825 personnes. Pour que le résultat trouvé à partir d'un échantillon soit jugé statistiquement significatif au niveau de la population toute entière, un test statistique est nécessaire. Les procédures mises en œuvre sont nécessairement

plus complexes que celles utilisées dans les tests d'égalité des moyennes et aucune procédure standard ne s'est encore complètement imposée. En cette matière, il est possible de mettre en œuvre des tests reposant sur l'estimation paramétrique de distributions de revenu (McDonald et Xu, 1995). Toutefois, 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. La méthode ici utilisée repose sur des tests non paramétriques. Bien que la définition de l'égalité des chances soit posée en termes de non-dominance, c'est sur la dominance que porte le test (cf. encadré 2).

Comparer les montants de pauvreté associés à chaque distribution

Soient deux distributions empiriques de revenu : par exemple, les deux distributions américaines de revenu pour le niveau inférieur et supérieur de formation du père. Le test de dominance d'ordre 1 amène à comparer les ordonnées de chacune des fonctions de répartition relatives à chacune de ces deux distributions, cette comparaison étant effectuée pour tous les revenus possibles. Trois possibilités sont envisageables : la fréquence cumulée est toujours plus grande pour l'une que pour l'autre distribution ; la fréquence cumulée est tantôt plus grande, tantôt plus petite ; les fréquences sont rigoureusement égales. Si ces distributions étaient établies à partir de données exhaustives sur l'ensemble de la population, il serait possible de conclure à l'égalité des chances dans les deux derniers cas de figure, et à l'inégalité dans le premier cas. L'utilisation de données d'échantillonnage requiert davantage de prudence et conduit à 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 de fréquence

8. Le critère adopté ici n'implique pas l'égalité des moyennes conditionnelles (critère de Van de Gaer). Roemer postule que les disparités de réussite pour une CSP d'origine donnée sont uniquement dues aux différences d'effort. L'analyse développée ici ne repose pas sur cette hypothèse. Le critère de Roemer exige alors que les déciles des distributions conditionnelles soient égaux. Le critère d'égalité des chances développé ici ne respecte pas cette condition. Le critère d'égalité des chances au sens fort respecte, quant à lui, les conditions de Van de Gaer et de Roemer.

9. Il est possible 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 particuliers.

cumulée calculées pour différents niveaux de revenu) dont on teste l'appartenance à une zone particulière (par exemple : toutes les différences sont positives).

Les relations de dominance peuvent s'exprimer au moyen de la notion de *montant de pauvreté*. On entend par indicateur de montant de pauvreté à l'ordre g , pour un seuil de revenu z , l'expression suivante :

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

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 atteignent le seuil de pauvreté z .

On démontre que si la distribution A domine à l'ordre 2 la distribution B, le montant de pauvreté moyen pour la distribution B est toujours supérieur au montant de pauvreté moyen pour la distribution A (Foster et Shorrocks, 1988). Plus généralement, il existe une relation d'équivalence entre dominance stochastique pour tout ordre de dominance et inégalité des montants de pauvreté. On peut en effet établir que pour tout $g \in \mathbb{N}_+$:

$$A \geq_{DS_g} B \Leftrightarrow \forall z \in \mathbb{R}_+, D_A^g(z) \leq D_B^g(z)$$

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

Encadré 2

CHOISIR LA DOMINANCE COMME HYPOTHÈSE NULLE

On examine dans ce qui suit les conséquences du choix de l'hypothèse nulle (à savoir la dominance ou la non-dominance) sur le résultat du test. Le choix de l'hypothèse nulle de dominance (resp. de non-dominance) maximise la probabilité de ne pas rejeter la dominance (resp. la non-dominance) et donc minimise la probabilité de ne pas rejeter la non-dominance (resp. la dominance).

Le schéma ci-dessous illustre cette différence dans le cas où la comparaison porte sur un vecteur à deux dimensions : l'écart de fréquence cumulée a été calculé en deux points du support de la distribution. Sans l'incertitude due à l'échantillonnage, on conclut à la dominance si le vecteur appartient à l'orthant positif. Dans les autres cas de figure, on conclut à la dominance.

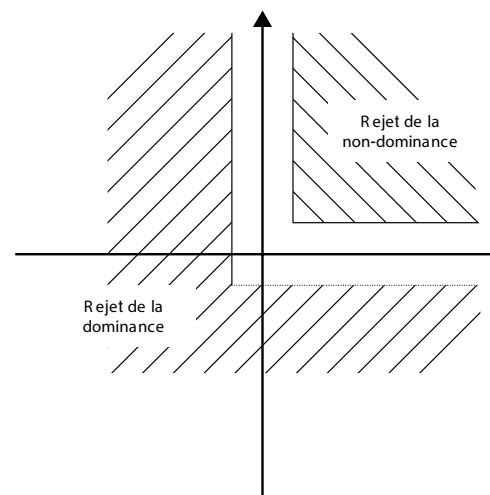
Si l'on retient la dominance comme hypothèse nulle, la région d'acceptation est plus grande que l'orthant positif : on accepte (1) de conclure à la dominance en considérant que des valeurs extérieures à l'orthant positif mais cependant proches de celui-ci peuvent résulter des particularités de l'échantillon tiré et ne pas refléter les caractéristiques de la population tout entière. La région permettant de conclure à l'absence de dominance s'en trouve réduite.

Si la non-dominance est choisie comme hypothèse nulle, on accepte de prendre le risque de ne pas la rejeter, même si les coordonnées du vecteur des 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 l'on accepte 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.

En conséquence, le fait de retenir la non-dominance comme hypothèse nulle (2) 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. On a choisi de retenir la dominance comme hypothèse nulle. Lorsque ce test conclut à l'égalité des chances entre deux distributions, renverser le test ne modifierait pas cette conclusion.

Zones de rejet des hypothèses nulles de dominance et de non-dominance



1. 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.
2. Howes (1993) considère ce cas de figure.

Le principe du test statistique (pour la procédure, se reporter à l'annexe 1) de dominance retenu ici (voir par exemple Davidson et Duclos, 2000) consiste à tester cette inégalité des montants de pauvreté pour un nombre fini de seuils de pauvreté non stochastiques. On a retenu les déciles et le 19^e vingtile des enquêtes *Revenus Fiscaux* correspondant aux années des enquêtes de BdF (10).

Des tests permettant d'isoler les cas d'égalité des chances au sens fort

La procédure retenue couple des tests de dominance faible et des tests d'égalité des distributions (11). On cherche à tester l'absence de relation de dominance. Habituellement, l'intérêt se focalise sur la recherche d'une relation de dominance. La procédure mise en œuvre ici diffère donc de la démarche usuelle. L'encadré 2 explicite les conséquences du choix de l'hypothèse nulle, celui de la dominance.

L'enchaînement des tests (12) est le suivant. Dans un premier temps, on teste l'hypothèse de dominance faible de la distribution A par la distribution B et réciproquement, 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 dominance sont acceptées ou si la dominance est rejetée dans les deux cas, on teste alors l'égalité des distributions. Si cette égalité 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, à l'égalité des chances. La même procédure est appliquée pour les distributions centrées. Cette procédure permet donc d'isoler l'égalité des chances au sens fort c'est-à-dire l'égalité des deux distributions.

Mesurer l'inégalité : revenus ou salaires ?

Les études portant sur les inégalités examinent en général soit le revenu du ménage soit le salaire de l'individu. On analyse ici l'inégalité des chances de revenu, ou plus exactement, de condition de vie. L'alternative, à savoir l'inégalité des chances de salaire, ne mesure que la réussite de l'individu sur le marché du travail, alors que le revenu du ménage permet de mesurer la réussite économique globale de l'individu

et recouvre à la fois la réussite professionnelle, le choix du conjoint et le patrimoine éventuellement hérité ou accumulé par l'individu. L'autre 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.

Les données d'enquête françaises qui se prêtent le mieux à l'étude du revenu des ménages sont les enquêtes *Budget de Famille* 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. On utilise 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.

L'information disponible permet une évaluation exhaustive des ressources perçues par les ménages. 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) (13). L'ensemble des revenus a été déflaté par l'indice des prix à la consommation de l'Insee (base 100 en 2002).

Deux mesures du revenu du ménage ont été distinguées : la première correspond au revenu primaire, 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). La seconde est le revenu disponible,

10. Plus exactement, les déciles de la distribution du revenu disponible par ménage par unité de consommation. Les valeurs de ces déciles sont dans le graphique III.

11. Il n'y a pas là de redondance. En effet, le test d'égalité des deux distributions est un test plus puissant que le test de dominance. En conséquence, la probabilité de ne pas rejeter l'hypothèse d'égalité 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'égalité. Si on attache une importance spécifique à l'égalité des distributions, il faut donc réaliser séparément le test d'égalité.

12. On a vérifié sur les données que l'ordre d'enchaînement des tests n'influe pas sur les résultats.

13. Le détail des sources de revenu déclarés à l'enquête et retenu dans cet article est présenté dans Lefranc et al. (2004).

égal au revenu net d'impôts et de transferts redistributifs. Dans les deux cas, la taille du ménage est prise en compte au moyen de l'échelle d'équivalence de l'OCDE (14). La pondération, enfin, permet aux données du niveau ménage d'être représentatives de la population totale des ménages français.

L'origine sociale appréhendée par la CSP regroupée du père

L'origine sociale du ménage est caractérisée par la catégorie socio-professionnelle du père du chef de ménage. Le code utilisé est la CSP à un chiffre, ce qui ne permet de distinguer que six groupes d'origine sociale : enfants de pères agriculteurs ; enfants d'artisans, commerçants et chefs d'entreprise ; enfants de cadres ; de professions intermédiaires ; d'employés ; d'ouvriers (15). Dans les enquêtes BdF, le chef de ménage et son conjoint sont interrogés sur la catégorie socio-professionnelle de leurs deux parents (CSP à un chiffre) (16). Il serait possible d'établir une nomenclature détaillée des milieux sociaux d'origine à partir de ces quatre variables. On a cependant renoncé à un tel code, car son information, plus riche, a pour contrepartie un découpage de l'échantillon des ménages enquêtés trop fin et conduisant à des estimations

peu précises des distributions de revenu conditionnelles à l'origine sociale. De plus, la CSP de la mère aurait posé des problèmes d'exogénéité par rapport à la présence d'enfants dans le ménage.

L'hypothèse d'égalité des chances est testée par rapport à la seule CSP à un chiffre du père, alors qu'il semblerait souhaitable de la tester par rapport à un code plus riche, tel que la CSP à un chiffre des deux parents du chef de famille. Cette relative pauvreté de la description du milieu d'origine n'a pas, cependant, des conséquences aussi néfastes qu'on pourrait le penser sur les tests de validité de l'hypothèse d'égalité des chances. On démontre que l'égalité des chances au sens fort par rapport à un conditionnement plus fin implique l'égalité des chances au sens fort pour un conditionnement plus grossier (cf. encadré 3). En conséquence, l'égalité

14. Cette échelle revient à diviser le revenu du ménage par la racine carrée du nombre d'individus du ménage. Les résultats ne changent pas si on utilise l'échelle de l'Insee ou l'échelle d'Oxford.

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

16. Pour toutes les enquêtes, sauf 1979, il s'agit de la catégorie socioprofessionnelle des parents lorsque l'individu enquêté avait 16 ans. En 1979, il s'agit de la dernière catégorie socioprofessionnelle d'activité des parents de l'individu.

Encadré 3

LES CONSÉQUENCES D'UNE DESCRIPTION PARTIELLE DU MILIEU D'ORIGINE

Le revenu des descendants est noté x , et le milieu social d'origine est maintenant fractionné en une partie observée s_1 , et une partie inobservée s_2 . Les observations ne portent que sur une distribution bivariable (x, s_1) , alors qu'il conviendrait d'observer une distribution trivariable (x, s_1, s_2) . On note la fonction de répartition de celle-ci F . F est supposée à support fini $X \times S_1 \times S_2$ inclus dans $[0, \bar{x}] \times [0, a_1] \times [0, a_2]$. F_x est la fonction de répartition de la distribution marginale du revenu et F_{s_1} , F_{s_2} sont les deux fonctions de répartition des distributions marginales des milieux d'origine. Soit F_x^{12} la fonction de répartition de x sachant s_1, s_2 (1).

