



# Documents de recherche

Direction des études analytiques

*Mesures d'inégalité divergentes—Théorie et résultats empiriques*

par Michael C. Wolfson

(Revisé)

No. 66



Statistics  
Canada

Statistique  
Canada

Canada

## **DIRECTION DES ÉTUDES ANALYTIQUES DOCUMENTS DE RECHERCHE**

La série de documents de recherche de la Direction des études analytiques permet de faire connaître, avant leur publication, les travaux de recherche effectués par le personnel de la direction, les boursiers invités et les universitaires associés. Cette série a pour but de favoriser la discussion sur divers sujets, notamment le travail, la dynamique des entreprises commerciales, les pensions, l'agriculture, la mortalité, la langue, l'immigration, la statistique informatique et la simulation. On incite les lecteurs à faire part aux auteurs de leurs commentaires, critiques ou suggestions. Une liste des titres figure à l'arrière de ce document.

Les documents de la série sont distribués aux bureaux régionaux de Statistique Canada, aux représentants statistiques des provinces, aux instituts de recherche et aux bibliothèques spécialisées. Vous pouvez vous procurer une copie du document par internet: ***[www.statcan.ca](http://www.statcan.ca)***.

Pour obtenir un ensemble de résumés des documents de la série ou un exemplaire des documents (en français ou en anglais), veuillez communiquer avec:

Comité de révision des publications  
Direction des études analytiques, Statistique Canada  
24<sup>e</sup> étage, Immeuble R.H. Coats  
Ottawa, Ontario, K1A 0T6  
(613) 951-6325

# Mesures d'inégalité divergentes – Théorie et résultats empiriques

par Michael C. Wolfson\*

**No. 66**  
*(Révisé)*

**11F0019MPF No.66**  
**ISSN:1200-5231**  
**ISBN: 0-662-99165-6**

Prix: 5 \$ l'exemplaire, 25 \$ par année

24 Immeuble R.-H. Coats, Ottawa, K1A 0T6  
\*Statistique Canada et Institut canadien  
de recherches avancées (613) 951-8216  
Télécopieur: (613) 951-5643

**juillet 1997**

Je tiens à exprimer mes sincères remerciements à Tony Atkinson, pour m'avoir suggéré une collaboration avec James Foster en vue d'une analyse plus approfondie de la question de la mesure de la polarisation, à James Foster pour notre travail commun d'élaboration des concepts de mesure, à Brian Murphy, Geoff Rowe et Milorad Kovacevic pour leur soutien concernant les résultats empiriques et les problèmes statistiques, ainsi qu'à mes collègues du programme sur la croissance économique de l'ICRA, qui ont suggéré une analyse critique de certaines publications récentes sur les inégalités salariales. Je demeure le seul responsable de toute erreur ou omission, ainsi que de toutes les opinions exprimées. Wolfson (1994) constitue une version sensiblement abrégée de la présente communication.

Les auteurs assument seuls la responsabilité des opinions formulés dans le présent document qui ne représentent pas nécessairement le point de vue de Statistique Canada.

*Also available in English*



## **Table des Matières**

<i>Resumé</i>	<i>iv</i>
<i>Introduction</i>	<i>1</i>
<i>Première divergence -- Concepts fondamentaux</i>	<i>1</i>
<i>Deuxième divergence – Validité conceptuelle</i>	<i>9</i>
<i>Autres divergences – Problèmes statistiques</i>	<i>12</i>
<i>Résultats empiriques</i>	<i>13</i>
<i>Conclusion</i>	<i>20</i>
<i>Bibliographie</i>	<i>24</i>
<i>Notes:</i>	<i>27</i>

## **Resumé**

Certains indicateurs statistiques largement utilisés concernant l'inégalité ou la « disparition de la classe moyenne » sont peut-être trompeurs. Parmi les divergences entre les conclusions tirées et les données citées, on retrouve l'absence de distinctions entre les concepts d'inégalité et de polarisation, ainsi que l'utilisation de mesures d'« inégalités » scalaires non conformes au classement selon les courbes de Lorenz. En outre, le fait de s'attarder uniquement à une sous-population comme les travailleurs à plein temps de sexe masculin et de ne pas tenir compte de la variabilité d'échantillonnage peut conduire à des affirmations inexactes quant aux tendances de l'inégalité. Ces divergences sont illustrées à l'aide de données canadiennes sur les revenus du travail pour la période allant de 1967 à 1994.

**Mots clés:** inégalité du revenu, bipolarisation du revenu, la classe moyenne en voie de disparition, erreur d'échantillonnage







## **Introduction**

La présente communication traite principalement des méthodes d'analyse de la distribution des revenus, en particulier des fondements qui sous-tendent les affirmations sur l'ampleur ou la tendance de l'inégalité. On observe depuis quelque temps un regain d'intérêt important pour l'analyse des tendances de la distribution des revenus. Malheureusement, l'attrait pour ce genre d'analyse s'est accompagné d'une diversification indisciplinée des méthodes statistiques. Résultat : il arrive que des conclusions non confirmées par les statistiques citées soient énoncées. On examine ici les sources de ces divergences entre les conclusions tirées et les données sous-jacentes.

Une importante divergence découle du sens fondamental donné à la notion d'inégalité dans la distribution des revenus. D'autres divergences entre les conclusions tirées et les données analysées tiennent aux mesures statistiques, aux populations et aux définitions de revenus particulières choisies pour cerner la réalité qu'on veut mesurer. Enfin, la marge d'erreur de mesures répandues qui servent à justifier des tendances de l'inégalité des revenus soulève elle aussi des préoccupations. Nous allons examiner tour à tour ces différents problèmes.

### **Première divergence -- Concepts fondamentaux**

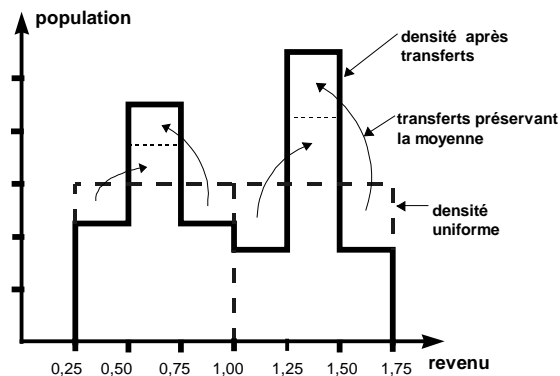
Depuis le début des années 1980, l'analyse des tendances dans la distribution des revenus prend de plus en plus la forme de discussions sur la « disparition de la classe moyenne » (p. ex. Kuttner, 1983; Thurow, 1984). On établit habituellement une équation entre ce concept relativement nouveau et celui d'une inégalité accrue des revenus. Toutefois, une telle équation suscite une question conceptuelle fondamentale. Dans leur récente enquête sur les tendances de l'inégalité des revenus aux États-Unis, Levy et Murnane (1992) cernent bien le problème : « ... une polarisation de la distribution des revenus signifie une diminution des emplois de la classe moyenne » (p. 1338) et plus loin « Malgré la variété des mesures scalaires (de l'inégalité), aucune ne semble convenir pour vérifier la proposition selon laquelle il y aurait une disparition de la classe moyenne. Cette proposition concerne une polarisation en vertu de laquelle les observations se déplacent du milieu de la distribution vers les deux extrémités. Les mesures d'inégalités courantes ne peuvent faire la distinction entre cette polarisation et d'autres sortes d'inégalités » (p. 1339). Par contre, les auteurs illustrent plus tard le manque de transparence dans ce domaine en affirmant, à l'encontre de leurs propos ci-dessus, que « si un creux apparaissait au milieu de la distribution des revenus des hommes, des mesures scalaires de l'inégalité permettraient d'enregistrer le phénomène » (p. 1351). [traduction]

Les expressions le plus couramment utilisées dans les études portant sur la polarisation et la disparition de la classe moyenne font ressortir la nécessité de préciser les concepts : « un creux au milieu de la distribution », « particuliers se déplaçant du milieu de la distribution vers les extrémités », etc. Ces expressions sous-

entendent que plus une distribution des revenus est polarisée, plus elle *s'étale* vers l'extérieur, de sorte qu'il existe moins de particuliers ou de familles ayant un revenu de niveau moyen. En outre, on a l'impression que cet étalement est également associé à un mouvement vers la *bi-modalité*, c'est-à-dire le regroupement de revenus anciennement de niveau moyen à des niveaux supérieurs ou inférieurs. Nous considérons ces deux notions comme essentielles au concept sous-jacent de la polarisation.

Du point de vue théorique, la polarisation définie de cette façon ne correspond pas à la notion d'inégalité officiellement définie dans la littérature (p. ex. Atkinson, 1970). La figure 1, tirée de Wolfson (1989) devrait dissiper tous les doutes. Ce graphique montre deux fonctions de densité hypothétiques de distribution des revenus. La première, qui est une densité uniforme, ou rectangulaire, sur l'intervalle 0,25 - 1,75, est représentée par un trait pointillé. La seconde, représentée par un trait continu, est manifestement bimodale, et est quelque peu renforcée au milieu. Selon notre interprétation du concept en vogue de la polarisation ou de la disparition de la classe moyenne, cette dernière densité est la plus polarisée.

**Figure 1 -- Polarisation et inégalité**



Cependant, la seconde densité bimodale a également été construite de telle sorte qu'en vertu de *n'importe quelle* mesure d'inégalité conforme aux critères de Lorenz – l'« étalon-or » du concept d'inégalité – elle présente une égalité supérieure. En d'autres termes, la densité bimodale a une courbe de Lorenz qui est plus près de la droite à 45 degrés que la courbe de Lorenz de la densité uniforme. La preuve formelle découle simplement du fait que la distribution bimodale peut être « dérivée » à partir de la distribution uniforme (de plusieurs façons, notamment) par deux ensembles de transfert de redistribution progressifs préservant la moyenne, au sens d'Atkinson (1970), comme le montrent les flèches de la figure 1.

Un ensemble de transferts de revenus égalisateurs s'effectue d'une partie des personnes de l'intervalle 0,75-1,00 de l'échelle de revenus (disons, les P) vers un nombre identique de personnes de l'intervalle inférieur 0,25-0,50 (les Q). Les P donnent aux Q une partie de leurs revenus égale à la moitié de la différence entre les

revenus des deux groupes – 0,25 en moyenne – de sorte qu’aussi bien les P que les Q passent à l’intervalle de revenus 0,50-0,75 dans la distribution bimodale. De même, un sous-ensemble des personnes se trouvant dans la partie supérieure de la distribution des revenus, soit l’intervalle 1,50-1,75 (disons, les M) donnent en moyenne 0,25 de leurs revenus à un nombre égal de personnes se trouvant dans la partie centrale supérieure de la distribution (les N), dont les revenus se situent entre 1,00 et 1,25. Par suite de ces transferts progressifs, aussi bien les M que les N passent à l’intervalle de revenus 1,25-1,50 de la distribution bimodale. Ainsi, par construction, la distribution bimodale est en même temps plus polarisée et plus égale que la distribution uniforme à partir de laquelle elle a été formée. La polarisation et l’inégalité sont donc clairement des concepts différents, comme il avait d’abord été signalé dans Love et Wolfson (1976), et comme le rappelle l’OCDE (1993)<sup>1</sup>.

