Wealth inequality: Theory, measurement and decomposition

James B. Davies *Department of Economics, University of Western Ontario*
Nicole M. Fortin *Vancouver School of Economics, University of British Columbia*
Thomas Lemieux *Vancouver School of Economics, University of British Columbia*

Abstract. This paper reviews the basic principles of inequality measurement, underlining the advantages and shortcomings of alternative measures from a theoretical standpoint and in the context of the study of the distribution of wealth. Adopting the two most popular measures, the Gini index and the P-shares, the paper documents wealth inequality in Canada using the 1999, 2005 and 2012 Survey of Financial Security (SFS). It carries out several decompositions with covariates, featuring DFL-type reweighting methods and Gini and P-shares RIF regressions. The latter parallel decompositions deepen our understanding of how changes in socio-demographic characteristics, including the compensating role of family formation and human capital, impact wealth inequality.

Résumé. Cet article débute par une revue des principes fondamentaux de la mesure de l'inégalité de la richesse, en soulignant les avantages et inconvénients de diverses mesures d'un point de vue théorique. Adoptant les deux mesures les plus populaires, l'indice de Gini et les fractiles, cet article documente l'évolution de l'inégalité de la richesse au Canada en utilisant les données des enquêtes sur la sécurité financière (ESF) de 1999, 2005 et 2012. Puis il réalise plusieurs types de décompositions, incluant les méthodes basées sur la repondération de type DFL et les RIF régressions pour l'indice de Gini et les fractiles. Ces dernières approfondissent notre compréhension de la façon dont les changements de caractéristiques sociodémographiques ont un impact sur l'inégalité de la richesse, notamment le rôle compensatoire de la formation des ménages et du capital humain.

JEL classification: D31, D63

1. Introduction

Over the last three decades, much empirical analysis has centred on changes in wage and earnings inequality fuelled by technological change, insti-

We would like to thank Pablo Gutiérrez for outstanding research assistance on this project. We also acknowledge SSHRC and the Bank of Canada Fellowship Program for financial support.
Corresponding author: Thomas Lemieux, thomas.lemieux@ubc.ca

Canadian Journal of Economics / Revue canadienne d'économique, Vol. 50, No. 5
December 2017. Printed in Canada / Décembre 2017. Imprimé au Canada

0008–4085 / 17 / 1224–1261 / © Canadian Economics Association

tutional transformation and globalization. In Canada, individual-based labour force surveys (LFSs) and household-based surveys (such as SLID) provided researchers with the information to evaluate the relative importance of these explanatory factors on a more or less continuous basis. A compendium of analyses in Green et al. (2016) summarizes the Canadian story in terms of income inequality.[1]

Since the early 2000s, however, the concerns about increasing inequality have shifted towards top income groups and researchers (Saez and Veall 2005, Veall 2012) have turned to data sources (such as LAD) based on income tax data, which are available yearly. The sustained increases in top incomes since the 1990s raise the question of why these were not accompanied by similarly strong increases in top wealth shares, as measured in standard household surveys of assets and debts, given that those richer individuals consume a relatively smaller share of their income. A possible answer to this question, for the United States, was provided by Saez and Zucman (2016). They extended the use of US income tax data to the analysis of wealth inequality by capitalizing the incomes reported by individual taxpayers. Their results indicate that the upsurge of top incomes combined with an increase in saving rate inequality led to larger increases in top wealth shares than shown by the Federal Reserve Board's Survey of Consumer Finance (SCF), with the top 1% share reaching 42% in 2013. However, these results are subject to the general limitations of the income capitalization method (see, e.g., Atkinson and Harrison 1978) and are also constrained by their reliance on income tax records (Kopczuk 2015, Bricker et al. 2016). Bricker et al. carefully examine the reasons for the difference in results between Saez and Zucman (2016) and those based on the SCF. They also provide their own estimates based on multiple data sources and refinements, concluding that the share of the top 1% in 2013 was just 33%. Like the SCF, their results do show an upward trend in wealth inequality over the last three decades, albeit one less pronounced than that found by Saez and Zucman (2016).

The Canadian case is different from that of the United States. While the US evidence shows an upward trend in wealth inequality, Canada's Survey of Financial Security (SFS) does not show such a trend as we find here. Further, adjusting the SFS top wealth shares to make them consistent with the "rich lists" published by *Forbes* magazine and other media outlets does not disturb the lack of trend (Davies and Di Matteo 2017). Beginning in 2006 and continuing after the global financial crisis (GFC) of 2007–2008, there was a collapse of house prices in the United States, while there has been a more or less continuous rise of house prices in Canada. Since housing is relatively less important for the wealthy than for the middle class, this contrast helps to explain the lack of a rise in top wealth shares in Canada over this period. Weaker performance of the Canadian stock market since the GFC vs. its US counterpart further helps to explain the lack of trend in Canada, as stocks are most important for the wealthy. However, the contrast

1 For example, Fortin and Lemieux (2015) find that the oil boom of the mid-2000s had an important mitigating effect on increasing wage inequality.

between the lack of an upward trend in Canada vs. rising top wealth shares in the United States holds over the last three decades, not just the last five to 10 years. A full explanation for the lack of trend in Canadian wealth inequality therefore requires a look at more than just the behaviour of asset prices.