Pour les besoins de l'exercice, il est plus facile de s'intéresser à l'égalité des chances au sens fort. On souhaite pouvoir constater **la pleine égalité des chances au sens fort** qui demande que :

$$\forall (x, s_1, s_2) \in X \times S_1 \times S_2, F_x^{12}(x | s_1, s_2) = F_x(x)$$

Faute d'observer s_2 , il n'est possible que de vérifier l'**égalité partielle des chances au sens fort**, définie par :

$$\forall (x, s_1) \in X \times S_1, F_x^1(x | s_1) = F_x(x)$$

La fonction de répartition de x conditionnellement à s_1 s'écrit également :

$$F_x^1(x | s_1) = F_{x2}^1(x, a_2 | s_1)$$

et, par définition d'une distribution conditionnelle :

$$F_{x2}^1(x, a_2 | s_1) = \int_0^{a_2} F_x^{12}(x | s_1, s_2) dF_2^1(s_2 | s_1)$$

On en déduit

$$F_x^1(x | s_1) = \int_0^{a_2} F_x^{12}(x | s_1, s_2) dF_2^1(s_2 | s_1)$$

ce qui permet d'énoncer :

la pleine égalité des chances au sens fort est une condition suffisante de l'égalité partielle des chances au sens fort.

1. Les probabilités conditionnelles sont notées avec la (les) variable(s) conditionnante(s) en exposant(s) et la (les) variable(s) conditionnée(s) en indice(s).

des chances au sens fort par rapport à la CSP des deux parents implique l'égalité des chances au sens fort par rapport à la seule CSP du père. Cette proposition est équivalente à sa contraposée : la non-égalité des chances au sens fort mesuré par rapport à la seule CSP du père est une condition suffisante de la non-égalité des chances au sens fort par rapport à tout conditionnement plus fin, et donc, en particulier, par rapport à la CSP du père et de la mère. En conséquence, toutes les fois que les données invalident l'hypothèse d'une parfaite égalité des chances, une information plus fine sur le milieu social conduirait au même résultat.

Réduire les risques de biais

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. 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 4 231 ménages en 1979, 4 428 en 1984, 3 529 en 1989, 4 644 en 1994 et 3 984 en 2000.

L'importance des non-réponses aux questions sur les revenus du ménage pourrait aussi conduire à des biais non négligeables : ces non-réponses varient en effet suivant la CSP du chef de ménage, et, au sein de chaque groupe d'origine sociale, suivant les caractéristiques observables des ménages. Aussi procède-t-on au redressement des réponses manquantes au moyen d'une procédure réduisant au minimum ces risques de biais (cf. encadré 4).

Les données de revenu d'exploitation agricole ont fait l'objet de vérifications (cf. Lefranc *et al.* (2004)). En effet, un changement important semble s'être produit durant cette période qui a eu pour conséquence de modifier le classement des CSP. Il était indispensable de s'assurer que les changements dont on va faire état n'étaient pas le fruit d'une illusion statistique.

Encadré 4

IMPUTATION DES REVENUS NON RENSEIGNÉS

Ignorer les non-réponses conduit à des biais

L'analyse des échantillons de l'enquête *Budget de Famille* fait apparaître, pour les vagues 1979, 1984 et 1989, 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 cette étude, à un phénomène aléatoire de non-réponse. Tout d'abord, comme le montrent Lefranc *et al.* (2004) le taux de non-réponse varie fortement en fonction de la catégorie socioprofessionnelle 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 individus : par exemple, on sous-estimerait implicitement la proportion de descendants 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 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.

Une procédure d'imputation qui élimine les biais

Pour ces raisons, en cas de non-réponse totale ou partielle, on a procédé à l'imputation du revenu du ménage par la méthode dite du résidu simulé. On estime tout d'abord, pour les ménages déclarant un revenu (en niveau ou en tranche), une équation de revenu, où le revenu est régressé sur une série de caractéristiques observables (âge, sexe, catégorie professionnelle, dernier diplôme obtenu, consommation, nationalité, structure familiale, zone géographique d'habitation, etc.) (1). 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.

1. Le détail des équations d'imputation est présenté dans Lefranc *et al.* (2004).

La hiérarchie des CSP d'origine change peu, mais les écarts se réduisent

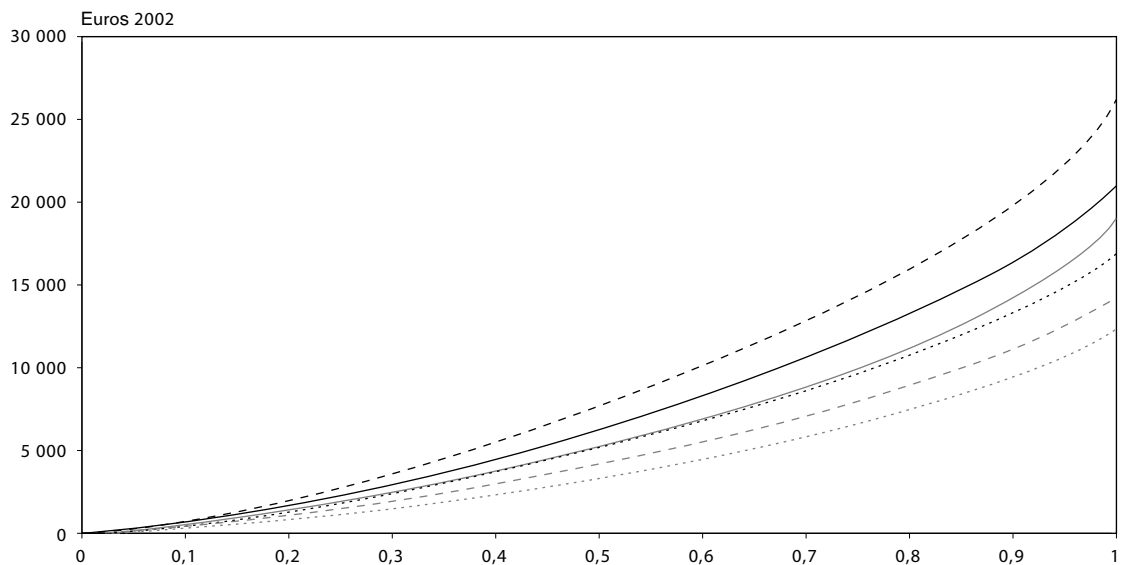
Pour étudier l'évolution de l'inégalité des chances depuis 1979, on procède tout d'abord à l'analyse des relations de dominance entre les différentes loteries sociales, pour chacune des vagues d'enquête disponibles. Deux enseignements principaux se dégagent alors. En premier lieu, la profonde permanence 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. En second lieu, les écarts entre les différentes loteries sociales tendent à se réduire au cours de la période étudiée, ce qui traduit une diminution du *degré* d'inégalité des chances.

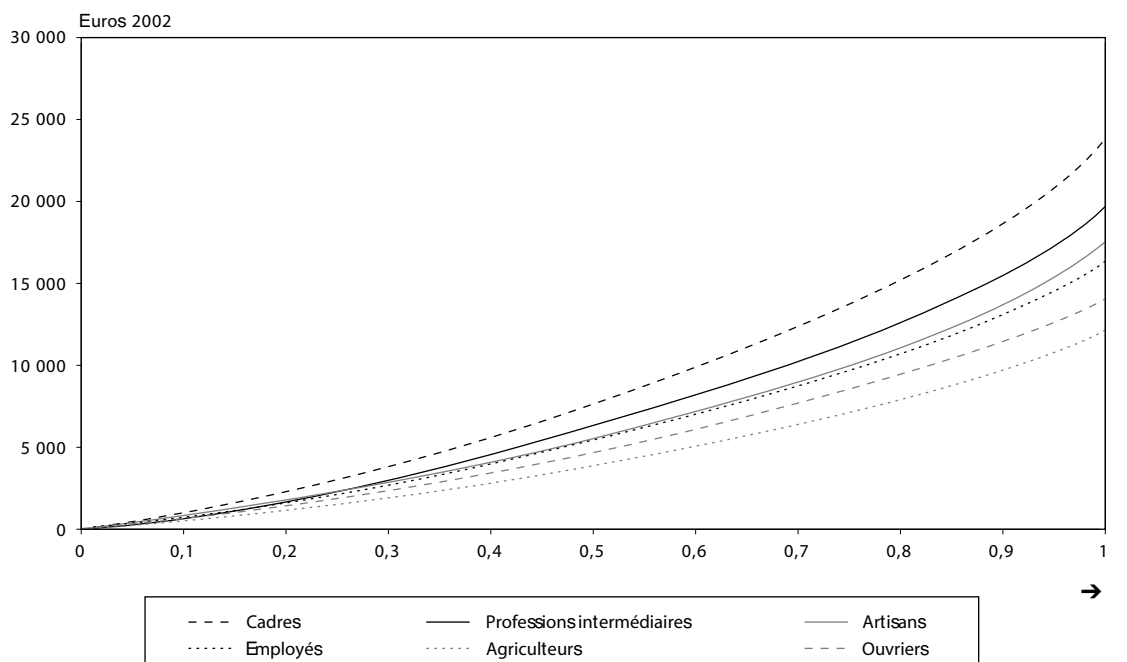
Le classement des distributions de revenu offertes aux individus en fonction de leur groupe social d'origine est particulièrement net (cf. graphiques II et III). La distribution de revenu des descendants de cadres domine très

Graphique II
Courbes de Lorenz généralisées : revenu primaire et revenu disponible

A - 1979 : avant impôts et transferts



B - 1979 : après impôts et transferts



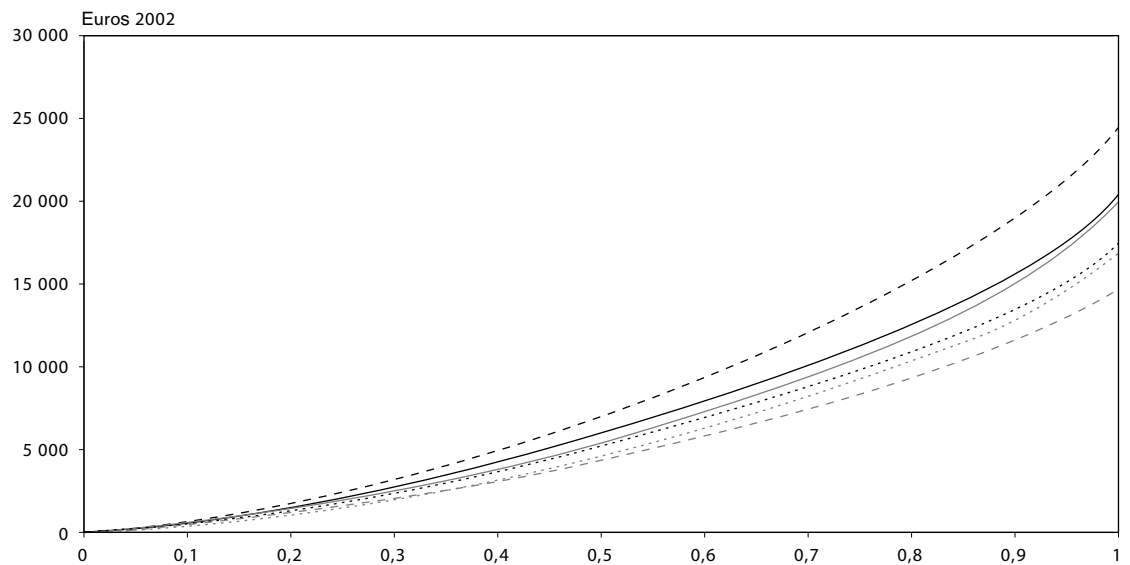
largement celle des autres groupes sociaux. Ils sont suivis, d'assez loin, par les individus issus des professions intermédiaires. Les descendants d'employés et d'artisans occupent une place intermédiaire et les distributions de revenu offertes aux individus de ces deux groupes apparaissent assez proches. Enfin, en bas de la hiérarchie se positionnent les descendants

d'ouvriers et d'agriculteurs, ces derniers nettement dominés par l'ensemble des autres milieux sociaux.