Ce résultat n’offre pas de réponse, toutefois, quant à la méthode statistique qui conviendrait pour mesurer la polarisation. Dans la littérature traitant de la disparition de la classe moyenne, outre les mesures d’inégalité, on trouve des mesures comme les parts de revenus par quintiles, ou encore la fraction de la population appartenant à divers intervalles de revenus définis en fonction du revenu moyen ou médian – par exemple la proportion de la population ayant des revenus dans un intervalle de plus ou moins 25 % de la médiane. En fait, la figure 1 a été construite d’une manière peu valorisante pour ce genre de mesures.

Puisque la distribution est symétrique, la moyenne est égale à la médiane, c’est-à-dire 1,0. On peut montrer que la part *des revenus* du tiers central de la distribution bimodale est plus faible que la part des revenus du tiers central de la distribution uniforme, mais que la part des revenus des deux tiers centraux est plus élevée dans le cas de la distribution bimodale. Ainsi, les parts de revenus de divers groupes de quantiles centraux ne sont *pas* nécessairement des mesures compatibles avec une formalisation pertinente du concept de polarisation ou de disparition de la classe moyenne. Par conséquent, dans les articles qui prétendent analyser la disparition de la classe moyenne au moyen d’indicateurs d’inégalité comme les parts de revenus par quintiles (p. ex. Levy, 1987, Beach, 1988), la mesure utilisée est tout simplement inapte à détecter le phénomène qu’on affirme étudier.

De même, la proportion de la *population* ayant des « revenus moyens » est une mesure qui résiste mal au test imposé par notre exemple, car cette proportion augmente ou diminue selon la façon dont on définit « moyen ». L’examen de la figure 1 permet de le constater facilement. La population ayant des revenus dans un intervalle de plus ou moins 25 % de la moyenne (= médiane) diminue de toute évidence par suite des transferts, mais la population ayant des revenus dans un intervalle de plus ou moins 50 % de la moyenne (= médiane) augmente. Par conséquent, les mesures fondées sur la part de la population ayant des revenus « près de la moyenne » ne sont *pas* nécessairement elles non plus compatibles avec une définition pertinente de la polarisation. Par exemple, Thurow (1984) a fondé son analyse de la disparition de la classe moyenne sur la proportion de la population ayant des revenus entre 75 % et 125 % de la médiane, alors que Blackburn et Bloom (1985), dans une analyse

semblable, se sont intéressés à la proportion ayant des revenus entre 60 % et 225 % de la médiane.

Nous sommes donc confrontés à un nombre croissant d'articles qui cherchent à analyser les phénomènes de l'inégalité et de la disparition de la classe moyenne, au moyen d'une diversité incohérente d'indicateurs statistiques. Toutefois, l'axiome le plus fondamental de la théorie formelle de la mesure de l'inégalité – la condition de Pigou-Dalton relative aux transferts, qui équivaut, du point de vue formel, aux critères des courbes de Lorenz – est incompatible avec le concept de polarisation, ou les notions à peu près équivalentes d'étalement vers l'extérieur par rapport au point milieu, ou de bimodalité, inhérentes aux phénomènes de disparition de la classe moyenne. Une solution évidente consisterait à formaliser le concept de la polarisation d'une manière analogue au développement théorique des mesures d'inégalité.

La série de graphiques présentés à la figure 2 brosse effectivement un tableau sommaire d'une telle formalisation. On y découvre une dualité, ou une complémentarité, très intéressante entre la polarisation et l'inégalité. Le point de départ commun est une fonction de densité cumulative (fdc) pour la distribution des revenus (graphique 2.1 – il s'agit en fait d'une paire de distributions montrant le revenu disponible familial « équivalisé » pour le Canada et les États-Unis en 1988, le revenu canadien étant plus égal et moins polarisé).

Pour les mesures d'inégalité, le fondement conceptuel est associé à la courbe de Lorenz, montrée dans le coin inférieur gauche (graphique 2.3a). Sur le plan graphique, il s'avère utile de montrer une étape intermédiaire entre la fdc et la courbe de Lorenz. Cette étape consiste à permuter les axes de la fdc de façon que les centiles de population se trouvent sur l'axe horizontal et les revenus, sur l'axe vertical. On obtient ainsi le « défilé de nains (et de quelques géants) » de Jan Pen (1973), que présente le graphique 2.2a, après division de chaque revenu individuel par le revenu moyen. Cette courbe de « défilé » normalisée est ensuite intégrée depuis l'origine vers la droite, ce qui donne la courbe de Lorenz habituelle (graphique 2.3a).

La formalisation du concept de polarisation peut emprunter une voie semblable et parallèle de transformations graphiques de la paire initiale de fdc. Pour commencer, il est utile de noter deux démarches fondamentales pour mesurer la taille de la classe moyenne. La première débute dans l'« espace revenu » et recherche, pour une fourchette donnée de revenus (« M » sur l'axe horizontal du graphique 2.1), la part de la population totale qui a un revenu dans cette fourchette (« S » sur l'axe vertical). L'autre démarche (de dualité) a comme point de départ une fourchette de population dans l'« espace population » (« S ») et vise à cerner l'ampleur de la gamme de revenus couvrant cette population (« M »). La plupart des statistiques de ce genre en usage à l'heure actuelle (p. ex. la proportion de la population ayant des revenus entre 75 % et 150 % de la médiane) ont leur point de départ dans l'« espace revenu »; par contre, il sera plus commode d'élaborer la notion formelle de la polarisation à partir de l'« espace population ». Nous amorçons donc la présentation graphique de la même manière, en permutant les axes des fdc du graphique 2.1 pour obtenir le défilé de nains

de Pen, conformément à la construction du graphique 2.2a, mais nous poursuivons ensuite avec la série d'opérations suivante :

- au lieu de diviser les revenus par la moyenne, les revenus des particuliers sur l'axe vertical sont normalisés au moyen d'une division par la médiane;
- l'axe horizontal est ensuite déplacé vers le haut pour toucher le « défilé » normalisé selon la médiane qui en résulte, au point milieu de l'axe horizontal, le 50<sup>e</sup> centile de la population, qui correspond (par définition) au revenu médian (lequel, par suite de la normalisation, est maintenant égal à l'unité);
- la première moitié de la courbe de « défilé » visant les 50 % de la population dont le revenu est inférieur à la médiane (qui se trouvent désormais sous l'axe horizontal) fait enfin l'objet d'une rotation autour de l'axe horizontal.

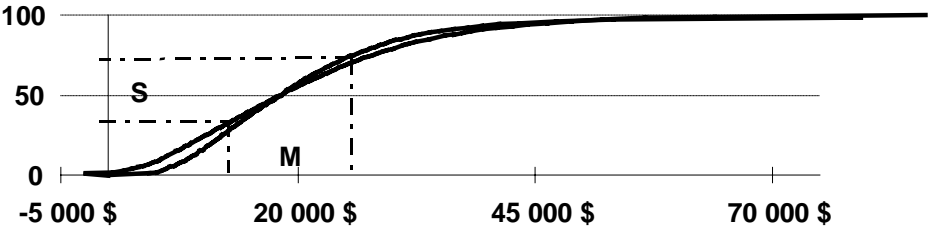
Le résultat est une courbe qui ressemble à un goéland disproportionné (graphique 2.2b). Cette courbe montre, pour n'importe quel centile de population le long de l'axe horizontal, la mesure dans laquelle le revenu de ce centile (exprimé en proportion de la médiane) s'écarte de la médiane. Il s'agit donc d'une indication de l'étalement vers l'extérieur, par rapport au point milieu (c'est-à-dire le 50<sup>e</sup> centile), de la distribution des revenus. (Pour une fourchette centrale donnée S le long de l'axe horizontal, M représente maintenant la somme des sommets de cette courbe d'étalement aux extrémités de la fourchette.) Une distribution moins étalée vers l'extérieur (c'est-à-dire ayant une classe moyenne plus peuplée) aura une courbe plus basse (et, si elle est plus basse partout, la courbe de Lorenz sera plus haute).

Pourtant, le concept de la polarisation renferme aussi un second aspect, celui de la bimodalité. La courbe d'étalement ou la « distance par rapport à la médiane » du graphique 2.2b ne le prend pas en compte parce qu'un transfert progressif d'un seul côté de la médiane donnera lieu à une deuxième courbe qui croise la première. Cependant, un tel transfert, comme celui de la paire montrée dans la figure 1 ci-dessus, augmentera le mode de son côté de la médiane et renforcera sans équivoque la bimodalité.

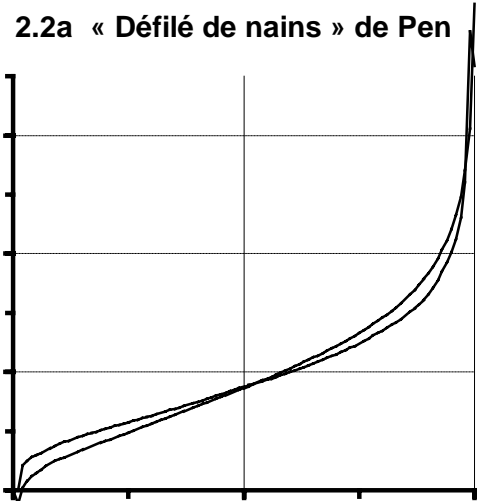
Il existe une transformation simple de la courbe d'étalement dans le graphique 2.2b qui la rendra simultanément sensible à ces deux attributs de la distribution, à savoir l'étalement vers l'extérieur et la bimodalité. Cette transformation correspond à la notion du passage du premier au deuxième ordre de la dominance stochastique. Sur le plan formel, nous intégrons la courbe d'étalement dans les deux directions à partir du point milieu le long de l'axe horizontal (où, par construction, la hauteur de la courbe est zéro) pour obtenir la courbe d'« étalement cumulatif » ou de polarisation (Foster et Wolfson, 1992) dans le graphique 2.3b. Cette courbe de polarisation classe n'importe quelle paire de distributions exactement de la même façon que les courbes d'étalement lorsqu'elles ne se croisent pas. De plus, elle classe les distributions dont les courbes d'étalement se croisent uniquement par suite d'une bimodalité accrue, exactement de la manière souhaitée. Cette courbe de polarisation joue donc le même rôle d'« étalon-or » pour le concept de polarisation que la courbe de Lorenz pour l'inégalité.