Here we document wealth inequality in Canada as reflected in the 1999, 2005 and 2012 Surveys of Financial Security. We carry out several decompositions with covariates. The methods applied include the widely used DFL reweighting method (DiNardo et al. 1996), the RIF Gini regressions proposed by Firpo et al. (2009) and new RIF P-shares regressions. The latter parallel decompositions deepen our understanding of how changes in socio-demographic characteristics affect wealth inequality. We find that an important reason for the lack of trend in wealth inequality in Canada over the last decades is that changes in family formation and in human capital investment have had approximately offsetting impacts.

The study of the measurement of wealth inequality has a long tradition in the Canadian literature (Podoluk 1974; Davies 1979, 1993; Harrison 1980; Oja 1983; Osberg and Siddiq 1988; Siddiq and Beach 1995; Di Matteo 1997, 2016; Morissette et al. 2006; Morissette and Zhang 2007; Brzozowski et al. 2010). We are thus able to draw on the contributions of many authors who have discussed the properties, advantages and disadvantages of different approaches.

The paper is organized as follows. Section 2 reviews the theoretical foundations of inequality measurement and addresses the classical analysis of the decomposability of the summary measures by income or wealth components and population subgroups. In section 3, we provide a summary of decomposition methods using covariates for the case of changes in wealth inequality. Section 4 features an empirical application of the use of decomposition methods, using the 1999, 2005 and 2012 SFS surveys, as well as the 1984 Survey of Consumer Finances (SCF). Section 5 concludes.

2. Theoretical foundations: Inequality measurement

The theory of inequality measurement is a rich and highly developed area. Our emphasis here is on inequality measures or practices that are useful in the study of wealth inequality.

The discussion in this section is couched in terms of the distribution of wealth in a finite population with n members, which we will refer to as individuals. (This is a help in exposition. In the next section, continuous distributions are used since they are more appropriate in a statistical context.) The theory can of course be applied to variables other than wealth, including income or labour earnings. And the units studied could be families or households. Denoting individual i's wealth as y_i, we order the individuals such that $y_1 \leqslant y_1 \leqslant \cdots \leqslant y_n$ and let $Y = (y_1 y_2, \ldots, y_n)$. Mean wealth is $\bar{y}$.

One of the foundations of the theory of inequality measurement is the Pigou–Dalton principle of transfers (Dalton 1920), or *principle of transfers* for short. It says that if wealth is transferred from a richer person to a poorer person, without reversing their ranks, inequality goes down. And of course, a transfer in the opposite direction makes inequality go up. If there are only two individuals, any redistribution of wealth has an unambiguous effect on inequality. Either it is a transfer from a richer to a poorer, or vice versa. But if there are more than two people, it is easy to construct distributional changes that cannot be judged as unambiguously inequality-reducing or increasing just by applying the principle of transfers. Let $n=3$. Transfer a small amount from individual 2 (the middle person) to individual 1 (the poorest). At the same time make a transfer of an equal amount from individual 2 to individual 3 (the richest). The first transfer is equalizing while the second is disequalizing.

When elements of the same redistributive package have conflicting effects on inequality, different observers will have different opinions about whether overall inequality has risen or fallen. In the above example, suppose the three individuals have wealth levels (\$100,000; \$200,000; \$300,000) and the amounts transferred are \$1,000 in each case. Some observers, perhaps the majority, would feel that equalizing the distribution at the bottom, by transferring \$1,000 from the middle person to the bottom person is a more significant change than the increase in inequality caused by taking \$1,000 from the middle person to give to the top person. We can say that they believe inequality is falling. Other observers will have the opposite opinion. In the theory of inequality measurement, the first group of observers are said to be *transfer sensitive.*

The principle of transfers and transfer sensitivity each contribute to ranking distributions according to their level of inequality. Suppose we have two wealth distributions, Y and Y' with the same mean. If the Lorenz curve for Y, say, is nowhere below that of Y' and it is above that of Y' at least somewhere[2], then distribution Y can be derived from Y' by a series of equalizing transfers. The converse is also true, so Lorenz curves can always tell us whether there is an unambiguous inequality ranking of different distributions that have the same mean.