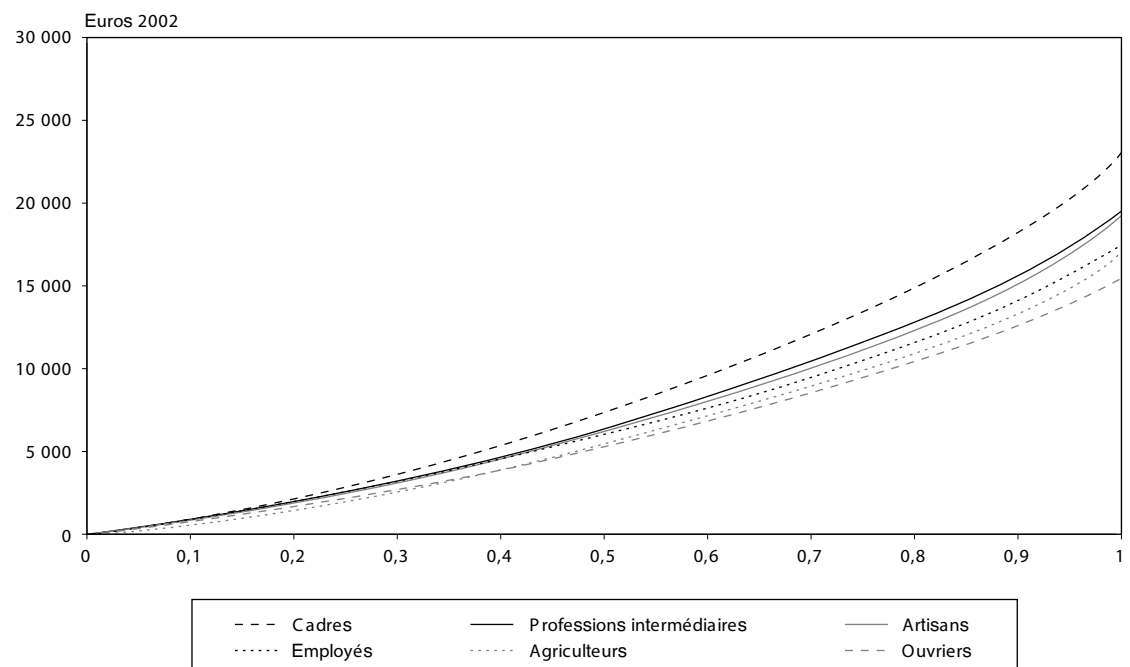
En 1979, cette inégalité des chances n'est que très peu corrigée par la politique redistributive. Celle-ci a pour effet de ne réduire que légèrement les écarts entre les courbes de Lorenz

Graphique II (suite)

C - 2000 : avant impôts et transferts



D - 2000 : après impôts et transferts



Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

généralisées des différents milieux d'origine (cf. graphique II). Les relations de dominance observées avant impôts et transferts sociaux se maintiennent toutes après redistribution.

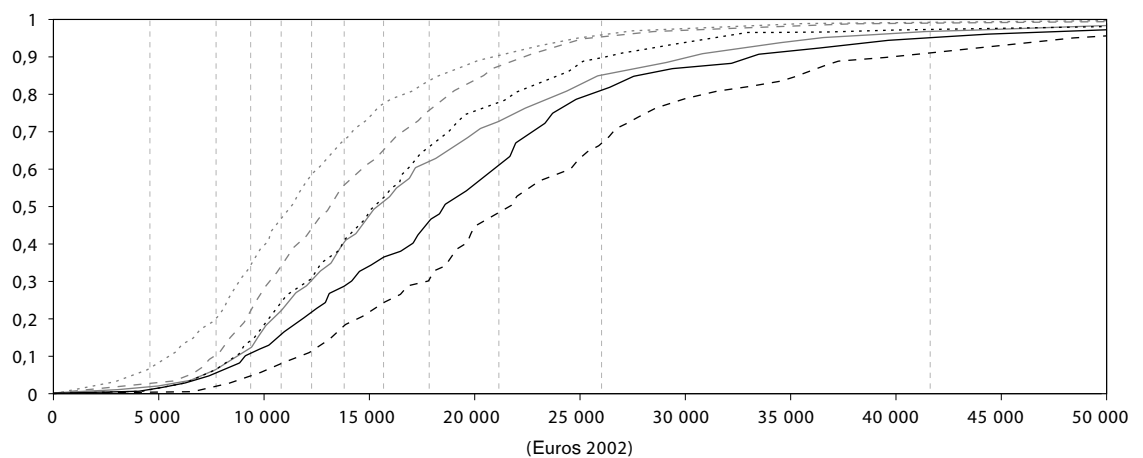
Ce classement est validé par les résultats des tests de dominance stochastique présentés dans les tableaux 1 et 2. Les critères de dominance permettent presque toujours d'établir un classement des différentes distributions : sur l'ensemble des vagues d'enquête, la proportion des cas où on ne conclut ni à l'égalité, ni à la dominance n'est que de 3 % (4/150 exactement). Ce faible

pourcentage justifie *ex post* le recours aux instruments de dominance stochastique pour définir l'égalité des chances. Ensuite, les relations de dominance stricte sont toutes obtenues avec un critère de dominance d'ordre 1, un critère particulièrement robuste, puisque la loterie dominante est préférée quelque soit le goût du risque du décideur.

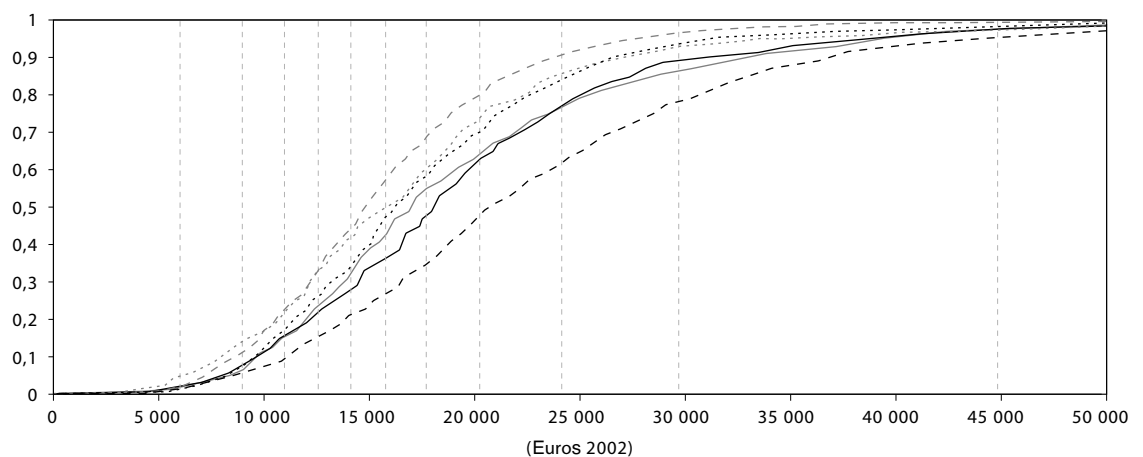
Les évolutions à l'œuvre entre 1979 et 2000 ne modifient que très partiellement les relations de dominance observées en début de période. Trois conclusions s'imposent au vu des tests : en premier

Graphique III
Distribution du revenu disponible selon le milieu d'origine

A - 1979 : après impôts et transferts



B - 2000 : après impôts et transferts



---	—	—
...	...	---
Cadres	Professions intermédiaires	Artisans
Employés	Agriculteurs	Ouvriers

Lecture : en 1979, les descendants d'agriculteurs ont 90 % de chances d'avoir un revenu disponible inférieur à 20 000 euros et les descendants de cadres, environ 45 %. Les traits pointillés verticaux correspondent aux seuils de pauvreté où sont effectués les tests statistiques (se reporter au texte).

Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

lieu, la position dominante des enfants de cadres par rapport aux autres milieux d'origine se maintient au fil des ans. En second lieu, la hiérarchie des groupes intermédiaires s'atténue. Ceci tient, pour partie, à l'amélioration de la position relative des descendants d'artisans au cours de la période : alors qu'ils étaient « dominés » en termes de revenu par les descendants de professions intermédiaires en début de période, ce n'est plus le cas à partir de 1989. Enfin, les positions relatives s'inversent en bas

de la hiérarchie des milieux sociaux d'origine : en 1994, la distribution de revenu conditionnelle des descendants d'agriculteurs devient non comparable à celle des descendants d'ouvriers, pour finalement la dominer en 2000. Les enfants d'ouvriers occupent désormais le dernier rang : cette catégorie socio-professionnelle est dominée, en termes de revenu, par l'ensemble des autres milieux d'origine (cf. tableau 1). Enfin, ce classement n'est que très peu affecté par la politique de

Tableau 1
Tests de dominance stochastique (avant impôts et transferts)

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers
1979						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	=	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1984						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	=	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1989						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	=	=	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	=	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1994						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	?
Artisans	-	-	< ₁	=	> ₁	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	=	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
2000						
Agriculteurs	-	< ₁	< ₁	< ₁	?	> ₁
Artisans	-	-	< ₁	=	> ₁	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-

Lecture : le tableau donne la situation de la catégorie en ligne par rapport à la colonne. <₁ : la ligne est dominée par la colonne à l'ordre 1. >₁ : la ligne domine la colonne à l'ordre 1. = : égalité des distributions. ? : non comparabilité.
Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

redistribution : les résultats des tests ne changent que dans 7 cas sur 150 (cf. tableau 2).

Ainsi, l'égalité des chances n'aurait-elle que peu progressé entre 1979 et 2000. Dans les deux premières vagues, 1979 et 1984, la comparaison bilatérale entre deux classes sociales, n'aboutit jamais à l'égalité des chances, à une exception près. Pour chacune des trois autres vagues, deux ou trois comparaisons au plus suggèrent l'éga-

lité des chances. L'incertitude provient essentiellement de la montée des deux classes de non-salariés par rapport aux classes salariées.

Ces classements suggèrent une hiérarchie à trois niveaux qui persiste tout au long de la période : un groupe dominant, formé des enfants de cadres, un groupe intermédiaire, de plus en plus homogène, formé des enfants d'employés, d'artisans et de professions intermédiaires ;

Tableau 2
Tests de dominance stochastique (après impôts et transferts)

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers
1979						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	=	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaire	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1984						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	< ₁	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1989						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	< ₁
Artisans	-	-	< ₁	< ₁	?	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
1994						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	> ₁
Artisans	-	-	< ₁	?	> ₁	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	=	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-
2000						
Agriculteurs	-	< ₁	< ₁	< ₁	< ₁	> ₁
Artisans	-	-	< ₁	=	> ₁	> ₁
Cadres	-	-	-	> ₁	> ₁	> ₁
Professions intermédiaires	-	-	-	-	> ₁	> ₁
Employés	-	-	-	-	-	> ₁
Ouvriers	-	-	-	-	-	-

Lecture : le tableau donne la situation de la catégorie en ligne par rapport à la colonne. <₁ : la ligne est dominée par la colonne à l'ordre 1. >₁ : la ligne domine la colonne à l'ordre 1. = : égalité des distributions. ? : non comparabilité.
Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

enfin, en bas de cette hiérarchie, un groupe formé des enfants d'ouvriers et d'agriculteurs.

La comparaison des courbes de Lorenz généralisées, pour la période initiale et la période terminale (cf. graphique II), conduit toutefois à nuancer ce constat. Si, comme l'indiquent les tests de dominance, le classement relatif des courbes évolue peu, l'écart qui les sépare se réduit nettement. La compression des courbes de Lorenz généralisées correspondant aux différents groupes sociaux d'origine permet de conclure à un changement de nature plus cardinale qu'ordinaire : les classements relatifs demeurent, mais le degré d'inégalité des chances diminue.