# Figure 2 Inégalité et polarisation : voies parallèles de développement graphique

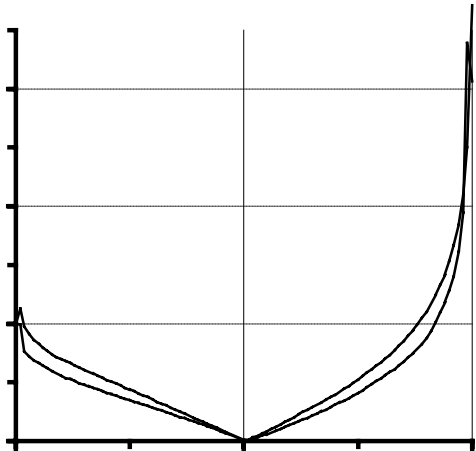
## 2.1 Deux fonctions de densité cumulatives



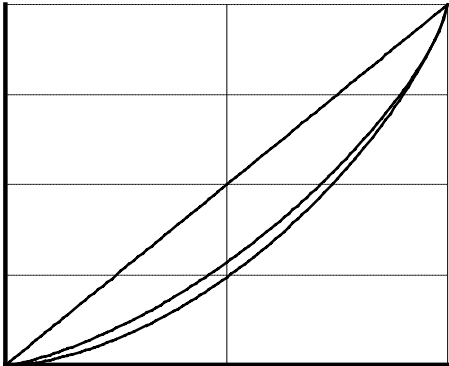
2.2a « Défilé de nains » de Pen



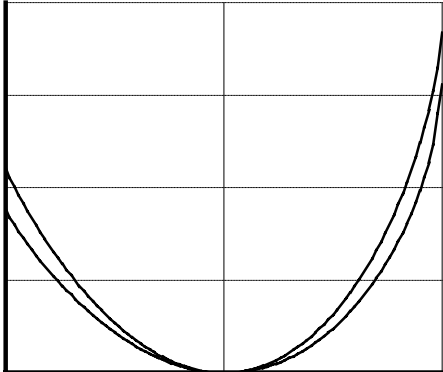
2.2b Première courbe de polarisation



2.3a Intégration à partir de la gauche pour obtenir la courbe de Lorenz



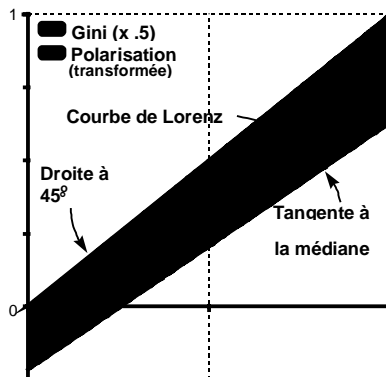
2.3b Intégration à partir du milieu pour obtenir la deuxième courbe de polarisation



Il s'ensuit que la surface sous cette courbe de polarisation, P, est un indice scalaire de polarisation, tout comme le coefficient de Gini, qui correspond à (deux fois) la surface entre la droite à 45 degrés et la courbe de Lorenz, est un indice scalaire d'inégalité. Toutefois, comme pour les courbes de Lorenz, il est possible d'avoir des courbes de polarisation qui se croisent. Ainsi, les courbes de polarisation (à l'instar des courbes de Lorenz) ne produisent qu'un classement partiel des densités de distribution des revenus en ce qui a trait à la taille de leur classe moyenne et à leur degré de bimodalité, tandis que les surfaces sous la courbe de polarisation P produisent un classement complet.

Les deux pistes de développement, qui avaient comme point de départ commun la fonction de densité cumulative de distribution des revenus, peuvent être de nouveau réunies grâce à une extension fort intéressante du concept de courbe de Lorenz. On peut voir une courbe de Lorenz type à la figure 3. L'ajout principal est la tangente de la courbe de Lorenz au 50<sup>e</sup> centile de population, et le prolongement de l'axe vertical jusqu'au croisement de cette tangente. Nous allons montrer que la courbe de polarisation que nous venons de décrire est étroitement liée à la courbe de Lorenz. Si nous normalisons de nouveau l'axe vertical de la courbe de polarisation (graphique 2.3b) en multipliant par le ratio entre la moyenne et la médiane, puis si nous faisons pivoter l'axe horizontal jusqu'à ce qu'il ait la même pente que la tangente de la courbe de Lorenz au 50<sup>e</sup> centile de population, la courbe de polarisation transformée que nous obtenons est identique à la courbe de Lorenz!

**Figure 3 -- Nouvelle mesure de polarisation fondée sur la courbe de Lorenz**



De même, la surface P sous la courbe de polarisation de la figure 2.3b correspond à la surface gris pâle de la figure 3 et est un indicateur scalaire de l'ampleur de la polarisation ou de la taille de la classe moyenne. Plus précisément, la surface gris pâle de la figure 3 entre la tangente et la courbe de Lorenz est égale à  $T - \text{Gini} / 2$ , de sorte que la surface P\* de la figure 2.3b est égale à  $(T - \text{Gini} / 2) / m \tan$ , où  $m \tan = \text{« tangente à la médiane »} = m / \mu = \text{pente de la tangente de la courbe de Lorenz au } 50^{\text{e}} \text{ centile de population}$ ,  $m = \text{médiane}$ ,  $\mu = \text{moyenne}$  et  $T = \text{surface du trapèze défini par la droite à}$

45 degrés et la tangente à la médiane = distance verticale entre la courbe de Lorenz et la droite à 45 degrés au 50<sup>e</sup> centile =  $0,5 - L(0,5)$  = différence entre 50 % et la part des revenus de la moitié inférieure de la population (cette dernière quantité,  $L(0,5)$ , sera appelée « part à la médiane »).

P a la valeur minimum zéro pour une distribution parfaitement égale des revenus et vaut 0,25 pour une distribution parfaitement bimodale dans laquelle la moitié de la population a un revenu nul et l'autre moitié, un revenu de  $2\mu$  (la médiane étant réputée correspondre à  $\mu$  dans ce cas)<sup>2</sup> : Afin d'avoir un intervalle de valeurs semblable à celui du coefficient de Gini (c'est-à-dire l'intervalle  $[0,1]$  s'il n'y a pas de revenus négatifs), nous définissons arbitrairement P comme ayant quatre fois la superficie normale, si bien que sa formule devient  $P = 2(2T - \text{Gini}) / m\text{tan}$ .

La courbe de Lorenz augmentée de la figure 3 montre clairement où peuvent survenir des conflits entre l'inégalité et la polarisation, et pourquoi ces concepts ont été si souvent confondus. S'il y a un « transfert d'égalisation » des revenus (conforme à la condition de Pigou-Dalton relative aux transferts) entre une personne se trouvant au-dessus de la médiane vers une personne dont le revenu se trouve sous la médiane (et que le transfert n'est pas assez grand pour que l'une ou l'autre personne traverse la médiane), l'inégalité et la polarisation diminuent toutes deux. Selon le diagramme, ce transfert de revenu fera en sorte que la courbe de Lorenz se rapproche de la droite à 45 degrés, ce qui réduira le coefficient de Gini, et déplacera parallèlement vers le haut la tangente au 50<sup>e</sup> centile de la population de manière à prouver que P chute aussi<sup>3</sup>. Aussi est-il clair que, dans de nombreux cas, les classements d'inégalité et de polarisation concorderont.

Il y aura conflit entre les deux concepts, toutefois, si les transferts d'égalisation s'opèrent entièrement d'un côté de la médiane – exactement comme c'était le cas dans l'exemple de la figure 1. Dans ces cas, la tangente à la médiane n'est pas modifiée par le transfert, mais la partie de la courbe de Lorenz se trouvant du côté concerné de la médiane se rapproche de la droite à 45 degrés. Un tel déplacement de la courbe de Lorenz réduit nécessairement le coefficient de Gini, ce qui accroît donc la mesure de polarisation P. Une telle divergence entre l'inégalité et la polarisation pourrait évidemment n'être qu'une curiosité théorique. Il importe de se demander si, en pratique, nous pouvons observer des tendances divergentes de ces deux types d'attributs des distributions des revenus. La réponse est oui, et des exemples seront présentés plus loin.

Le fait que le concept formalisé d'inégalité ne soit pas toujours en harmonie avec le concept de polarisation oblige à s'interroger, encore une fois, sur le fondement axiomatique des mesures d'inégalité. Plus précisément, cela soulève des questions quant à la condition de Pigou-Dalton relative aux transferts. Comme l'ont noté Amiel et Cowell (1989; note 14), Pigou avait des doutes sur sa validité. En outre, dans un sondage fait par ces auteurs auprès d'un groupe de près de mille étudiants du premier cycle en sciences économiques (qui, pour la plupart, n'avaient pas encore étudié ce sujet), une majorité d'étudiants ne retenaient pas cet axiome dans leur conception de l'inégalité. À tout le moins, un tel résultat porte à croire que pour répondre aux attentes



du grand public, l'analyse des tendances de la distribution des revenus devrait accorder autant d'importance à des mesures sommaires fondées sur des concepts comme la polarisation qu'à des mesures d'inégalité conformes au critère de Lorenz. En fait, la mesure formalisée de polarisation présentée ici est peut-être plus proche de l'idée courante d'inégalité dans le grand public que des mesures formelles d'inégalité fondées sur les concepts de Pigou-Dalton-Lorenz-Gini.

## **Deuxième divergence – Validité conceptuelle**

Outre la divergence conceptuelle fondamentale entre les idées relatives à l'inégalité et la polarisation, il existe d'autres divergences liées à la notion de « validité conceptuelle ». Une de ces divergences a trait à l'adéquation d'une mesure statistique donnée (*indicator*) à cerner le concept sous-jacent visé (*indicatum*). Par exemple, une série de questions permettant uniquement de savoir si la personne travaillait contre rémunération, non accompagnées de questions permettant de savoir si elle cherchait activement un travail rémunéré, ne serait pas un moyen valable de cerner le concept sous-jacent de la participation au marché du travail. De nombreuses mesures statistiques employées actuellement dans l'intention d'évaluer l'inégalité des revenus ne satisfont pas à ce critère de validité conceptuelle.

Prenons seulement le concept le plus largement reconnu, selon lequel l'inégalité se mesure par le classement partiel correspondant à la proximité entre la courbe de Lorenz et la droite à 45 degrés (sans tenir compte des critiques faites plus haut sur le plan de la polarisation). La dérogation la plus grave à la validité conceptuelle survient quand il y a une incohérence – par exemple si une distribution plus égale selon le classement de Lorenz est, selon la mesure utilisée, davantage *inégale*. Malheureusement, plusieurs mesures répandues échouent à cet égard. La plus courante est la variance des logarithmes (p. ex. Karoly, 1992; Davis, 1992; Katz et Murphy, 1992). Comme il est indiqué dans Love et Wolfson (1976) ainsi que dans Cowell (1977), pour des transferts d'*égalisation* supérieurs à environ 2,7 fois le revenu moyen (souvent aux environs du 95<sup>e</sup> centile), la variance des logarithmes indiquera une *augmentation* de l'inégalité<sup>4</sup>. Encore une fois, il est raisonnable de se demander s'il s'agit simplement d'une curiosité théorique; les résultats empiriques présentés plus loin donnent à entendre qu'il n'en est rien.

Un autre ensemble courant d'indicateurs d'inégalité est formé de ratios ou de différences entre des points de démarcation de quantiles de revenus, ou de leurs logarithmes. Par exemple, Davis (1992) s'intéresse au logarithme du ratio entre un centile de salaire supérieur et un centile de salaire inférieur (p. ex. ratios 90<sup>e</sup>/10<sup>e</sup>, 90<sup>e</sup>/médiane, médiane/10<sup>e</sup>), tandis que l'OCDE (1993) et Atkinson et coll. (1995) utilisent de tels ratios, mais sans prendre le logarithme. Ces mesures ne répondent pas elles non plus au critère de base de la validité conceptuelle. Imaginons, pour donner un exemple numérique simple, une société qui comprend trois personnes ayant comme revenus (1,5,9). Si la personne du milieu donne une unité de revenu à la personne du bas, la distribution résultante (2,4,9) est évidemment plus égale selon le

critère de la courbe de Lorenz. Toutefois, le ratio entre le 90<sup>e</sup> centile et la médiane augmente, ce qui indique un changement en direction opposée.