Ces évolutions traduisent-elles des changements dans le risque ou dans le rendement associés à telle ou telle origine sociale ? Une relation de dominance à l'ordre 2 peut provenir, en effet, soit d'un supplément de rendement à risque constant, soit d'une diminution du risque à rendement donné. Cela conduit à étudier l'évolution du risque inhérent à chaque loterie et celle du rendement espéré (revenu moyen conditionnel au milieu d'origine).

Des loteries sociales presque aussi risquées

Les courbes de Lorenz des différentes loteries sociales apparaissent profondément similaires (cf. graphiques IV et V). Avant impôts et transferts, les tests de dominance concluent à l'égalité au sens fort des distributions centrées autour de leurs moyennes dans neuf comparaisons (sur 15) en 1979, dans cinq cas en 1984, 12 en 1989, 11 en 1994 et 10 en 2000. On ne conclut à la dominance que dans un cas en 1979, trois cas en 1984, sept en 1989, deux en 1994, et jamais en 1989 et en 2000 (cf. annexe 2). Le degré de risque associé aux différentes loteries sociales de

revenu primaire semble donc très proche pour tous les milieux d'origine (17).

La redistribution ne modifie que peu cette conclusion, comme l'indique le tableau 3 qui donne les écarts maximum et minimum entre les courbes de Lorenz exprimés en points de Gini.

Les écarts minimum et maximum sont, pour chaque vague, d'un ordre de grandeur comparable, que l'on raisonne en revenu primaire ou en revenu disponible. Il y a au plus six points de Gini entre la distribution la moins risquée et la distribution la plus risquée.

Les distributions de revenu disponible diffèrent de celles de revenu primaire. Après redistribution, la courbe de Lorenz relative des enfants d'ouvriers domine plus souvent celle des autres catégories. Ainsi la distribution de revenu 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 descendants de cadres, est souvent dominée, donc plus risquée. Enfin, les loteries les plus risquées, après impôts et transferts, sont celles des enfants de non-salariés, notamment les agriculteurs en 1984, et les artisans en 1994, ce qui semble conforme à l'intuition.

C'est justement dans la comparaison entre salariés et non-salariés que le classement en termes de risque peut contredire le classement en termes de rendement (cf. *infra*). Par exemple, en 1994, le revenu disponible moyen des enfants d'artisans est plus élevé que celui des enfants d'employés. Toutefois, le test ne permet pas de

17. On peut aussi s'interroger sur l'évolution au cours du temps du risque associé à la distribution de chaque milieu d'origine. Les tests (non repris ici) indiquent que pour chaque groupe, le niveau de risque est resté globalement stable sur la période 1979-2000, à l'exception des enfants d'ouvriers pour qui la dispersion de la distribution de revenu conditionnel diminue légèrement.

Tableau 3
Écart entre les courbes de Lorenz (en points de Gini)

	Avant impôts					Après impôts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Écart minimum	0,000	0,007	0,002	0,008	0,001	0,005	0,001	0,001	0,000	0,007
Écart maximum	0,051	0,053	0,046	0,053	0,048	0,046	0,053	0,065	0,056	0,058

Lecture : L'indice de Gini est un outil classique de mesure des inégalités. Il est égal à deux fois la surface délimitée par la courbe de Lorenz et la première bissectrice. Par construction, l'indice de Gini est compris entre 0 (distribution uniforme : tous les ménages disposent du même revenu) et 1 (distribution où tous les ménages sauf un ont un revenu nul). Plus l'indice de Gini est proche de 1, plus l'inégalité mesurée est importante.

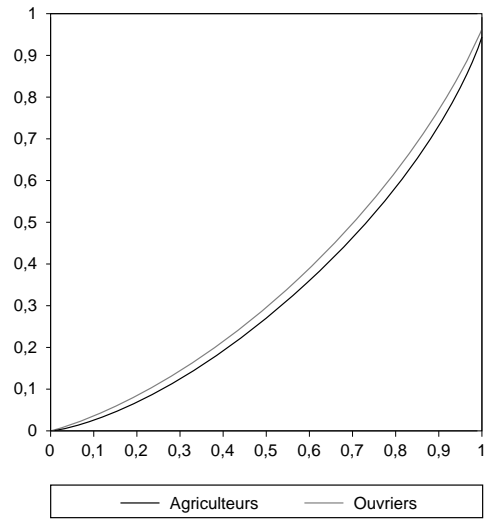
En 1979 l'écart maximum entre les courbes de Lorenz des distributions conditionnelles avant impôts est de 5,1 points de Gini.

Champ : personnes de référence du ménage de 30 à 50 ans.

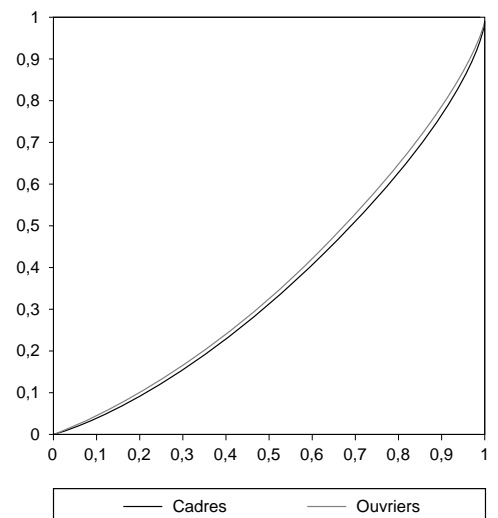
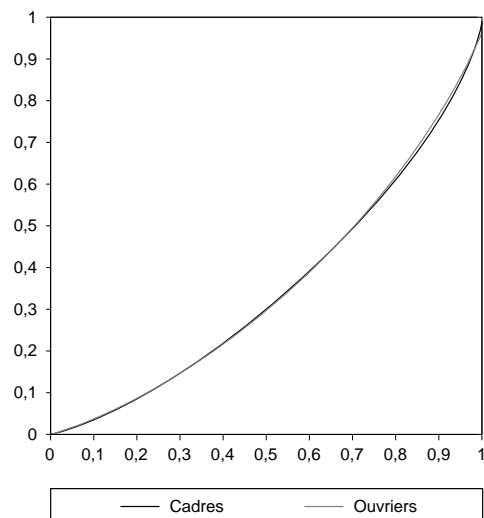
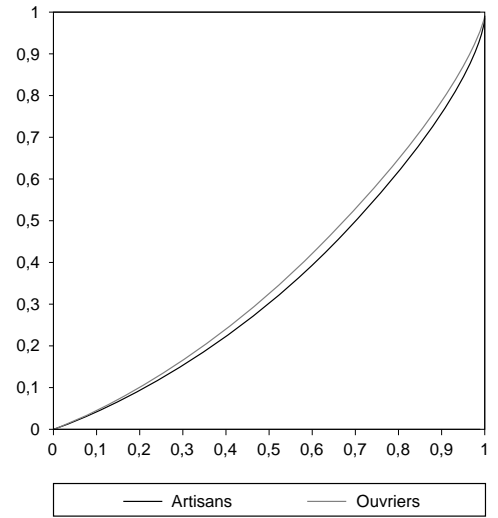
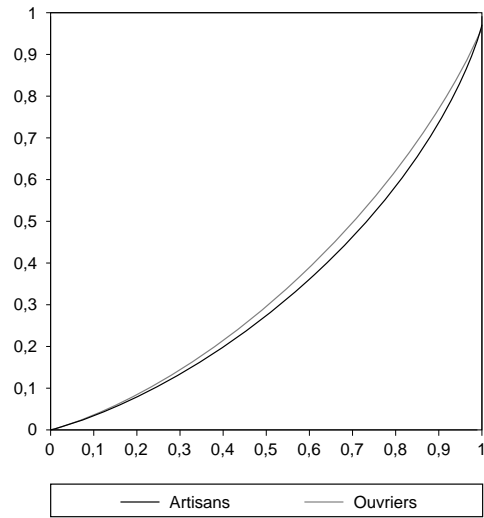
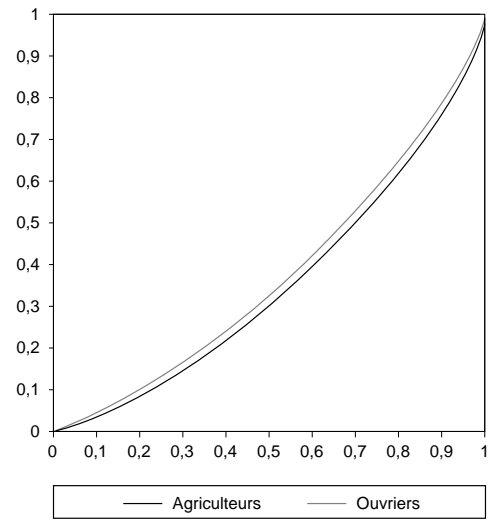
Source : enquêtes Budget de Famille, Insee.

Graphique IV
Courbes de Lorenz en 1979

A - Avant impôts et transferts



B - Après impôts et transferts



conclure à la dominance, en raison justement du plus grand degré de risque de la loterie des enfants d'artisans.

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

$$x_{is} = E(x|s) \times \varepsilon_i \quad (1)$$

où x_{is} désigne le revenu de l'individu i dont le milieu d'origine est s , $E(x|s)$ l'espérance de revenu conditionnel et ε_i un terme aléatoire distribué indépendamment de l'origine sociale. Dans la mesure où les loteries sont presque aussi risquées, se prononcer sur l'égalité des chances en

comparant les revenus moyens conditionnels, comme le propose Van de Gaer, n'exposerait ici qu'à des erreurs de faible importance (18).

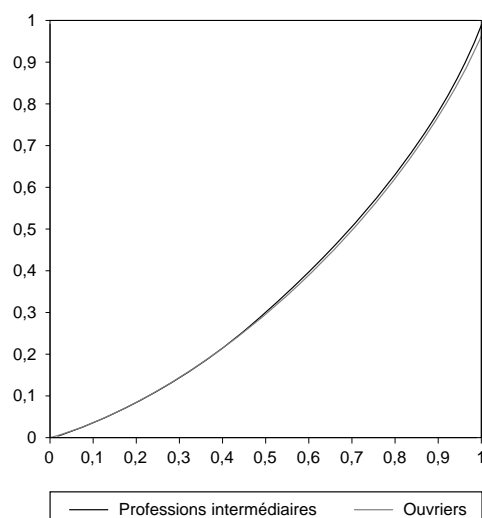
Un resserrement des revenus moyens conditionnellement à l'origine sociale

Que l'on raisonne avant ou après impôts et transferts, le tableau 4 indique un resserrement assez spectaculaire des revenus moyens conditionnels au cours de la période : le rapport entre le revenu moyen de la catégorie la plus favorisée

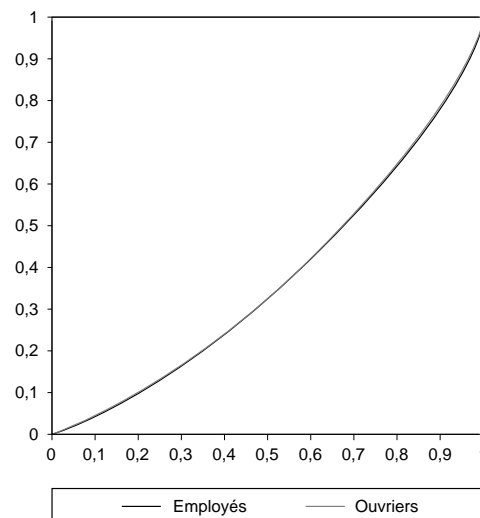
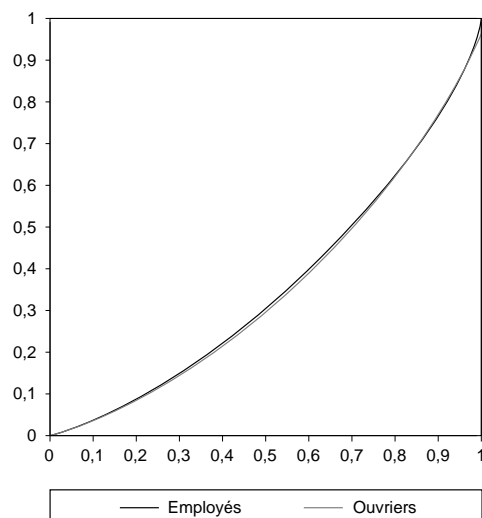
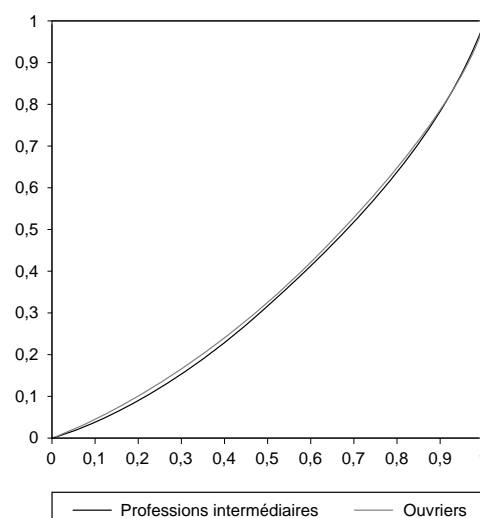
18. Le faible écart de risque entre les différentes loteries permet de mieux comprendre pourquoi les tests de dominance stochastique concluent, en général, à une dominance à l'ordre 1.