Cet exemple simple est évidemment extrême. Dans de nombreux cas, il sera beaucoup plus probable de voir de telles variations des ratios entre centiles de revenus reflétées par des courbes de Lorenz qui se croisent. Une solution minimale à cette forme de non-validité conceptuelle consiste simplement à ne pas appeler ces statistiques des indicateurs, ou des mesures, d'inégalité. Il existe, en fait, une interprétation directe de ces mesures : le ratio entre tout centile de revenus et la médiane est simplement la hauteur de la "courbe de défilé" de Jan Pen (1973) normalisée selon la médiane (plutôt que selon la moyenne) au centile de population concerné (figure 2.2a).

Il serait plus indiqué d'utiliser un ensemble rigoureusement sélectionné de ces indicateurs, plutôt qu'un ou deux pris isolément. Par exemple, Juhn et coll. (1993) complètent leur analyse de la tendance de l'inégalité fondée sur la variance des logarithmes en présentant des graphiques qui illustrent l'évolution du logarithme des salaires pour les centiles individuels compris entre le 10<sup>e</sup> et le 90<sup>e</sup>. Comme, dans l'ensemble, ces auteurs concluent que les variations proportionnelles des salaires de caractère monotone augmentent selon le centile, ces variations satisfont en réalité au critère de l'« étalon-or » de Lorenz en ce qui concerne l'inégalité. (Par contre, leur exclusion des déciles inférieur et supérieur signifie qu'ils ont omis de l'analyse le décile supérieur, soit la partie de la distribution où apparaîtra en toute probabilité l'incohérence de la variance des logarithmes avec la courbe de Lorenz.)

Puisque de nombreuses mesures d'inégalité largement utilisées manquent de validité conceptuelle, quelles statistiques devraient être utilisées? Il est généralement peu pratique de produire une multitude de courbes de Lorenz sur du papier-calque (ou sur un écran d'ordinateur) et de les comparer visuellement. Une solution consiste à choisir soigneusement un petit groupe de mesures d'inégalité et de polarisation.

Pour les mesures d'inégalité de Lorenz (c'est-à-dire les indicateurs conformes au classement partiel engendré par les courbes de Lorenz), un choix raisonnable, pour reprendre les termes de Cowell (1977), est d'utiliser trois mesures d'inégalité distinctes, l'une sensible dans la partie inférieure de la distribution, une autre sensible au milieu et la troisième sensible dans la partie supérieure, chacune étant strictement conforme aux classements des courbes de Lorenz<sup>5</sup>. Si les trois mesures concordent lorsqu'on compare deux distributions de revenus, on peut alors être modérément assuré que leurs courbes de Lorenz ne se croisent pas (du moins, pas de façon sensible). Toutefois, si les mesures ne concordent pas, nous savons que les courbes de Lorenz se croisent, et donc qu'aucun classement exempt d'ambiguïté n'est possible.

Dans un ouvrage antérieur (Wolfson, 1986), nous avons donné à entendre que nos mesures préférées d'inégalité sensibles dans les parties inférieure, centrale et supérieure étaient respectivement la mesure exponentielle (Exp), le coefficient de Gini et le (carré du) coefficient de variation (CV). La seule mesure peu connue est peut-être la mesure exponentielle<sup>6</sup>. Elle a été proposée justement parce qu'elle a l'avantage, par

rapport à d'autres mesures sensibles dans la partie inférieure, p. ex. Theil-entropie, Theil-Bernouilli (aussi appelée écart logarithmique moyen) et mesures de la famille décrite dans Atkinson (1970), de ne pas monter en flèche lorsqu'il y a des revenus de zéro ou voisins de zéro. En fait, cet ensemble de trois mesures d'inégalité accorde la priorité à la validité conceptuelle, à la définition adaptée à tous les niveaux de revenus concevables (y compris zéro et les montants négatifs) et à la probabilité élevée de témoigner du croisement des courbes de Lorenz et, par le fait même, de classements ambigus des courbes de Lorenz. Malheureusement, ces priorités excluent la plupart des mesures d'inégalité qui sont « facilement » décomposables (Shorrocks, 1980).

Néanmoins, tel qu'il est illustré empiriquement ci-dessous, cet ensemble rigoureusement sélectionné de mesures parvient difficilement à repérer les courbes de Lorenz qui se croisent. Ainsi, dans un autre ordre de validité conceptuelle, si le concept fondamental de l'inégalité est considéré comme le classement *parital* des distributions engendrées par les courbes de Lorenz, il faut se demander si cet ensemble étroit de mesures d'inégalité valides (dans le premier sens de la validité conceptuelle) traduit fidèlement des classements ambigus et non ambigus de l'inégalité.

Pour ce qui concerne le concept de polarisation, le candidat évident pour le choix d'une mesure valide (au premier sens) est la mesure P définie plus haut. Un ensemble de mesures beaucoup plus simples et plus commodes est obtenu en considérant la proportion de la population ayant un revenu, disons, entre 75 % et 150 % de la médiane, ou d'autres intervalles semblables (p. ex. entre 60 % et 225 %). Toutefois, comme nous l'avons signalé dans notre analyse de la figure 1, de telles mesures individuelles ne sont *pas* nécessairement compatibles avec le concept formel de polarisation qui a été élaboré, à savoir le classement *parital* découlant des courbes de polarisation du type exposé à la figure 2.3b. Il s'agit donc de mesures *invalides* de polarisation au premier sens du terme.

Par ailleurs, il est possible que les courbes de polarisation se croisent, tout comme les courbes de Lorenz. Par conséquent, le classement complet engendré par notre statistique P peut dissimuler des courbes de polarisation qui se croisent. En outre, malgré la relation très étroite entre la courbe de Lorenz et la courbe de polarisation illustrée à la figure 3, des courbes de Lorenz qui se croisent ne correspondent *pas* nécessairement à des courbes de polarisation connexes qui se croisent (et inversement). Intuitivement, d'après la figure 3, rappelons-nous que la courbe de polarisation dépend non seulement de la « courbure » de la courbe de Lorenz, mais aussi de la pente et de la hauteur de la tangente de la courbe de Lorenz à la médiane. Plus formellement, si nous désignons la courbe de polarisation au centile de population p par P(p) et la courbe de Lorenz par L(p), on a  $P(p) = (\mu / m) [L(p) - L(0,5)] + (0,5 - p)$ . La conséquence évidente est que le critère par excellence (l'« étalon-or ») pour les classements de polarisation devrait être l'examen des courbes de polarisation P(p).

Enfin, en ce qui a trait aux ratios comme le revenu au 90<sup>e</sup> centile par rapport au revenu à la médiane ou au 10<sup>e</sup> centile – par exemple, ceux décrits dans Atkinson et coll. (1995) –, rappelons qu'ils ne sont *pas* nécessairement conformes – et peuvent même être non

conformes – au classement partiel fondé sur les courbes de Lorenz. Il peut aussi être démontré qu'ils ne sont pas non plus nécessairement conformes au classement partiel fondé sur les courbes de polarisation. Ces statistiques, par conséquent, n'ont aucun avantage compensatoire à offrir pour la mesure des concepts formels d'inégalité et de polarisation, surtout lorsqu'elles sont employées isolément, sans mesures valides. Elles décrivent simplement quelques ordonnées d'une courbe de « défilé de nains » de Pen normalisée selon la médiane (figure 2.2a). Leur utilisation continue s'explique pour deux raisons : elles sont faciles à comprendre et sont largement disponibles (c'est d'ailleurs pourquoi elles ont été utilisées dans OCDE, 1993).

## **Autres divergences – Problèmes statistiques**

D'autres divergences des mesures d'inégalité peuvent être attribuables à des problèmes statistiques. L'un d'entre eux concerne la variabilité d'échantillonnage. L'immense majorité des analyses des tendances de l'inégalité des revenus omettent complètement d'examiner les erreurs d'échantillonnage sous-jacentes. On trouve des exceptions notables dans Burkhauser et coll. (1996), Bishop et coll. (1993) et Karoly (1992), mais même ces auteurs ne tiennent pas compte de la complexité des échantillons sous-jacents. Par exemple, l'enquête « Current Population Survey » de mars aux États-Unis et l'Enquête canadienne sur les finances des consommateurs sont chacune le fruit d'un plan de sondage complexe, c'est-à-dire un échantillonnage par grappes à plusieurs degrés. Love et Wolfson (1976), par une méthode de répétition de la moitié de l'échantillon, ont estimé la variance du coefficient de Gini et ont constaté que le plan de sondage complexe produisait des variances, pour le revenu familial, égales à près du double de celles auxquelles on se serait attendu si les données avaient été recueillies par un échantillonnage aléatoire simple.

Plus récemment, Kovacevic et Binder (1997) ont appliqué, au niveau des grappes, une méthode fondée sur des équations d'estimation, afin d'estimer la variance de plusieurs mesures d'inégalité et de polarisation, ainsi que les ordonnées de Lorenz et les courbes de polarisation, en tenant compte entièrement du plan de sondage complexe. Leur analyse a été fondée sur les mêmes données sur la distribution des salaires de 1991, tirées de l'Enquête canadienne sur les finances des consommateurs, qui ont servi à la présente analyse (tailles d'échantillons de l'ordre de 50 000 personnes). Pour l'essentiel, les intervalles de confiance à 95 % implicitement associés à des statistiques de distribution sommaires comme le coefficient de Gini et l'indicateur de polarisation  $P$  donnent à entendre qu'en règle générale, seulement les deux premiers chiffres de toute mesure d'inégalité d'usage courant ont une certaine fiabilité statistique et que, dans le cas d'une mesure sensible dans la partie supérieure comme le carré du coefficient de variation, seul le premier chiffre offre une certaine fiabilité statistique.

Il existe bien d'autres facteurs statistiques qui peuvent influencer sur les résultats<sup>7</sup>. Devant l'intérêt actuel pour la question de la distribution des salaires, il importe de s'interroger sur la population choisie pour l'analyse. La conclusion générale d'études comme celles de Davis (1992) et de Karoly (1992), c'est que l'inégalité des salaires s'amplifie depuis plus de dix ans. Toutefois, les populations incluses dans les analyses vont de l'ensemble des personnes ayant un revenu du travail positif (employés ou travailleurs

autonomes) à une population comprenant seulement les travailleurs à plein temps de sexe masculin travaillant toute l'année<sup>8</sup>. De plus, dans Davis (1992), en raison de la quantité limitée de données internationales disponibles, la qualité des comparaisons entre pays est peut-être entachée du fait que les données de chaque pays portent sur des populations très différentes. Le phénomène qui nous intéresse ici, soit la divergence des mesures d'inégalité, peut donc exister pour la simple raison que les comparaisons ne concernent pas des populations semblables.

## Résultats empiriques

Les travaux ont relevé cinq sources possibles de divergence dans les résultats des mesures d'inégalité ont été relevées :

- différences conceptuelles entre l'inégalité de Lorenz et la polarisation;
- ensembles restreints de mesures scalaires qui n'arrivent pas à identifier les ambiguïtés dans les classements partiels (c'est-à-dire les courbes de Lorenz ou courbes de polarisation qui se croisent);
- problèmes de validité conceptuelle de prétendues mesures d'inégalité qui ne sont pas conformes au concept de Lorenz;
- différences entre les populations observées;
- variabilité d'échantillonnage.

Dans la présente section, nous utilisons des résultats empiriques tirés d'une série chronologique de données des Enquêtes canadiennes sur les finances des consommateurs pour illustrer et évaluer ces divergences.

Nous allons utiliser deux ensembles chronologiques de données de distribution, soit un pour la distribution du revenu du travail des travailleurs à plein temps de sexe masculin et un pour la distribution du revenu du travail de l'ensemble des personnes ayant un revenu du travail annuel d'au moins 5 % du salaire moyen. Ces deux populations seront désignées respectivement "Hommes à plein temps" et "Ensemble des PEA" (PEA = personnes effectivement actives). Le revenu du travail comprend les salaires et traitements, la rémunération et les indemnités militaires, ainsi que le revenu de travail autonome (qui peut être négatif). Les Hommes à plein temps étaient âgés de 18 à 64 ans, travaillaient au moins 48 semaines et déclaraient travailler surtout à plein temps. L'Ensemble des PEA comprenait les femmes et incluait lui aussi les personnes âgées de 18 à 64 ans; toutefois, aucune autre condition ne visait les heures hebdomadaires ou le nombre annuel de semaines de travail.

Le tableau 1 présente les données relatives à ces deux populations, pour certaines années<sup>9</sup>. Il montre deux ensembles distincts de statistiques. Le premier mesure l'inégalité et le second, la polarisation. Les trois premiers indicateurs d'inégalité

correspondent à notre ensemble privilégié de mesures valides, à savoir le (carré du) coefficient de variation (CV), le coefficient de Gini et la mesure exponentielle, sensibles respectivement dans les parties supérieure, centrale et inférieure. Deux autres mesures sensibles dans la partie inférieure sont ajoutées, soit les indicateurs Theil-entropie et Theil-Bernouilli<sup>10</sup>. Enfin, les trois dernières mesures d'inégalité ne sont *pas valides* :

- la variance des logarithmes;
- le ratio  $[\text{LN}(D9 / \text{médiane}) / \text{LN}(\text{médiane} / D1)]$ , désigné « SJDavis » puisqu'il a été mis en relief dans l'analyse de l'inégalité de Davis (1992);
- le ratio  $D9 / D1$  utilisé par OCDE (1993) et Atkinson et coll. (1995)

où D9 et D1 représentent respectivement les revenus des 90<sup>e</sup> et 10<sup>e</sup> centiles. Nous avons inclus ces mesures pour illustrer leur *invalidité* conceptuelle en pratique.

Le deuxième ensemble de statistiques a trait au concept de polarisation. Les cinq premières statistiques donnent la proportion de la population appartenant à la « classe moyenne » en termes de revenus du travail, mais sous deux angles différents. Les deux premières, appelées « part de la population (%) dans l'intervalle de revenus (de) 75 %-150 % ou 60 %-225 % de la médiane », donnent respectivement les proportions des personnes ayant des revenus situés entre 75 % et 150 % de la médiane et les proportions des personnes ayant des revenus situés entre 60 % et 225 % de la médiane. Ces statistiques mesurent la taille de la classe moyenne en termes d'un intervalle de revenus normalisés selon la médiane. Pour les trois statistiques qui suivent, les axes sont en fait permutés, la classe moyenne étant définie dans l'« espace population » plutôt que dans l'« espace revenu ». Ces statistiques sont basées sur des intervalles symétriques de centiles de population – plus ou moins 10 %, 20 % et 30 % du 50<sup>e</sup> centile – et désignées respectivement « 40 %-60 % de la population », « 30 %-70 % de la population » et « 20 %-80 % de la population ». Pour chacun de ces intervalles de l'« espace population », l'intervalle correspondant de revenus divisé par la médiane est la statistique indiquée. Par exemple, la valeur de 0,341 donnée pour l'intervalle « 40 %-60 % de la population » (pour l'Ensemble des PEA en 1967) résulte du calcul suivant : revenu au 60<sup>e</sup> centile moins revenu au 40<sup>e</sup> centile divisé par la médiane. Même si la figure 1 examinée plus haut montre que l'une ou l'autre de ces statistiques considérée isolément peut être trompeuse, une concordance entre plusieurs données a beaucoup de chances d'indiquer un changement de polarisation sans équivoque, selon le concept que nous avons formalisé.

Tableau 1

## Certains indicateurs d'inégalité et de polarisation

	1967	1973	1981	1986	1988	1990	1991	1992	1993	1994
<b>Ensemble des PEA</b>										
<b>Inégalités valides</b>										
Carré du CV (0,0429)*	0,577	0,605	0,542	0,617	0,628	0,642	0,664	0,660	0,670	0,643
Gini (0,0027)*	0,363	0,378	0,377	0,396	0,395	0,398	0,402	0,404	0,405	0,399
Exponentiel (0,0012)*	0,446	0,451	0,450	0,458	0,457	0,458	0,460	0,461	0,462	0,459
<b>Inégalités - autres</b>										
Theil-entropie	0,243	0,259	0,251	0,276	0,277	0,280	0,288	0,290	0,292	0,283
Theil-Bernouilli	0,288	0,309	0,311	0,340	0,337	0,339	0,347	0,357	0,357	0,346
<b>Inégalités invalides</b>										
Variance des logarithmes	0,699	0,746	0,770	0,830	0,818	0,818	0,839	0,873	0,869	0,847
S.J.Davis	0,471	0,517	0,502	0,523	0,528	0,549	0,536	0,515	0,541	0,517
90 <sup>e</sup> /10 <sup>e</sup> centile	8,220	9,090	9,710	10,450	10,410	10,280	10,540	11,150	11,120	10,660
<b>Polarisation</b>										
part de la population (%) dans l'intervalle de revenus										
75-150 % de la médiane	41	37	36	32	32	34	32	32	32	33
60-225 % de la médiane	67	64	62	59	59	58	58	58	58	59
intervalle de revenus/médiane couvrant le groupe central										
40-60 % de la population	0,341	0,400	0,396	0,434	0,424	0,419	0,437	0,440	0,453	0,447
30-70 % de la population	0,700	0,792	0,815	0,912	0,900	0,877	0,889	0,926	0,935	0,886
20-80 % de la population	1,123	1,255	1,311	1,403	1,393	1,388	1,410	1,412	1,449	1,385
part à la médiane	0,243	0,229	0,226	0,213	0,215	0,214	0,210	0,208	0,207	0,211
médiane/moyenne	0,900	0,871	0,883	0,861	0,856	0,853	0,852	0,859	0,843	0,863
Polarisation (P, 0,0042)*	0,337	0,376	0,384	0,415	0,410	0,409	0,417	0,420	0,429	0,413
* erreur-type de 1991										
<b>Hommes – temps plein</b>										
<b>Inégalités</b>										
Carré du CV	0,328	0,273	0,234	0,288	0,306	0,327	0,331	0,332	0,355	0,337
Gini	0,261	0,245	0,242	0,263	0,266	0,274	0,272	0,273	0,275	0,274
Exponentiel	0,412	0,407	0,405	0,412	0,413	0,415	0,414	0,415	0,416	0,415
<b>Inégalités – autres</b>										
Theil-entropie	0,135	0,116	0,109	0,130	0,133	0,141	0,139	0,141	0,141	0,145
Theil-Bernouilli	0,140	0,117	0,120	0,143	0,142	0,151	0,145	0,149	0,149	0,149
<b>Inégalités invalides</b>										
Variance des logarithmes	0,308	0,248	0,273	0,331	0,316	0,339	0,312	0,328	0,328	0,318
S.J.Davis	0,774	0,827	0,775	0,654	0,701	0,744	0,726	0,688	0,713	0,717
90 <sup>e</sup> /10 <sup>e</sup> centile	3,370	3,100	3,250	3,680	3,590	3,690	3,780	3,760	3,690	3,710
<b>Polarisation</b>										
part de la population (%) dans l'intervalle de revenus										
75-150 % de la médiane	60	60	58	56	55	53	54	54	54	54
60-225 % de la médiane	81	84	83	79	79	79	78	78	78	78
intervalle de revenus/médiane couvrant le groupe central										
40-60 % de la population	0,190	0,194	0,201	0,220	0,224	0,249	0,237	0,230	0,230	0,246
30-70 % de la population	0,415	0,418	0,430	0,474	0,484	0,496	0,498	0,500	0,500	0,504
20-80 % de la population	0,705	0,702	0,712	0,763	0,787	0,803	0,809	0,799	0,799	0,831
part de la médiane	0,319	0,328	0,327	0,313	0,311	0,306	0,307	0,307	0,307	0,305
médiane/moyenne	0,917	0,920	0,936	0,945	0,924	0,913	0,912	0,923	0,923	0,919
Polarisation (P)	0,219	0,216	0,221	0,236	0,241	0,248	0,250	0,246	0,246	0,251

Les trois dernières statistiques liées à la polarisation découlent toutes de la courbe de polarisation/Lorenz présentée à la figure 3. Les deux premières ne sont que des ingrédients dans le calcul de l'indice sommaire de polarisation P, bien qu'il s'agisse d'indicateurs intéressants en soi. La première est la « part à la médiane » vue plus haut, c'est-à-dire la part des revenus revenant à la moitié inférieure de la population. Il s'agit exactement de la hauteur de la courbe de Lorenz à mi-chemin de l'axe horizontal, c'est-à-dire au 50<sup>e</sup> centile (par conséquent, il faudrait plutôt la considérer comme un indicateur d'inégalité). Par ailleurs, la quantité (0,5 - part à la médiane) correspond à la distance verticale entre la droite à 45 degrés et la courbe de Lorenz au 50<sup>e</sup> centile, donc à la surface T du trapèze renfermant la courbe de Lorenz à la figure 3. La deuxième statistique est le ratio entre le revenu médian et le revenu moyen  $m/\mu$ . Outre son interprétation graphique, c'est-à-dire la pente de la tangente de la courbe de Lorenz au 50<sup>e</sup> centile ( $m_{tan}$ ), ce ratio est un indicateur de l'asymétrie de la distribution. Enfin, la dernière statistique est la mesure de polarisation P que nous avons définie plus haut. Plus P est élevé, plus la polarisation est accentuée, et plus la classe moyenne est réduite.

Examinons maintenant les résultats du tableau 1 pour observer les divers types de divergences des mesures d'inégalité. Nous laissons de côté, pour le moment, la variabilité d'échantillonnage et supposons que les distributions ont été observées avec une infinie précision.

On peut observer le premier type de divergence entre les concepts d'inégalité de Lorenz et de polarisation. En règle générale, les mesures de ces deux concepts vont dans la même direction lorsqu'on compare par paires les distributions de revenus adjacentes. Toutefois, entre 1973 et 1981, pour les deux définitions de population active, toutes les mesures d'inégalité conformes au concept de Lorenz régressent sauf celles de Theil-Bernouilli. En même temps, toutes les mesures de polarisation sauf une augmentent – les parts des deux groupes de population ayant un revenu moyen (75 %-150 % et 60 %-225%) baissent; on observe un élargissement de tous les intervalles de revenus (sauf un) dont on a besoin pour délimiter diverses populations moyennes symétriques, ainsi que de l'indice sommaire de polarisation. Ainsi, selon ce qui est de loin le style analytique le plus courant dans la littérature, qui fait appel à un petit nombre de mesures sommaires plutôt qu'à la distribution intégrale, nous sommes en présence d'un cas où la polarisation et l'inégalité vont dans des directions opposées.