Graphique IV (suite)

A - Avant impôts et transferts



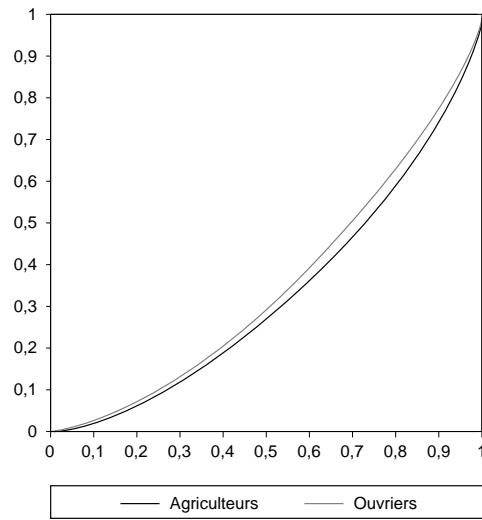
B - Après impôts et transferts



Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

Graphique V
Courbes de Lorenz en 2000

A - Avant impôts et transferts



B - Après impôts et transferts

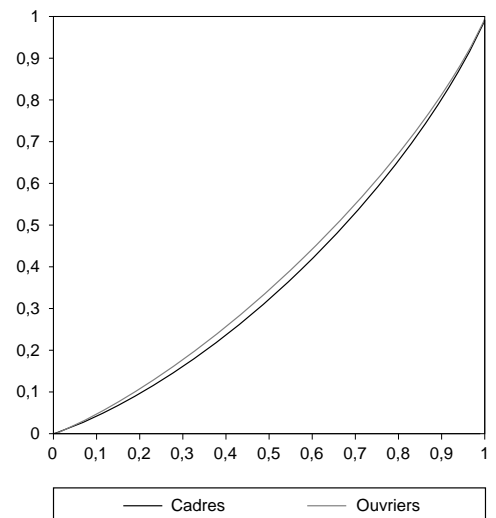
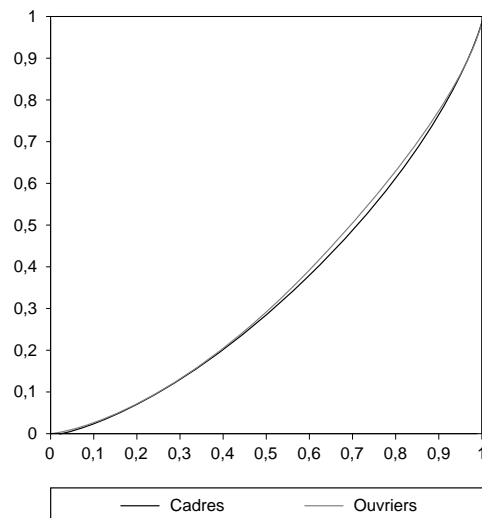
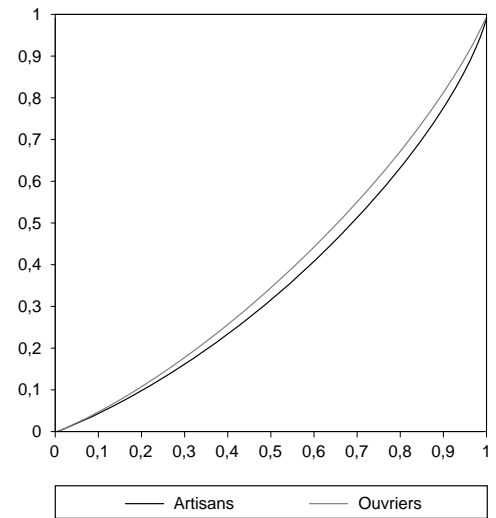
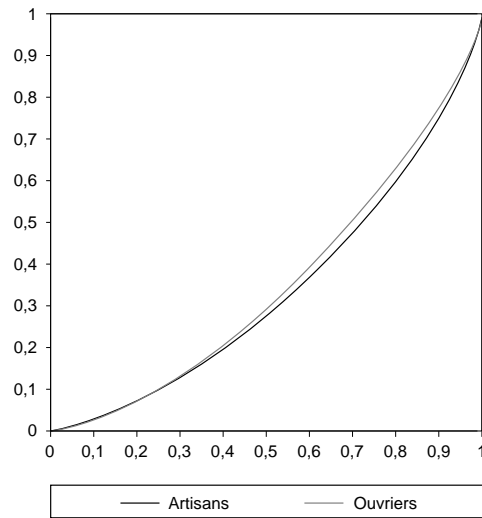
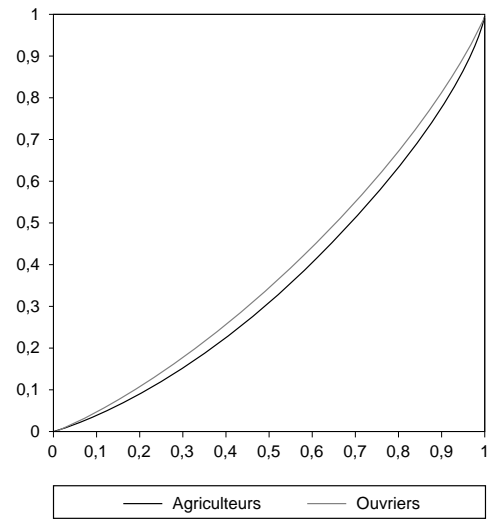


Tableau 4
Revenu moyen en fonction de l'origine sociale individuelle

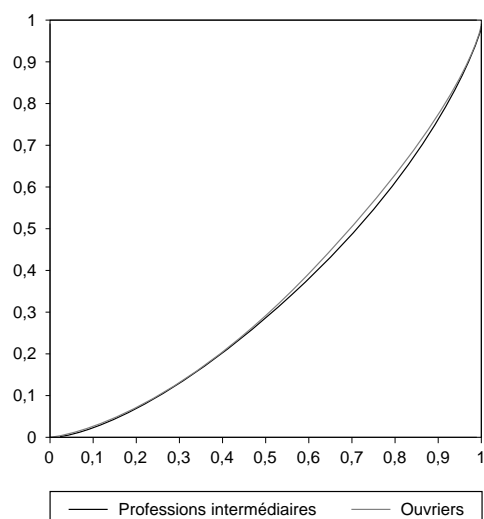
En euros 2002

	Revenu moyen									
	Avant impôts et transferts					Après impôts et transferts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Agriculteurs	12 874	14 079	14 058	16 839	17 541	12 914	14 018	14 219	16 728	17 395
Artisans	19 295	17 846	19 199	21 328	20 174	18 137	16 814	18 464	20 214	19 691
Cadres	26 375	26 126	26 435	25 038	24 543	24 490	23 694	23 975	22 946	23 033
Professions intermédiaires	21 225	21 752	20 668	20 122	20 511	20 055	20 292	19 748	19 156	19 534
Employés	17 379	18 195	19 753	18 633	17 720	16 884	17 486	18 970	18 053	17 747
Ouvriers	14 612	15 320	15 440	15 406	15 008	14 592	15 252	15 305	15 588	15 709
Ensemble	16 503	17 154	17 676	18 774	18 313	16 070	16 590	17 161	18 178	18 180

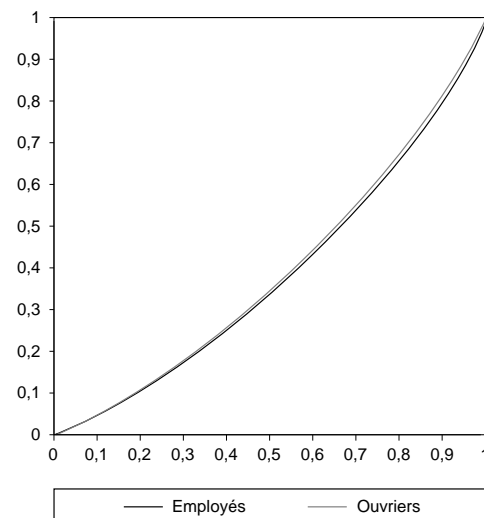
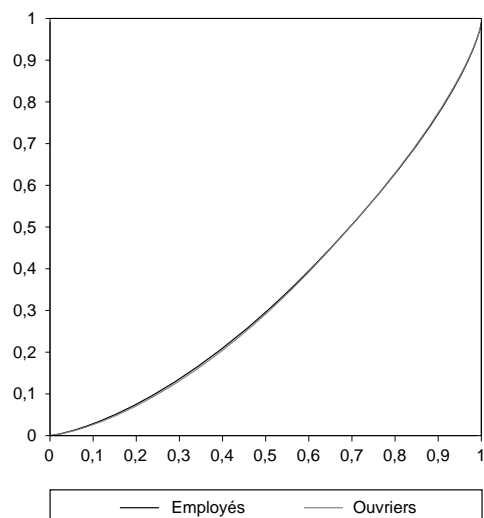
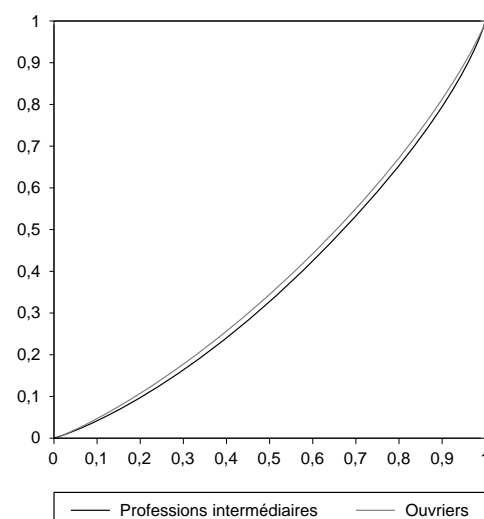
Lecture : la catégorie sociale est celle du père de l'enquêté.
 Champ : personnes de référence du ménage de 30 à 50 ans.
 Source : enquêtes Budget de Famille, Insee.

Graphique V (suite)

A - Avant impôts et transferts



B - Après impôts et transferts



Champ : personnes de référence du ménage de 30 à 50 ans.
 Source : enquêtes Budget de Famille, Insee.

(les descendants de cadres) et celui de la catégorie la moins favorisée (les descendants d'agriculteurs en 1979 et les descendants d'ouvriers en 2000) passe, entre 1979 et 2000, de 1,89 à 1,46 en raisonnant en revenu disponible et de 2,04 à 1,63 pour le revenu primaire. En l'espace d'une génération, l'avantage relatif des descendants de cadres, par rapport à la catégorie la plus désavantagée, s'est réduit de moitié.

L'érosion de l'avantage des descendants de cadres est plus accentuée par rapport à tous les autres milieux d'origine que par rapport aux seuls enfants de salariés : en limitant la comparaison à ces derniers, le revenu disponible moyen des descendants de cadres par unité de consommation a régressé de 5,9 % sur la période étudiée. Dans le même temps, celui des descendants de professions intermédiaires est resté pratiquement constant et celui des enfants d'employés et d'ouvriers ont progressé respectivement de 5,1 % et de 7,6 %. La progression est donc en relation inverse avec le classement de départ des CSP d'origine. Les comparaisons menées en termes de revenu primaire aboutissent à la même conclusion.