Par contre, ce cas illustre également le deuxième type de divergence, où même nos trois mesures valides de l'inégalité, choisies avec soin, n'arrivent pas à repérer les courbes de Lorenz qui se croisent. Le tableau 2 présente une série d'ordonnées de courbes de Lorenz, avec un ensemble de symboles indiquant des écarts statistiquement significatifs, dont nous reparlerons plus loin. Parallèlement, le tableau 3 montre une série d'ordonnées correspondantes des courbes de polarisation.

Le fait qu'une des mesures d'inégalité valides (la plus sensible dans la partie inférieure, bien qu'elle ne fasse pas partie des trois mesures sélectionnées), à savoir la mesure Theil-Bernouilli, évolue différemment des autres mesures valides au tableau 1 révèle correctement la présence des courbes de Lorenz qui se croisent entre 1973 et



1981, tel qu'il est indiqué au tableau 2 pour l'« Ensemble des PEA ». (N'oublions pas que nous avons fait abstraction, pour le moment, de la variabilité d'échantillonnage.) Dans la plupart des cas, cependant, les trois indicateurs montrent correctement si les courbes de Lorenz sous-jacentes se croisent ou non. Par exemple, les trois évoluent inégalement de 1986 à 1988, puis de 1991 à 1992; les courbes de Lorenz se croisent elles aussi pendant ces périodes (compte non tenu de la variabilité d'échantillonnage).

L'harmonisation générale des indicateurs de polarisation (dont tous sauf un sont *invalides* au sens strict) lorsqu'il s'agit de classer la distribution de l'« Ensemble des PEA » de 1973 comme étant plus polarisée que celle de 1981 témoigne correctement de la « dominance de la polarisation » en termes des courbes de polarisation sous-jacentes montrées au tableau 3 (bien que nous y voyions un conflit entre l'indicateur « 40 %-60 % de la population » et le classement sans équivoque des courbes de polarisation selon l'« étalon-or »).

Ainsi, notre exemple du premier type de divergence (entre l'inégalité et la polarisation) s'avère plus apparent que réel. On l'observe lorsqu'il s'agit de comparer les trois mesures d'inégalité valides à notre série d'indicateurs de polarisation; en l'occurrence, toutefois, les trois mesures scalaires sommaires valides de l'inégalité n'ont pas réussi à déceler le croisement des courbes de Lorenz sous-jacentes.

Le troisième type de divergence a trait à la validité conceptuelle des mesures individuelles. S'agissant de la variance des logarithmes, il existe au moins deux cas de divergences apparentes au tableau 1. Pour l'Ensemble des PEA, la variance des logarithmes varie en direction opposée à toutes les mesures d'inégalité conformes au concept de Lorenz au cours de la période allant de 1991 à 1994. Dans ce cas, le tableau 2 montre un faible croisement des courbes de Lorenz : la courbe de 1991 est un peu supérieure autour du 15<sup>e</sup> centile, alors que celle de 1994 est partout ailleurs au même niveau ou à un niveau supérieur. Une divergence apparente analogue pour la variance des logarithmes touche la population des Hommes à plein temps pour la période allant de 1986 à 1991; encore une fois, cependant, les courbes de Lorenz sous-jacentes se croisent à peine dans le quintile inférieur (phénomène non illustré dans le tableau).

Les deux autres indicateurs d'inégalité *invalides* peuvent également donner des résultats trompeurs. Par exemple, le ratio D9 / D1 régresse de 1988 à 1990 pour l'Ensemble des PEA, alors que les mesures d'inégalité valides progressent toutes; de plus, comme le montre le tableau 2, les courbes de Lorenz ne se croisent pas, de sorte que le classement de l'inégalité est sans équivoque. Parallèlement, la mesure SJDavis diminue pour l'Ensemble des PEA de 1990 à 1991, même si toutes les mesures d'inégalité valides augmentent et que, encore une fois, les courbes de Lorenz ne se croisent pas. Ainsi, ces deux exemples démontrent qu'en réalité un ensemble de mesures d'inégalité scalaires valides ainsi que l'étalon-or de la dominance de Lorenz sont *invalides*.

Un quatrième type de divergence est lié au choix de la population. Comme il a été signalé plus haut, des chercheurs comme Davis (1992) se sont intéressés à des

populations de travailleurs à plein temps de sexe masculin. En principe, toutefois, le fait de se limiter à un tel sous-ensemble de la population active peut amener à négliger les effets de tendances actuelles comme le taux d'activité croissant des femmes, l'augmentation du travail à temps partiel et les changements touchant le travail autonome. En règle générale, l'inégalité parmi les Hommes à plein temps est moins prononcée et plus stable dans le temps que parmi l'Ensemble des PEA (tableau 1). Le cas le plus évident de tendances divergentes selon la population choisie survient entre 1967 et 1973. L'inégalité varie considérablement dans des directions opposées pour les deux populations. En outre, il s'agit de classements sans équivoque, car les deux paires de courbes de Lorenz sous-jacentes ne se croisent pas. En outre, la polarisation est stable ou en régression au cours de cette période pour les Hommes à plein temps, alors qu'elle est nettement en hausse pour l'Ensemble des PEA.

Le dernier type de divergence survient quand les auteurs décèlent des tendances dans les données, alors qu'il n'en existe aucune parce que les variations ne sont pas significatives sur le plan statistique. Le tableau 1 indique des valeurs de trois chiffres pour la plupart des statistiques, mais nous avons vu plus haut qu'en raison de l'erreur d'échantillonnage, deux chiffres au maximum sont statistiquement significatifs, et souvent un seul, dans le cas de mesures sensibles dans la partie supérieure comme le carré du CV.

Plus précisément, le tableau 4, tiré de Kovacevic et Binder (1997, et communication personnelle) donne des estimations cohérentes du coefficient de variation (CV) pour les données de 1991 sur l'Ensemble des PEA des tableaux 1 à 3, en tenant compte de la complexité inhérente du plan de sondage<sup>11</sup>. Le tableau montre aussi l'intervalle de confiance de 95 % (I.C. 95 %, en supposant la normalité), de même que les effets du plan liés à la complexité du plan de sondage (soit les ratios des variances estimées correctement aux variances qui auraient été estimées si l'enquête avait fait appel à un échantillonnage aléatoire simple). Au moyen de ces estimations de la variabilité d'échantillonnage, les colonnes de symboles entre les paires adjacentes d'ordonnées de la courbe de Lorenz et de la courbe de polarisation dans les tableaux 2 et 3 respectivement indiquent alors si oui (+) ou non (-) la courbe de Lorenz ou de polarisation à gauche est supérieure à celle qui se trouve à droite à un centile de population donné, ou bien si la différence n'est pas significative (=), d'après l'hypothèse simplificatrice voulant que l'erreur-type de chacune des années de l'enquête soit la même que celle de 1991. (Cette hypothèse sous-estime vraisemblablement les erreurs-types des années antérieures où la taille de l'échantillon était plus faible.)

Les effets du plan sont appréciables au tableau 4. Même dans l'analyse menée avec soin par Bishop et coll. (1993), qui a tenté de tenir compte des complexités du plan de sondage au moyen de la méthode de Beach et Kaliski (1986), les effets du plan de sondage correspondent à environ le tiers de ceux calculés ici (Kovacevic, communication personnelle reproduisant la démarche Beach-Kaliski pour ces données). Dans un exemple connexe, Kovacevic et Binder (1997) calculent un CV pour le revenu du travail médian de 1991 d'à peu près 1 % dans notre échantillon de plus de 50 000 personnes, alors que Bishop et coll. (1993, tableau A4) présentent un CV de 0,6 %, dans un échantillon d'environ 15 000 personnes, pour le revenu familial médian par habitant de 1981. Si les auteurs ont tellement sous-estimé la variabilité d'échantillonnage pour le Canada, c'est qu'ils ont utilisé une version grand public du fichier de micro-données qui, dans le but de protéger la confidentialité des données fournies par les répondants, ne donne aucune information sur le plan de sondage, notamment sur les grappes.

Si ces sous-estimations énormes de la variabilité d'échantillonnage valent aussi pour les autres ensembles de données des pays visés par l'étude LIS et analysés dans Bishop et coll., un bon nombre des rapports de dominance « statistiquement significatifs » établis par ces auteurs ne le seront plus. On pourrait en dire autant pour un grand nombre d'autres études sur l'inégalité.

Si nous revenons aux données analysées dans la présente communication, ces erreurs-types laissent croire que pour l'Ensemble des PEA, les seuls changements par paires qui sont significatifs sur le plan statistique au tableau 1 sont les augmentations au cours des périodes allant de 1967 à 1973 et de 1981 à 1986 pour les indicateurs Gini et Exp, et entre 1967 et 1973 pour l'indicateur P. À leur tour, ces augmentations appréciables de l'inégalité de 1967 à 1973 et de 1981 à 1986 sont confirmées par l'examen, au tableau 2, des courbes de Lorenz sous-jacentes qui, pour l'essentiel, ne se croisent pas.

Toutefois, la conclusion tirée du tableau 1 selon laquelle il n'y a eu aucune tendance significative de l'inégalité entre 1973 et 1981 ne convient pas, parce que les courbes de Lorenz sous-jacentes se croisent. Les coefficients Gini et Exp sont pratiquement constants, mais cette constance masque des variations compensatoires et statistiquement significatives dans les distributions sous-jacentes au cours de cette période. De plus, l'indice sommaire de polarisation est trompeur à un autre niveau. L'indicateur P n'affiche aucune variation significative de la polarisation entre 1973 et 1981, ni entre 1986 et 1988, au niveau de 95 %, alors que le tableau 3 montre des changements sans équivoque et significatifs sur le plan statistique, notamment dans la moitié inférieure de la distribution.

Il ressort clairement de cette méthodologie que les seuls indicateurs fiables des tendances de l'inégalité ou de la polarisation sont les courbes sous-jacentes complètes de Lorenz et de polarisation.

Si nous examinons attentivement les courbes de Lorenz au tableau 2, nous observons, dans l'ensemble, l'inégalité croissante – et statistiquement significative – pour la moitié

de la série de comparaisons par paires : un changement ambigu de 1973 à 1981, des niveaux à peu près constants au cours des périodes allant de 1986 à 1988, de 1988 à 1990 et de 1992 à 1993, et une diminution de 1993 à 1994. Dans l'ensemble, l'inégalité s'est accrue de 1967 à 1986, mais elle est restée inchangée sur le plan statistique de 1986 à 1994.

Du côté de la polarisation, les conclusions sont à peu près semblables. La polarisation n'a cessé de s'intensifier de 1967 à 1986, exception faite d'un faible croisement des courbes de polarisation de 1981 à 1986. Toutefois, de 1986 à 1994, la polarisation est restée stable dans l'ensemble. Plus particulièrement, la comparaison des ordonnées des courbes de polarisation pour les années (non adjacentes) 1986 et 1994 ne fait ressortir aucune différence significative du point de vue statistique (exception faite des 45<sup>e</sup> et 55<sup>e</sup> centiles).

Ensemble, les tableaux 2 et 3 montrent qu'aussi bien les courbes de Lorenz que les courbes de polarisation peuvent se croiser en pratique, et que de tels croisements se produisent bel et bien, de sorte qu'un certain nombre de classements sont ambigus. Ces tableaux montrent aussi que les deux types de courbes ne se croisent *pas* de la même façon ni aux mêmes endroits, malgré leur étroite relation.

Enfin, cette expérience d'analyse des tendances observées des mesures sommaires d'inégalité et de polarisation s'est révélée quelque peu frustrante. Souvent, les variations des mesures sommaires ne sont pas statistiquement significatives. Il arrive parfois, cependant, qu'une stabilité apparente révélée par un ensemble de trois mesures d'inégalité, spécialement choisies pour capter les variations d'inégalité dans l'échelle globale des revenus, masquent d'importantes variations dans la distribution sous-jacente des revenus du travail, même si ces variations se traduisent par des courbes de Lorenz qui se croisent.

La principale raison qui a justifié, au départ, l'emploi de mesures sommaires, c'est leur commodité, en comparaison de la tâche fastidieuse qui consiste à comparer des courbes de Lorenz complètes – bien qu'une telle comparaison soit le critère par excellence (l'« étalon-or ») pour juger de l'évolution de l'inégalité. Toutefois, compte tenu de la puissance des ordinateurs actuels et de la vaste disponibilité de micro-données, la lourdeur de la tâche n'est plus un argument de poids. Le tableur utilisé pour générer la totalité des tableaux et figures ci-dessus n'est pas très complexe et se trouve bien en deçà de la capacité de la plupart des PC utilisés aujourd'hui. (On peut en obtenir un exemplaire en s'adressant à [Wolfson@statcan.ca](mailto:Wolfson@statcan.ca)).

## **Conclusion**

Certains indicateurs statistiques largement utilisés concernant l'inégalité ou la « disparition de la classe moyenne » sont peut-être trompeurs. Premièrement, les concepts fondamentaux d'inégalité et de polarisation sont distincts et ne produisent pas toujours les mêmes classements des distributions. Deuxièmement, certaines mesures comme la variance des logarithmes et le ratio des D9 / D1 ne mesurent pas ce que la

majorité des gens croient qu'elles mesurent – elles devraient être exclues de l'analyse de l'*inégalité*. Par ailleurs, même des mesures valides peuvent donner des résultats trompeurs comparativement à l'« étalon-or » de la dominance de Lorenz. Outre ces problèmes fondamentaux liés à la clarté des concepts et à la validité conceptuelle, certaines affirmations faites à propos de tendances de l'inégalité sont peut-être inexactes parce qu'elles omettent de tenir compte de la variabilité d'échantillonnage, ou encore elles devraient faire l'objet des réserves qui s'imposent quand l'analyse porte uniquement sur une sous-population, par exemple les travailleurs à plein temps de sexe masculin.

Pour tous ces cas de divergence possible entre les données citées et les conclusions tirées, nous avons donné des exemples montrant l'importance des problèmes. Les conséquences à tirer peuvent être résumées en quelques recommandations touchant les analyses rigoureuses de la distribution des revenus : (1) surtout lorsqu'on aborde des enjeux tels que la « disparition de la classe moyenne », ajouter l'indicateur P ou d'autres mesures de polarisation connexes à l'ensemble des mesures statistiques servant à l'analyse; (2) bannir la variance des logarithmes et les ratios de revenus, par exemple entre le 90<sup>e</sup> et le 10<sup>e</sup> centile, de toute analyse de l'inégalité; (3) sauf si une analyse explicite laisse prévoir une plus grande précision, ne considérer que les deux premiers chiffres de toute statistique d'inégalité (et le premier chiffre seulement dans le cas des mesures sensibles dans la partie supérieure, comme le coefficient de variation) comme étant statistiquement significatifs ou supposer que l'effet du plan de sondage est de trois, à titre de règle approximative, lorsque les données sur les distributions des revenus proviennent d'un échantillon aux grappes complexes; (4) s'efforcer d'utiliser des populations complètes et uniformes à des fins de comparaison, ou au moins présenter ces résultats en référence quand l'analyse porte sur des sous-populations; (5) essayer toujours d'examiner les courbes de Lorenz et de polarisation sous-jacentes comme s'il s'agissait de l'« étalon-or » pour des classements sans équivoque; et (6) tirer parti de la puissance des ordinateurs pour produire des ensembles plus complets d'indicateurs statistiques et de nouvelles méthodes de visualisation des tendances au moyen de tableaux ou de diagrammes.

Tableau 2		Ordonnées des courbes de Lorenz (%) à certains centiles de population, Ensemble des PEA, Canada, 1967 - 1994									
Centile de population	e.-t. (%)	Année									
		1967	1973	1981	1986	1988	1990	1991	1992	1993	1994
5	0,005	0,5 =	0,5 =	0,5 +	0,4 =	0,4 =	0,4 =	0,4 =	0,4 =	0,4 =	0,4 =
10	0,016	1,3 =	1,3 +	1,2 =	1,2 =	1,2 =	1,2 =	1,2 +	1,1 =	1,1 -	1,2
15	0,028	2,6 +	2,5 +	2,4 +	2,3 =	2,3 =	2,3 =	2,3 +	2,1 -	2,2 =	2,2
20	0,043	4,4 +	4,1 +	4,0 +	3,7 -	3,8 =	3,8 +	3,7 +	3,5 =	3,5 -	3,7
25	0,059	6,7 +	6,2 +	6,0 +	5,5 =	5,6 =	5,6 =	5,5 +	5,3 =	5,3 -	5,5
30	0,078	9,4 +	8,7 +	8,5 +	7,8 =	7,9 =	7,9 +	7,7 +	7,5 =	7,5 -	7,7
35	0,096	12,5 +	11,7 +	11,4 +	10,5 =	10,6 =	10,6 +	10,4 +	10,1 =	10,1 -	10,4
40	0,115	16,1 +	15,0 +	14,7 +	13,6 =	13,8 =	13,8 +	13,4 =	13,2 =	13,2 -	13,5
45	0,135	20,0 +	18,8 +	18,5 +	17,2 =	17,4 =	17,4 +	17,0 +	16,7 =	16,7 -	17,1
50	0,153	24,3 +	22,9 =	22,6 +	21,3 =	21,5 =	21,4 +	21,0 =	20,8 =	20,7 -	21,1
55	0,173	29,0 +	27,5 =	27,3 +	25,8 =	25,9 =	25,9 +	25,5 =	25,3 =	25,2 -	25,7
60	0,189	34,1 +	32,5 =	32,3 +	30,8 =	30,9 =	30,8 =	30,5 =	30,2 =	30,1 -	30,7
65	0,208	39,5 +	38,0 =	37,9 +	36,3 =	36,3 =	36,3 =	35,9 =	35,7 =	35,6 -	36,2
70	0,222	45,4 +	43,9 =	44,0 +	42,4 =	42,4 =	42,2 =	41,8 =	41,7 =	41,6 -	42,2
75	0,239	51,7 +	50,4 =	50,5 +	49,1 =	49,0 =	48,7 =	48,4 =	48,4 =	48,2 -	48,8
80	0,256	58,5 +	57,4 =	57,8 +	56,4 =	56,2 =	55,9 =	55,6 =	55,7 =	55,6 =	56,0
85	0,269	65,9 +	65,1 -	65,7 +	64,5 =	64,3 =	63,9 =	63,7 =	63,8 =	63,7 =	64,1
90	0,281	74,1 =	73,7 -	74,6 +	73,6 =	73,3 =	72,9 =	72,8 =	72,9 =	72,9 =	73,2
95	0,286	83,8 =	83,8 -	84,9 +	84,1 =	83,8 =	83,5 =	83,4 =	83,5 =	83,5 =	83,7
97	0,280	88,5 =	88,5 -	89,7 +	88,9 =	88,7 =	88,5 =	88,4 =	88,5 =	88,5 =	88,6
99	0,242	94,4 =	94,5 -	95,4 +	94,9 =	94,7 =	94,7 =	94,5 =	94,6 =	94,4 =	94,6

Tableau 3		Ordonnées des courbes de polarisation (%) à certains centiles de la population, Ensemble des PEA, Canada, 1967 - 1994									
Centile de population	e.-t. (%)	Année									
		1967	1973	1981	1986	1988	1990	1991	1992	1993	1994
5	0,142	18,5	- 19,3	- 20,0	- 20,7	+ 20,3	= 20,4	- 20,8	- 21,3	+ 20,9	= 21,0
10	0,134	14,4	- 15,2	- 15,8	- 16,6	+ 16,3	= 16,3	- 16,8	- 17,1	+ 16,8	= 16,9
15	0,122	10,9	- 11,6	- 12,1	- 12,9	+ 12,6	= 12,6	- 13,1	= 13,2	= 13,1	= 13,1
20	0,106	7,9	- 8,4	- 8,9	- 9,6	+ 9,3	= 9,4	- 9,7	= 9,9	+ 9,6	- 9,8
25	0,088	5,4	- 5,8	- 6,2	- 6,6	+ 6,4	= 6,5	- 6,8	= 7,0	+ 6,7	- 6,9
30	0,067	3,4	- 3,7	- 4,0	- 4,3	+ 4,1	= 4,2	- 4,4	= 4,5	+ 4,3	- 4,5
35	0,047	1,9	- 2,1	- 2,3	- 2,5	+ 2,3	= 2,3	- 2,6	= 2,5	+ 2,4	- 2,6
40	0,029	0,9	= 0,9	- 1,1	= 1,1	= 1,0	- 1,1	= 1,1	- 1,2	= 1,1	- 1,2
45	0,011	0,2	- 0,3	- 0,4	+ 0,2	+ 0,2	- 0,3	= 0,3	+ 0,2	- 0,3	- 0,4
50	0	0,0	= 0,0	= 0,0	= 0,0	= 0,0	= 0,0	= 0,0	= 0,0	= 0,0	= 0,0
55	0,023	0,2	- 0,3	= 0,3	+ 0,2	+ 0,1	- 0,3	= 0,3	= 0,2	- 0,3	= 0,3
60	0,046	0,9	- 1,0	= 1,0	= 1,0	= 1,0	= 1,0	- 1,1	+ 0,9	- 1,2	= 1,1
65	0,093	1,9	- 2,3	= 2,3	= 2,4	= 2,3	= 2,5	= 2,5	= 2,3	- 2,7	= 2,5
70	0,129	3,5	- 4,1	= 4,2	- 4,5	= 4,4	= 4,4	= 4,4	= 4,3	- 4,8	+ 4,4
75	0,188	5,5	- 6,6	= 6,6	- 7,3	= 7,1	= 7,0	= 7,2	= 7,1	- 7,6	+ 7,1
80	0,259	8,0	- 9,6	= 9,9	- 10,8	= 10,5	= 10,5	= 10,6	= 10,6	- 11,4	+ 10,4
85	0,334	11,2	- 13,4	= 13,8	- 15,2	= 15,0	= 14,9	= 15,1	= 15,1	- 16,0	+ 14,8
90	0,420	15,4	- 18,3	= 18,9	- 20,8	= 20,5	= 20,4	= 20,8	= 20,6	- 21,9	+ 20,4
95	0,538	21,1	- 24,9	= 25,6	- 28,0	= 27,8	= 27,8	= 28,2	= 28,0	- 29,5	+ 27,5
97	0,595	24,4	- 28,3	= 29,0	- 31,5	= 31,5	= 31,7	= 32,1	= 31,8	- 33,4	+ 31,2
99	0,683	28,9	- 33,2	= 33,5	- 36,5	= 36,5	= 37,0	= 37,3	= 36,9	- 38,4	+ 36,1