Cette conclusion s'impose aussi quand on compare descendants de cadres et descendants de non-salariés. Les différences d'ordre de grandeur sont impressionnantes : les enfants d'agriculteurs améliorent leur revenu moyen de 34,6 % au cours de cette période et le revenu des enfants d'artisans progresse de 8,5 %. En conséquence, les enfants de non-salariés voient leur situation relative s'améliorer, non seulement vis-à-vis des descendants de cadres, mais encore par rapport à l'ensemble des enfants de salariés. Ce mouvement a contribué à resserrer les inégalités au cours de la période, dans la mesure où la situation de départ des enfants de non-salariés était modeste : les enfants d'artisans n'occupaient que le troisième rang tandis que ceux d'agriculteurs fermaient la marche (19).

Le resserrement des revenus moyens selon la CSP de destination...

Deux facteurs peuvent expliquer l'évolution du revenu moyen conditionnel aux différents milieux sociaux. Tout d'abord, l'évolution de la mobilité sociale intergénérationnelle entre catégories socio-professionnelles : par exemple, la baisse de l'avantage relatif des descendants de cadres peut provenir d'une diminution de la probabilité que ces derniers deviennent eux-mêmes cadres. En second lieu, l'évolution des revenus moyens perçus pour chaque destination sociale

possible : ainsi, même si la distribution des CSP de destination des enfants d'agriculteurs reste inchangée, il est possible que le revenu moyen conditionnel de ces derniers progresse si le revenu moyen des agriculteurs augmente au cours de la période. Le premier effet est désigné sous le nom d'effet mobilité. Il provient de l'évolution de la reproduction sociale en termes de CSP. Le second, qualifié d'effet rentabilité, repose sur l'évolution de la hiérarchie des CSP de destination en termes de revenu.

L'effet rentabilité n'est autre que le resserrement des hiérarchies de revenu constaté au cours de la période (cf. tableau 5). 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 croît plus fortement. La réduction des disparités conditionnellement à l'origine sociale est cependant 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 presque constant au cours de la période : il passe de 2,28 en 1979 à 2,33 en 2000. De même, l'écart de revenu disponible est passé de 2,04 à 1,92. Dans le même temps, le revenu moyen primaire des enfants de cadres rapporté à celui des enfants d'ouvriers régresse de 1,80 à 1,63 et le ratio des revenus disponibles moyens tombe de 1,67 à 1,46 (20). L'évolution des revenus par CSP de destination ne suffit donc pas à expliquer la réduction du degré d'inégalité des chances. Celle-ci est en effet en partie imputable à un effet mobilité (cf. *infra*).

... se conjugue à la mobilité sociale pour réduire l'inégalité selon l'origine

Cet effet mobilité se mesure à l'aide de matrices de transition (cf. tableau 6). Celles-ci indiquent,

19. L'analyse de l'évolution du revenu des non-salariés doit évidemment être menée avec prudence, du fait de la fragilité des données relatives à ce groupe, en début de période, soulignée ci-dessus. Il est en effet possible que les phénomènes de non-réponse conduisent, même après imputation, à surestimer la croissance du revenu des non-salariés, en début de période. Cependant, la croissance des revenus estimée sur cet 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 montrent Lefranc et al. (2004) 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 l'enquête ont été recueillies avec plus de soin, comme en témoigne le faible pourcentage d'imputations réalisées pour ces années-là.

20. À titre d'exemple, en fin de période, l'écart de revenu moyen entre enfants de cadres et enfants d'ouvriers ne représente que la moitié de l'écart de revenu moyen entre cadres et ouvriers.

au moins pour certains milieux d'origine, une augmentation de la mobilité sociale. C'est le cas notamment pour les enfants de cadres 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 51 % en 1979 à 41 % en 2000. En

revanche, les enfants d'agriculteurs connaissent 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 diminue.

Tableau 5
Revenu moyen en fonction de la catégorie sociale individuelle

En euros 2002

	Revenu moyen									
	Avant impôts et transferts					Après impôts et transferts				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Agriculteurs	9 367	8 939	9 765	13 377	17 858	9 614	9 221	10 221	13 571	17 449
Artisans	17 090	14 545	19 853	21 432	19 833	15 797	13 584	18 229	19 801	18 972
Cadres	28 513	29 492	29 868	31 423	30 642	26 104	26 410	26 952	28 198	27 604
Professions intermédiaires	19 048	19 961	20 247	19 890	20 185	18 304	18 959	19 285	19 077	19 672
Employés	14 009	14 680	14 054	14 008	13 327	13 999	14 620	14 198	14 354	14 199
Ouvriers	12 264	12 548	12 673	12 975	13 113	12 738	13 141	13 364	13 752	14 357
Ensemble	16 503	17 154	17 676	18 774	18 313	16 070	16 590	17 161	18 178	18 180

Lecture : la catégorie sociale est celle de l'enquête.
Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

Tableau 6
Matrices de transition socioprofessionnelle

A - 1979

En %

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers	Total
Agriculteurs	22,5	7,4	4,7	11,6	14	39,6	21,6
Artisans	0,9	23,4	19,2	21,9	14,7	19,5	13,9
Cadres	0,9	3,3	51	26,1	13,3	5,0	6,9
Professions intermédiaires	0,0	4,5	28,7	40,1	13,8	12,7	7,3
Employés	0,6	5,7	14,4	29,5	19,7	29,8	11,1
Ouvriers	0,5	8,3	6,4	20,9	16,4	47,4	38,9
Total	5,3	9,3	13,4	21,8	15,6	34,3	100

B - 2000

En %

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers	Total
Agriculteurs	21,8	4,9	10	15,9	14,6	32,5	10,5
Artisans	0,4	11,8	20,5	25,1	19,1	22,9	12,2
Cadres	0,5	4,4	41,5	31,8	11,9	9,7	15
Professions intermédiaires	1,0	5,5	22,2	34,9	17,5	18,6	8,5
Employés	0,7	5,6	13,7	27,5	23,2	29	18,4
Ouvriers	0,5	4,7	6,6	19,2	19,8	49	35,1
Total	2,8	5,8	16,6	24,3	18,4	31,8	100

Lecture : les pourcentages sont en ligne. La catégorie sociale du père est en ligne, celle du fils en colonne. Exemple : en 2000, 10,5 % sont enfants d'agriculteurs. Parmi ces derniers, 21,8 % sont agriculteurs et 4,9 %, artisans.
Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

Pour quantifier la contribution respective de l'effet mobilité et celle de l'effet rentabilité, on utilise une décomposition de l'évolution du revenu conditionnel moyen inspirée d'Oaxaca (1973) et de Blinder (1973) (cf. encadré 5).

Les résultats de cette décomposition sont donnés dans le tableau 7. Elle confirme ce que suggérait la réduction des disparités selon l'origine plus forte que selon la CSP, constatée dans le cas des fils de cadres (cf. *infra*). L'effet mobilité et l'effet rentabilité contribuent tous les deux à une baisse du revenu moyen de cette catégorie. L'effet le plus déterminant est dans ce cas l'effet mobilité, qui explique 62 % de la variation du

revenu moyen. L'érosion de l'avantage relatif des enfants de cadres s'explique donc principalement par une transmission du statut de cadre à la génération suivante moins fréquente que par le passé.

L'effet mobilité (21) et l'effet rentabilité contribuent tous les deux positivement à la progression de la position sociale relative des descen-

21. Ces résultats sont à mettre en rapport avec ceux de Goldthorpe (1995) qui trouve une amélioration de 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.

Tableau 7
Décomposition de l'évolution du revenu moyen - revenu après impôts et transferts

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers
Variation du revenu selon l'origine sociale (1979-2000)	4 481	1 553	- 1 457	- 521	863	1 118
Première décomposition (en %)						
Effet rentabilité	83	122	38	- 75	121	117
Effet mobilité	17	- 22	62	175	- 21	- 17
Seconde décomposition (en %)						
Effet rentabilité	85	73	38	- 100	118	115
Effet mobilité	15	17	62	200	- 18	- 15

Lecture : la variation est une variation de revenu moyen, en euros 2002. La catégorie sociale est celle du père. Les décompositions sont données dans l'encadré 5.

Champ : personnes de référence du ménage de 30 à 50 ans.

Source : enquêtes Budget de Famille, Insee.

Encadré 5

DÉCOMPOSITION DE L'ÉVOLUTION DU REVENU CONDITIONNEL À L'ORIGINE

Cette décomposition s'inspire de Blinder (1973) et d'Oaxaca (1973).

Le revenu moyen conditionnel s'exprime 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 catégorie s l'année t , α_{sk}^t est la proportion d'individus issus de la catégorie sociale s qui aboutissent à la catégorie sociale k l'année t et \bar{y}_{sk}^t est le revenu moyen des individus de CSP k , issus de la catégorie sociale s l'année t .

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

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

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

où α_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 .

Dans chacune de ces décompositions, le premier terme correspond à un effet mobilité. Il mesure quel aurait été l'effet d'un changement de la matrice de mobilité sociale entre catégories sociales pour une distribution de revenus \bar{y}_{sk} inchangée. Le second terme correspond à 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 que l'on prend pour référence la matrice de mobilité ou la distribution de revenus conditionnels de la date initiale.

dants d'agriculteurs. Cependant, la majeure partie de cette progression s'explique par l'effet rentabilité. Le revenu des agriculteurs augmente très fortement au cours de la période 1979-2000 (+ 90,6 %), l'essentiel de la hausse intervenant dans la décennie 1990. Ceci bénéficie alors largement aux 22 % d'enfants d'agriculteurs devenus eux-mêmes agriculteurs.

De la même façon, une grande partie de la hausse du revenu des descendants d'artisans découle de la progression notable du revenu des artisans (environ 20 %), plus rapide que celle des salariés. L'effet mobilité est ambigu. Il est positif si l'on raisonne par rapport à la structure initiale des revenus par catégorie sociale (décomposition 2) et négatif si l'on raisonne par rapport à la structure terminale des revenus par catégorie sociale (décomposition 1).

Enfin, toutes les catégories de salariés connaissent un effet mobilité négatif ou nul. L'effet mobilité a le même ordre de grandeur pour les professions intermédiaires que pour les cadres (911 euros pour les professions intermédiaires et 903 euros pour les cadres). Par contre, l'effet rentabilité est positif, dans la mesure où le revenu des catégories modestes de salariés a continué de progresser plus vite que celui des cadres.

Si l'on compare 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, l'ensemble des milieux sociaux d'origine bénéficie de la mobilité sociale, à l'exception des descendants de cadres. Les pertes de revenu moyen liées à la mobilité sociale frappant cette catégo-

rie 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. En fin de période, ce sont les enfants d'employés qui tirent le mieux parti de la mobilité sociale en termes de revenu (cf. tableau 8).

Avant de s'interroger sur la contribution éventuelle de l'inégalité des chances à l'inégalité des revenus, il convient sans doute de s'interroger sur la robustesse des résultats obtenus jusqu'ici.

Une réduction de l'inégalité des chances, ou bien une modification de la composition des CSP ?

Les évolutions à l'œuvre entre 1979 et 2000 conduisent donc à une réduction du degré d'inégalité des chances, mesurée à partir du revenu des descendants, en appréhendant l'origine sociale par la CSP à un chiffre du père. Cette conclusion rejoint les constats des sociologues qui enregistrent une augmentation de la fluidité sociale en France entre 1970 et 1993 (Vallet, 2001 ; Goux et Maurin, 2003 ; Firth, 2003).

Ces résultats sont toutefois susceptibles d'être affectés par une déformation de la composition des catégories sociales d'origine. Ce problème provient de la conjonction d'une description assez fruste du milieu social d'origine (CSP à un seul chiffre) et d'une période d'étude de vingt ans suffisamment longue pour que des changements de grande ampleur dans la structure sociale se soient produits. Trois catégories sociales sont susceptibles d'être affectées.