Tableau 4 – Variabilité d'échantillonnage pour certaines mesures

Mesure	Valeur	CV %	I.C. 95 %	Effet du plan
CV <sup>2</sup>	0,761	5,64	0,0858	1,87
Gini	0,412	0,66	0,0054	3,54
Exp	0,460	0,26	0,0024	3,53
P	0,399	1,07	0,0085	2,18
Lorenz (0,25)	0,055	1,06	0,0012	2,96
Lorenz (0,50)	0,210	0,72	0,0030	2,96
Lorenz (0,75)	0,487	0,49	0,0048	3,47
Lorenz (0,90)	0,728	0,38	0,0055	3,84
Polarisation (0,20)	0,097	1,09	0,0021	3,03
Polarisation (0,40)	0,011	2,63	0,0006	2,90
Polarisation (0,60)	0,011	4,08	0,0009	2,82
Polarisation (0,80)	0,115	2,25	0,0052	2,82
Polarisation (0,90)	0,208	2,02	0,0084	3,00

## **Bibliographie**

- Amiel, Y. and F. Cowell (1989), "Measurement of Inequality: Experimental Test by Questionnaire", Discussion Paper No. TIDI/140, October, Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics, London.
- Atkinson, A.B. (1970), "On the Measurement of Inequality", Journal of Economic Theory, Vol.-2.
- Atkinson, A. B. L. Rainwater, and T. M. Smeeding (1995), Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS), OECD Social Policy Studies No. 18, Paris.
- Beach, C.M. (1988), "The 'Vanishing' Middle Class?: Evidence and Explanations", Queen's Papers in Industrial Relations, Industrial Relations Centre, Queen's University at Kingston.
- Beach, C.M. and S.F.Kaliski (1986), "Lorenz Curve Inference with Sample Weights: an Application to the Distribution of Unemployment Experience", Applied Statistics, 35, 439-50.
- Bishop, J.A., J.P.Formby, and W.J.Smith (1993), "International Comparisons of Welfare and Poverty: Dominance Orderings for Ten Countries", Canadian Journal of Economics, XXVI, No. 3.
- Blackburn, M. and D.E. Bloom (1985), "What is Happening to the Middle Class?", American Demographics, January.
- Burkhauser, R.V., A.D.Crews, M.C.Daly, and S.P.Jenkins (1996), "Where in the world is the Middle Class? A Cross-National Comparison of the Vanishing Middle Class Using Kernel Density Estimates", ESRC Working Paper 96-8, University of Essex.
- Cowell, F.A. (1977), Measuring Inequality, Oxford, Philip Allan Publishers.
- Davis, S.J. (1992), "Cross-Country Patterns of Change in Relative Wages", in O.J. Blanchard and S. Fischer (Eds), NBER Macroeconomics Annual, MIT Press.
- Foster, J. and M.C. Wolfson (1992), "Polarization and the Decline of the Middle Class: Canada and the U.S.", Vanderbilt University and Statistics Canada, mimeo.
- Juhn, C., K.M. Murphy and B. Pierce (1993), "Wage Inequality and the Rise in Returns to Skill", Journal of Political Economy, Vol. 101, No. 3



- Karoly, L. (1992), "Changes in the Distribution of Individual Earnings in the United States: 1967-1986", Review of Economics and Statistics.
- Katz, L. and K. Murphy, (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", Quarterly Journal of Economics, Vol. CVII, February, Issue 1.
- Kovacevic, M.S. and D.A. Binder (1997), "Variance Estimation for Measures of Income Inequality and Polarization – The Estimating Equations Approach", forthcoming, Journal of Official Statistics.
- Kuttner, B. (1983), "The Declining Middle", The Atlantic Monthly, July.
- Levy, F. (1987), "Changes in the Distribution of American Family Incomes, 1947 to 1984", Science, Vol.-236, pp.-923-27, May.
- Levy, F. and R.J. Murnane (1992), "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations", Journal of Economic Literature, September, Vol XXX, No 3.
- Love, R. and M.C. Wolfson (1976), Income Inequality: Statistical Methodology and Canadian Illustrations, Catalogue 13-559 Occasional, Statistics Canada, Ottawa, March.
- OECD (1993), Employment Outlook, Paris.
- Pen, J. (1973), "A Parade of Dwarfs (and a Few Giants)", in A.B. Atkinson (Ed.), Wealth, Income and Inequality, Penguin, Middlesex.
- Rowe, G. (1994), "Income Statistics from Survey Data: Effects of Respondent Rounding", mimeo, Social and Economic Studies Division, Statistics Canada, Ottawa, January.
- Shorrocks, A.F. (1980), "The Class of Additively Decomposable Inequality Measures", Econometrica, Vol. 48, No. 3
- Thurow, L. (1984), "The Disappearance of the Middle Class", New York Times, February 5, sect 3, p. 2.
- Wolfson, M.C. (1986), "Stasis Amid Change -- Income Inequality in Canada 1965-1983", Review of Income and Wealth, December.
- Wolfson, M.C. (1989), "Inequality and Polarization: Is There a Disappearing Middle Class in Canada?", in Proceedings of the Statistics Canada Symposium on Analysis of Data in Time, October 1989, Statistics Canada, Ottawa.

Wolfson, M.C. and J. Foster (1993), "Inequality and Polarization -- Concepts and Recent Trends", paper presented to the Winter Meetings of the American Statistical Association, Ft. Lauderdale, January.

Wolfson (1994), "When Inequalities Diverge", American Economic Review, May.

## Notes:

<sup>1</sup> Kolm (1966) était lui aussi conscient de cette distinction, encore qu'il ait employé une terminologie différente. Il parlait de «rectifiante» pour désigner la condition de transfert de Pigou-Dalton. Il faisait ensuite remarquer qu'une succession de changements «rectifiants» apportés à une distribution de revenus déterminée finissait par aboutir à l'égalité complète (c'est-à-dire à une courbe de Lorenz coïncidant avec la droite à 45 degrés), mais qu'il n'en découlait pas nécessairement une diminution monotone des inégalités (ce qu'il appelait «isophilie» ou amour de l'égalité). Plus précisément, une succession de transferts pouvait entraîner un élargissement des écarts de revenus entre certaines paires de particuliers — ce qui correspond au fond à notre notion de dispersion.

<sup>2</sup> Notons que 0,25 n'est pas nécessairement le maximum. P pourrait dépasser 0,25 si la moitié de la population avait un revenu moyen négatif. Pour la même raison, le coefficient de Gini peut dépasser 1,0 lorsqu'il y a des revenus négatifs. Pour une part à la médiane et une valeur  $m_{tan}$  données (toutes deux positives), P est réduit au minimum et s'approche de zéro pour une distribution trimodale dans laquelle une personne a un revenu négatif très élevé, une autre a un revenu positif très élevé et toutes les autres ont le même revenu, entre les deux premiers.

<sup>3</sup> Le déplacement est parallèle puisque la valeur de  $m_{tan}$  n'est pas touchée par le transfert. Il convient de noter qu'en fin de compte, un tel déplacement parallèle vers le haut provoque le croisement de la tangente à la médiane et de la droite à 45 degrés (soit au 0<sup>e</sup> centile si  $m/\mu < 1$ , sinon au 100<sup>e</sup> centile), ce qui donne à entendre une courbe de Lorenz sur au moins la moitié de sa longueur. Cela correspond à une série de transferts croisant la médiane et préservant la moyenne et la médiane, en vertu desquels la moitié de la distribution devient concentrée dans un seul revenu. Ainsi, on peut associer une réduction de P à un accroissement de bimodalité. Toutefois, étant donné que P est fondé sur l'ampleur à la fois de l'étalement et de la bimodalité, la réduction de l'étalement dans ce cas neutralise – et au-delà – l'augmentation de la bimodalité.

<sup>4</sup> La popularité soutenue de la variance des logarithmes pourrait s'expliquer par le fait qu'elle permet une décomposition commode. Toutefois, les indicateurs Theil-entropie et Theil-Bernouilli, ainsi que le carré du coefficient de variation (défini plus loin), offrent aussi des décompositions "intéressantes" et ne souffrent pas de la même lacune sur le plan de la validité conceptuelle.

<sup>5</sup> Une formulation légèrement différente de la sensibilité à divers points de l'échelle des revenus, avec des conclusions semblables quant au choix des mesures d'inégalité, a été présentée dans Love et Wolfson (1976).

<sup>6</sup>  $Exp = \sum p_i \exp(-y_i / \mu)$  où  $p_i$  est la proportion de la population dans le  $i^e$  groupe de revenus,  $y_i$  est le revenu moyen de ce groupe et  $\mu$  est le revenu moyen global.

---

<sup>7</sup> Par exemple, aucun compte n'est tenu ici de la constatation récente de Rowe (1994) selon laquelle les répondants indiquent souvent leurs revenus en montants arrondis (p. ex. au multiple de 100 \$ ou de 1 000 \$ le plus proche). Rowe estime que pour des données sur les revenus semblables à celles utilisées ici, ces valeurs arrondies entraînent des erreurs de réponse qui sont de même grandeur que les erreurs d'échantillonnage estimées par Kovacevic et Binder (1997), et qui doivent être ajoutées à ces dernières.

<sup>8</sup> Un autre problème peut résulter de l'interaction entre des mesures d'inégalité sensibles dans la partie inférieure et le choix de la population. Si toutes les personnes ayant un revenu strictement positif sont incluses, comme dans Karoly (1992), plutôt que celles gagnant au-delà d'un seuil minimum plus élevé, p. ex. 5 % du salaire moyen, des mesures comme les indicateurs Theil-entropie et Theil-Bernouilli pourraient présenter des variations trompeuses, en raison des fluctuations touchant les sous-populations n'ayant que quelques dollars de revenus. Ce problème a été observé avec des données canadiennes.

<sup>9</sup> Chaque distribution a d'abord été totalisée en vue d'obtenir les moyennes des 100 centiles. Toutes les statistiques présentées ont ensuite été calculées à partir de ces données intermédiaires. Il en résulte une certaine erreur d'approximation (de groupement); cependant, exception faite du carré du CV, il s'agit d'une erreur négligeable lorsqu'on la compare aux résultats tirés directement des micro-données brutes. Ces chiffres diffèrent aussi quelque peu de ceux présentés dans Wolfson (1994) parce qu'on a apporté un changement dans le traitement du revenu négatif découlant d'un emploi autonome, changement qui ne touche que la première moyenne de centiles.

<sup>10</sup> Theil-entropie =  $\sum (y_i / \mu) \ln (y_i / \mu)$ ; Theil-Bernouilli =  $-\sum p_i \ln (y_i / \mu)$ , où  $p_i$  est la proportion de la population dans le  $i^e$  groupe de revenus,  $y_i$  est le revenu moyen de ce groupe et  $\mu$  est le revenu moyen global.

<sup>11</sup> L'échantillon sous-jacent de 1991 renfermait 50 701 observations dans 4 201 grappes ou unités primaires d'échantillonnage, tirées à leur tour de 1 139 strates. Il convient de noter que ces statistiques ont été calculées à partir des micro-données complètes, de sorte qu'en comparant les valeurs présentées ici à celles des tableaux 1 à 3, on obtient une idée de l'erreur d'approximation qui résulte de l'utilisation des moyennes des 100 centiles pour calculer toutes les autres statistiques présentées.