Tableau 8
Rapport des revenus moyens conditionnels à l'origine et à la destination sociale

	Revenu avant impôts		Revenu après impôts	
	1979	2000	1979	2000
Agriculteurs	1,37	0,98	1,34	1,00
Artisans	1,13	1,02	1,15	1,04
Cadres	0,93	0,80	0,94	0,83
Professions intermédiaires	1,11	1,02	1,10	0,99
Employés	1,24	1,33	1,21	1,25
Ouvriers	1,19	1,14	1,15	1,09

Lecture : en 1979 le revenu moyen des descendants d'agriculteurs valait 1,37 fois celui des agriculteurs.
Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

La part des pères agriculteurs a fortement baissé, passant de 21 à 10 % (cf. tableau 6) et il y a tout lieu de penser que ce sont les agriculteurs les plus favorisés qui se sont maintenus. Le milieu social d'origine agriculteur évolue au cours du temps par perte des éléments les moins économiquement performants, et donc un *biais de composition* vers le haut pourrait affecter ce groupe social. Il est possible qu'un processus de même nature joue pour les artisans. Dans ces conditions, l'amélioration des perspectives des fils d'agriculteurs et d'artisans pourrait ne refléter que le déplacement vers le haut de la hiérarchie sociale des pères agriculteurs ou artisans. La dilution progressive du statut social de cadres, corrélative de la montée en puissance de cette catégorie dans la population active française, a pu également jouer. On peut en effet penser que l'augmentation de la part des cadres parmi les ascendants, passant de 7 à 15 % de la population, 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. Cette hétérogénéité accrue, voire ce biais de composition vers le bas, pourrait alors expliquer en partie l'érosion de l'avantage relatif des descendants de cadres observée plus haut.

Loin de se limiter aux seuls cadres, cette hétérogénéité croissante des positions individuelles au sein des différentes catégories socio-professionnelles a vraisemblablement 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. La thèse d'une hétérogénéité interne croissante des catégories

sociales apparaît d'ailleurs en accord avec certaines analyses antérieures (Herpin et Verger, 1999).

Cette diversification des positions individuelles au sein des différentes catégories socioprofessionnelles rendrait alors moins pertinent le conditionnement par la catégorie sociale d'origine. Cette hétérogénéité s'est sans doute 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 très agrégé des positions sociales, on peut alors ne pas saisir l'éventuelle permanence du rôle de l'héritage socio-économique, pour un certain nombre de sous-catégories sociales. Une appréhension plus fine du milieu socio-économique d'origine pourrait invalider le constat dressé ici d'une réduction de l'inégalité des chances.

Afin d'obtenir une description plus précise du milieu d'origine, on complète l'information sur la CSP du père de la personne de référence, par des informations sur la CSP du père de son conjoint (22). Les ménages dont la personne de référence est fils (ou fille) d'agriculteur ou de cadre sont décomposés en deux sous-catégories : un groupe plus favorisé et un groupe moins favorisé. Le tableau 9 montre l'évolution des revenus moyens de ces quatre sous-catégories sociales ainsi que les effectifs.

22. Si la CSP du père est clairement exogène à la réussite des enfants, cette réussite peut être corrélée au statut des beaux-parents.

Tableau 9
Revenu moyen après impôts et transferts selon la CSP du père et du beau-père

	Revenu moyen (en euros 2002)					Observations				
	1979	1984	1989	1994	2000	1979	1984	1989	1994	2000
Agriculteur - manuel	12 417	13 273	13 479	15 904	16 539	719	628	432	450	311
Agriculteur - non manuel	15 115	17 120	16 981	19 214	20 185	148	157	120	165	115
Cadre - non cadre	22 156	21 860	24 236	21 817	22 608	182	136	125	278	269
Cadre - cadre	28 010	25 087	23 674	23 721	23 482	134	151	130	400	303

Lecture : définition des catégories : « Agriculteur - manuel » : le père de la personne de référence (PR) est agriculteur et le père du conjoint (CJ) est agriculteur ou ouvrier ; « Agriculteur - non manuel » : le père de la PR est agriculteur et le père du CJ est employé, profession intermédiaire, artisan ou cadre ; « Cadre - non cadre » : le père de la PR est cadre, le père du CJ n'est pas cadre ; « Cadre - cadre » : le père de la PR est cadre, le père du CJ est cadre. Pour rendre comparables ces calculs à ceux du tableau 4, les célibataires n'ont pas été exclus de l'échantillon et on leur affecte un conjoint issu de la même CSP que la leur.

Champ : personnes de référence du ménage de 30 à 50 ans.
Source : enquêtes Budget de Famille, Insee.

Pour les ménages dont la personne de référence est descendant d'agriculteur, la part relative des individus issus du milieu plus favorisé (fils d'agriculteur et fille d'un père col blanc) augmente. Toutefois les revenus moyens des deux sous-catégories ont évolué dans les mêmes proportions. La progression des descendants d'agriculteurs ne s'explique donc pas uniquement par une augmentation du revenu ou de la part de la catégorie la plus favorisée. En d'autres termes, elle ne provient pas uniquement d'un effet de composition.

Parmi les ménages dont la personne de référence est descendant de cadre, on assiste également à une augmentation de la part relative du groupe le plus favorisé (de 42 à 52 %). Par contre, les revenus moyens de chaque sous-catégorie connaissent des évolutions radicalement divergentes. La catégorie la moins favorisée connaît une progression de son revenu moyen conforme à celle des autres catégories de salariés. Le revenu moyen des plus favorisés connaît une forte baisse pour devenir égal, à peu de choses près, à celui des moins favorisés. L'avantage ainsi procuré pour un descendant de cadre de s'apparier avec une personne de la même origine sociale a pratiquement disparu. La diminution de l'avantage relatif des descendants de cadres est donc concentrée sur la sous-catégorie dont les deux pères sont cadres. Un affaiblissement de la position relative des cadres dans l'échelle sociale affecterait à un double titre cette sous-catégorie et pourrait expliquer une telle évolution.

L'inégalité des chances mesurée par la CSP du père n'explique qu'une faible part de l'inégalité totale

On se propose enfin de mesurer quelle part de l'inégalité de revenu total provient de l'inégalité des chances. Deux quantifications simples de cette part sont proposées. La première utilise l'indice d'inégalité le plus utilisé, l'indice de Gini, la seconde l'outil favori des économètres, la décomposition de la variance. Les deux méthodes exploitent la similarité du degré de risque des différentes loteries sociales (cf. encadré 6). Gommer l'inégalité des chances revient alors à gommer l'effet moyen de l'origine sociale.

La première conduit à une contribution de l'inégalité des chances à l'inégalité totale de 5 %. Cette contribution est donc relativement faible. À titre de comparaison, l'action redistributive

de l'État conduit, en France, à une réduction de l'inégalité totale d'environ 15 %. Par ailleurs, avec des méthodes voisines et sur données brésiliennes, Bourguignon *et al.* (2003) obtiennent une contribution de l'inégalité des chances comprise entre 10 % et 20 % de l'inégalité totale brésilienne.

La seconde méthode consiste à utiliser une analyse de régression. Cela permet à la fois de tester la forme multiplicative de formation du revenu selon l'origine (équation 1 p. 68), et de mesurer la contribution de l'inégalité des chances à l'inégalité totale. Avec cette seconde approche, cette part passe de 11 % à 6 % au cours de la période 1979-2000. Cette seconde méthode confirme ainsi la faible part de l'inégalité des chances, mesurée par la CSP du père, dans l'inégalité totale.

La catégorie sociale du père est la seule circonstance dont on a cherché à mesurer l'impact. Bien d'autres circonstances jouent un rôle dans la réussite ou l'échec de la génération suivante, et la part de l'inégalité des chances dans l'inégalité totale est certainement sous-estimée du fait qu'elles ne sont pas prises en compte. Les estimations de cette part fournies précédemment doivent donc être considérées comme des bornes minimales. À titre de comparaison, la même décomposition de la variance des logarithmes a été appliquée au conditionnement étudié par Lefranc et Trannoy (2003), qui retiennent non pas un déterminisme social comme dans cet article (l'appartenance à un milieu social d'origine), mais un déterminisme individuel (le revenu des ascendants). Leur étude porte sur l'impact du revenu du père sur le revenu des enfants en France. Elle conclut à une contribution des inégalités de la circonstance « revenu du père » de l'ordre de 16 % (23). Cette contribution serait en première approximation constante entre 1964 et 1993.

*
* *

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. Les différences d'origine sociale conduisent à des différences significatives de niveau de vie. L'égalité des chances dans l'obtention du revenu ne semble donc pas pré-

23. Ce chiffre repose sur l'hypothèse que l'inégalité est restée constante d'une génération à l'autre.

valoir, tant en ce qui concerne le revenu primaire qu'en matière de revenu disponible après impôts et transferts redistributifs (24). L'inégalité des chances provient beaucoup plus des écarts de rendements espérés que des écarts de

risque inhérent à chaque loterie. Toutefois, seule une partie de l'inégalité de revenus se transmet d'une génération à l'autre.

24. Cette première conclusion recoupe par exemple celle de Vallet (2001) qui constate une forte stabilité des positions sociales.

Encadré 6

ESTIMER LA CONTRIBUTION DE L'INÉGALITÉ DES CHANCES À L'INÉGALITÉ TOTALE

Deux méthodes permettent de mesurer l'impact de l'inégalité des chances sur l'inégalité totale. Elles exploitent toutes deux la similarité du degré de risque des différentes loteries sociales.

La première utilise un indicateur d'inégalité, l'indice de Gini. On suppose qu'il est possible d'égaliser les revenus moyens conditionnels, grâce à une transformation homothétique des revenus des descendants d'une même CSP d'origine (1). Compte tenu de l'importante similarité des courbes de Lorenz mise en évidence précédemment, ceci conduit, en première approximation, à égaliser les distributions de revenus selon le milieu social d'origine et donc à assurer l'égalité des chances. L'inégalité résiduelle, qui correspond à la courbe de Lorenz des revenus centrés à leur moyenne conditionnelle, tous individus confondus, est alors aisée à calculer. 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 A livre une contribution de l'ordre de deux points de Gini, soit 5 % de l'inégalité totale.

La seconde façon de procéder consiste à utiliser une analyse de régression, à la fois pour tester la forme multiplicative du revenu suggérée plus haut (cf. équation 1 p. 68) et pour mesurer la contribution de l'inégalité des chances à l'inégalité totale. Pour cela on exprime la relation (1) en logarithme, avec η_i un terme d'erreur qui est un bruit blanc, soit :

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

avec $\beta = 1$

Une telle formulation évoque le modèle de Gibrat où le revenu d'une période passée est remplacé par l'espérance de revenu conditionnelle au milieu social d'origine qui joue le rôle de circonstance.

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 (cf. tableau B). Enfin, pour savoir si les résidus sont identiquement distribués entre les différentes origines sociales, on réalise des tests de Kolmogorov-Smirnov pour chaque paire de distributions des résidus. Sur 150 tests, on rejette l'hypothèse que les résidus sont identiquement distribués 24 fois (2), soit dans 16 % des cas seulement. Au total cette spécification multiplicative semble assez robuste, et la plupart des tests vont dans le même sens.

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

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

Cette écriture donne 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 du R^2 dans la régression donne exactement la contribution relative de la CSP du père à l'inégalité totale. Cette part diminue de 11 % à 6 % au cours de la période 1979-2000. Cette seconde méthode conduit à des résultats très voisins de la première : une part très faible de l'inégalité des chances par rapport à l'inégalité totale.

1. On ne tient pas compte ici des conditions et des répercussions économiques possibles de cette égalisation.
2. Sept fois avant impôts et 17 fois après impôts.

Tableau A
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 (en %)	6,56	4,71	5,24	4,46	5,08	6,64	4,89	5,50	4,60	5,33

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

Champ : personnes de référence du ménage de 30 à 50 ans.

Source : enquêtes Budget de Famille, Insee.



Encadré 6 (suite)

Tableau B
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	(0,046)	(0,052)	(0,055)	(0,061)	(0,060)	(0,047)	(0,052)	(0,054)	(0,056)	(0,057)
Observations	4 231	4 428	3 529	4 644	3 984	4 231	4 428	3 529	4 644	3 984
R^2	0,11	0,09	0,08	0,05	0,06	0,11	0,09	0,08	0,06	0,06
Test sur le coefficient $\hat{\beta} = 1$										
\hat{w}	1,67	2,84	0,22	0,43	0,14	0,93	0,59	1,41	2,49	3,46
Probabilité critique	0,196	0,091	0,639	0,511	0,708	0,334	0,442	0,235	0,114	0,063
Tests de normalité du résidu										
Probabilité critique	0,010	0,010	0,010	0,010	0,010	0,010	0,010	0,010	0,010	0,010

Lecture : la régression est en logarithme. Pour tester la normalité du résidu, on utilise le test non-paramétrique de Kolmogorov-Smirnov. L'hypothèse de normalité est rejetée à 1 %. Le test $\beta = 1$, effectué au moyen d'une statistique de Wald ($\hat{w} \sim \chi_1^2$) conduit à un non-rejet de l'hypothèse d'égalité.
 Champ : personnes de référence du ménage de 30 à 50 ans.
 Source : enquête Budget de Famille, Insee.

Par ailleurs, la politique redistributive contribue à réduire l'inégalité des chances mais dans une mesure relativement modeste : l'écart maximal de revenu moyen des descendants selon l'origine sociale est de 46 % en termes de revenu disponible, contre 63 % en termes de revenu primaire.

Une tendance à la réduction du degré d'inégalité des chances se dessine cependant. Ainsi, le milieu social le plus avantageux, celui des enfants de cadres, voit son avantage relatif par rapport à la catégorie la moins favorisée diminuer de moitié, au cours de la période 1979-2000, tandis que les fils d'agriculteurs et, dans une moindre mesure, les fils d'artisans entament un mouve-

ment ascendant. Ce résultat mériterait d'être confirmé par une appréhension plus fine du milieu socio-économique d'origine.

Enfin, 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 d'importance. □

Les auteurs remercient Gérard Forgeot de l'Insee pour la mise à disposition des données de la vague 2000 de l'enquête Budget de Famille, ainsi que François Bourguignon, Russell Davidson, Jean-Yves Duclos et Dirk Van de Gaer ainsi que les participants au séminaire Fourgeaud de la direction de la Prévision pour leurs remarques sur une version précédente de ce texte. Ils remercient également John Roemer de leur avoir fourni les données sur la Suède et les États-Unis.

BIBLIOGRAPHIE

- Arneson R. (1989)**, « Equality and Equal Opportunity of Welfare », *Philosophical Studies*, n° 56, pp. 77-93.
- Atkinson A.B., Glaude M., Olier L. et Piketty T. (2001)**, *Inégalités Économiques*, Rapport pour le Conseil d'Analyse Économique, La documentation Française.
- Beach C.M. et Davidson R. (1983)**, « Distribution-Free Statistical Inference with Lorenz Curves and Income Shares », *Review of Economic Studies*, vol. 50 n° 4, 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*, n° 82, pp. 262-272.
- Blinder A.S. (1973)**, « Wage Discrimination: Reduced Form and Structural Estimates », *Journal of Human Resources*, n° 8, pp. 436-455.
- Boudon R. (1974)**, « Education, Opportunity, and Social Inequality » in *Changing Prospects in Western Society*, MIT Press, Cambridge (MA).
- Bourguignon F. (1998)**, *Fiscalité et Redistribution*, Rapport pour le Conseil d'Analyse Économique, La documentation Française.
- Bourguignon F., Ferreira F.H. et Menendez M. (2003)**, « Inequality of Outcomes and Inequality of Opportunities in Brasil », *DELTA Working Paper*, n° 2003-24.
- Cohen G. (1989)**, « On the Currency of Egalitarian Justice », *Ethics*, n° 99, pp. 906-944.
- Dardanoni V. et Forcina A. (1999)**, « Inference for Lorenz Curve Orderings », *Econometrics Journal*, n° 2, pp. 49-75.
- Davidson R. et Duclos J.Y. (2000)**, « Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality », *Econometrica*, n° 68, pp. 1435-1464.
- Dworkin R. (1981)**, « What is Equality. Part 1: Equality of Welfare », *Philosophy and Public Affairs*, n° 10, pp. 185-246.
- Erikson R. et Goldthrope J.H. (1992)**, *The Constant Flux – A Study of Class Mobility in Industrial Societies*, Clarendon Press, Oxford.
- Firth D. (2003)**, « Overcoming the Reference Category Problem in the Presentation of Statistical Methodology », *Sociological Methodology*, n° 33, pp. 1-18.
- Foster J. et Shorrocks A.F. (1988)**, « Poverty Orderings and Welfare Dominance », *Social Choice and Welfare*, n° 5, pp. 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*, n° 36(1), pp. 61-79.
- Gollier C. (2003)**, *The Economics of Risk and Time*, MIT Press, Cambridge (MA).
- Goux D. et Maurin É. (1995)**, « Origine sociale et destinée scolaire », *Revue Française de Sociologie*, n° 36(1), pp. 81-23.
- Goux D. et Maurin É. (1997)**, « Meritocracy and Social Heredity in France: Some Aspects and Trends », *European Sociological Review*, n° 13, pp. 159-177.
- Goux D. et Maurin É. (2003)**, « On the Evaluation of Equality of Opportunity for Income : Axioms and Evidence », *Crest Working Paper*, Crest.
- Herpin N. et Verger D. (1999)**, « Consommation et stratification sociale selon le profil d'emploi », *Économie et Statistique*, n° 324-325, pp. 57-74.
- 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*, n° 54, pp. 1243-1248.
- Lefranc A. (1997)**, « Évolution des marchés du travail français et américains : quelques éléments d'analyse comparative », *Revue Économique*, n° 48(5), pp. 1041-1060.
- Lefranc A., Pistolesi N. et Trannoy A. (2004)**, « An Equalization of Opportunities for Income Acquisition in France? », miméo, THEMA.
- Lefranc A. et Trannoy A. (2003)**, « Intergenerational Earnings Mobility in France:

Is France more Mobile than US? » document de travail IDEP, n° 04-01, à paraître dans les *Annales d'Économie et de Statistique*.

McDonald J.B. et Xu Y.J. (1995), « A generalization of the Beta Distribution with Applications », *Journal of Econometrics*, n° 66, pp. 133-152.

Oaxaca R. (1973), « Male-Female Wage Differentials in Urban Labor Markets », *International Economic Review*, n° 14, pp. 693-709.

Ooghe E., Schokkaert E. et Van de Gaer D. (2003), « Equality of Opportunity vs. Equality of Opportunity Sets », *Public Economics Working Paper Series*, n° 240, 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 Extend do Fiscal Regimes Equalize Opportunities for Income Acquisition among Citizens? », *Journal of Public Economics*, n° 87, pp. 539-565.

Shorrocks A.F. (1983), « Ranking Income Distributions », *Economica*, vol. 50, n° 197, pp. 3-17.

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 », à paraître in *Social Mobility in Europe*, Richard Breen ed (2004), Oxford University Press.

Van de Gaer D. (1993), *Equality of Opportunity and Investment in Human Capital*, unpublished Ph. D. dissertation, K.U. Leuven.

Wolak F.A. (1989), « Testing Inequality Constraints in Linear Econometric Models », *Journal of Econometrics*, n° 41, pp. 205-235.

LA PROCÉDURE D'INFÉRENCE STATISTIQUE

Pour un seuil de pauvreté z donné, un estimateur sans biais et asymptotiquement normal du montant de pauvreté $D^g(z)$ est donné par :

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

où N est la taille de l'échantillon, x_j le revenu du j^{e} individu et $I(\cdot)$ 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 Σ sa matrice de variance-covariance asymptotique dont un estimateur est donné dans Davidson et Duclos (1).

On peut exprimer l'hypothèse de dominance à l'ordre g sous forme vectorielle. Soit δ le vecteur des différences des montants de pauvreté (2) calculées pour k seuils de pauvreté dans la population. On veut tester :

$$\mathbf{H}_0 : \delta \in \mathbb{R}_+^k \text{ contre } \mathbf{H}_1 : \delta \notin \mathbb{R}_+^k.$$

La procédure de test repose sur le vecteur $\hat{\delta} = (\hat{D}_B^g - \hat{D}_A^g)$ dont la matrice de variance-covariance est donnée par $\frac{\sum_A}{N_A} + \frac{\sum_B}{N_B}$ sous l'hypothèse d'indépendance des distributions A et B .

Ainsi, l'hypothèse nulle est 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 al.* (1992). On peut toutefois montrer qu'il 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. La seconde voie, retenue dans cet article, 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 consiste à examiner la distance du vecteur estimé $\hat{\delta}$ aux deux ensembles définissant respectivement l'hypothèse nulle et l'hypothèse alternative. Bien que les tests globaux de Wald utilisés ici soient plus complexes à mettre en pratique que les tests union-intersection, leur plus grande puissance (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. L'examen de l'égalité des chances combine ces deux types de tests.

Les tests d'égalité des distributions

L'égalité de deux distributions traduit une égalité des chances au sens fort, telle que discutée dans cet article.

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 $\mathbf{H}_0 : \delta = 0$. On peut alors montrer (Beach et Davidson, 1983 ; Davidson et Duclos, 2000) que sous l'hypothèse nulle, le vecteur des différences de montant de pauvreté estimé est asymptotiquement normal et on a :

$$\hat{\delta} \sim N\left(0, \frac{\sum_A}{N_A} + \frac{\sum_B}{N_B}\right)$$

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

$$T_1 = \hat{\delta}' \left(\frac{\sum_A}{N_A} + \frac{\sum_B}{N_B} \right)^{-1} \hat{\delta} \sim \chi_k^2$$

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

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 $\mathbf{H}_0 : \delta \in \mathbb{R}_+^k$ contre $\mathbf{H}_1 : \delta \notin \mathbb{R}_+^k$. 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 :

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

où $\|x\| = x' \Sigma^{-1} x$. Kodde et Palm (1986) démontrent que la statistique T_2 suit un mélange de loi du χ^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 δ soient strictement positifs. La loi de ce mélange de χ^2 n'est pas tabulée 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 (3).

1. Voir Davidson et Duclos (2000), théorème 1 p. 1441.

2. Le montant de pauvreté de la distribution B diminué du montant de pauvreté de la distribution A .

3. Il suffit de tirer 10 000 vecteurs normaux multivariés d'espérance 0 et de matrice de variance covariance δ , 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 $w(k, j, \Sigma)$).

TESTS DE DOMINANCE RELATIVE (AVANT IMPÔTS ET TRANSFERTS)

	Agriculteurs	Artisans	Cadres	Professions intermédiaires	Employés	Ouvriers
1979						
Agriculteurs	-	?	=	=	=	<
Artisans	-	-	=	?	?	?
Cadres	-	-	-	=	=	?
Professions intermédiaires	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
1984						
Agriculteurs	-	=	?	<	<	<
Artisans	-	-	=	?	=	=
Cadres	-	-	-	?	?	?
Professions intermédiaires	-	-	-	-	?	?
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
1989						
Agriculteurs	-	=	?	=	=	?
Artisans	-	-	=	=	=	=
Cadres	-	-	-	=	=	?
Professions intermédiaires	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
1994						
Agriculteurs	-	=	=	=	=	=
Artisans	-	-	=	<	=	<
Cadres	-	-	-	=	?	?
Professions intermédiaires	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-
2000						
Agriculteurs	-	=	?	?	?	<
Artisans	-	-	=	=	=	?
Cadres	-	-	-	=	=	=
Professions intermédiaires	-	-	-	-	=	=
Employés	-	-	-	-	-	=
Ouvriers	-	-	-	-	-	-

Lecture : le tableau donne la situation de la catégorie en ligne par rapport à la colonne. < : la ligne est dominée par la colonne. > : la ligne domine la colonne. = : égalité des courbes de Lorenz. ? : non comparabilité.

Champ : personnes de référence de 30 à 50 ans.

Source : enquêtes Budget de Famille, Insee.