What do we do if the distributions we want to compare have different means?[3] One approach is to restrict attention to *relative inequality*. If all individuals' wealth levels change by the same percentage, which alters the mean but has no effect on

2 Here "below" and "above" mean strictly below and strictly above, as they do throughout the paper. The Lorenz curve displays the relationship between the proportion of overall wealth accruing to the bottom p% of the population (with wealth below the p-th quantile q_p), $L(p)=\sum_{(y_i<q_p)} y_i / \sum_{i=1}^{N} y_i$, and the corresponding proportion of the population, p.

3 If population sizes differ, inequality comparisons can still be made if we accept the principle of population homogeneity, which says that inequality is not altered if the population is replicated. If one distribution has the population size m and the other has population n, replicating the first population n times and the second m times, generates two populations of the same size, mn, with the same inequality levels as the respective original distributions, which therefore can indeed be compared.

the Lorenz curve, we would say that inequality does not change. This is the usual approach in practice—we compare the Lorenz curves or Gini coefficients for countries, time periods or whatever and discuss the differences in inequality they show, ignoring differences in means. Differences in means are considered separately. That is the approach followed in this paper, but it is not the only possible one. The theory of inequality measurement has also been worked out for absolute inequality, in which case a uniform *absolute* change in all individuals' incomes is taken as having no effect on inequality (Kolm 1976a, 1976b).

What happens when Lorenz curves cross? In that case, the distributions cannot be ranked by the principle of transfers alone. But Shorrocks and Foster (1987) proved a remarkable and helpful result. If the Lorenz curves for two distributions, Y and Y', cross once with the Y Lorenz curve higher than the Y' curve in the lower range and if the coefficient of variation of Y, $CV(Y)$, is strictly less than $CV(Y')$ then all observers who are transfer sensitive will say the Y distribution is more equal than Y'.[4]

These results show that the Lorenz curve is more than a handy tool. It is central to relative inequality measurement. With very high-quality datasets, for example administrative data on a country's entire population, such as one sees in Scandinavia, the results can be readily applied. With data based on household surveys, which have much smaller sample sizes, as in the Statistics Canada SFS surveys used in this article, there is an issue of how confident one can be that one Lorenz curve lies completely above the other, or that they intersect a particular number of times. These issues were addressed by Beach and Davidson (1983) and Beach and Richmond (1985), who developed methods of statistical inference and joint confidence intervals for income shares and Lorenz curves.[5] In Canada, these methods were applied to the study of historical wealth inequality by Siddiq and Beach (1995).

As we show in section 4, while the Lorenz curve has central importance in inequality measurement, simply viewing Lorenz curves does not reveal the detailed information one needs to carefully assess differences in distributions. For this reason, analysts have long looked at decile and quintile shares, supplemented in many cases by top shares such as those of the top 1% and 5%. These are examples of what are referred to as percentile shares (or P-shares hereinafter) in the modern literature. They are the shares of individuals between particular percentiles. Use of P-shares has refocused attention away from the "standard" top shares, e.g., to the shares of individuals between the 95th and 99th percentiles and between the 90th and 95th percentiles. While using a battery of summary inequality indexes is still useful in comparing inequality across a large number of distributions, for example in international datasets, carefully examining the

4 Davies and Hoy (1995) extended this result to the case where Lorenz curves may intersect any number of times.

5 There is a large literature on the statistical aspects of inequality and social welfare comparisons. Cowell and Mehta (1982) was an important early contribution. See Bishop, Chakraborti et al. (1991); Bishop, Formby et al. (1991); Davidson and Duclos (2000); and Cowell and Victoria-Feser (1996, 2008) for examples of later work.

shape of the wealth distribution with the help of P-shares is indispensable for understanding the sometimes subtle differences in a small set of distributions, which is our goal in this article.

P-shares have a special advantage in the context of wealth inequality. This is that interior P-shares are unaffected, in relative terms, by errors in the estimated extremes of the distribution. Suppose, for example, that it is known that a particular wealth survey underestimates the share of the top 1% by at least 10 percentage points but that the remainder of the distribution is well captured. Then the estimated P-shares below the 99th percentile are all too high, but we do not know by how much since the precise degree of error in the top 1% share is not known. The required adjustment in the P-shares below the top 1% would be equi-proportional, however, so their relative differences would be unaffected.

2.1. Classical decomposition

An important question about any inequality measure or index is whether it can be decomposed. In this section we discuss "classical decomposition," which breaks up total inequality into the contributions either of subgroups or wealth components. The classical approach is applied in section 4 to the decomposition of wealth inequality in Canada by family types. The next two sections set out and apply regression-based decomposition methods using covariates.

2.1.1. Decomposition by subgroups

Let a population be composed of m subgroups with population n_j, mean wealth $\bar{y}_j$ and wealth vector Y_j, all for $j=1,\dots,m$. Let $I(\cdot)$ be an inequality index, with overall value $I(Y)$, and $I(Y_j)$ for a subgroup. Then we say that I is *decomposable* if it can be written as:

$$I\left(Y\right)=I\left(n_1,\dots,n_m;\bar{y}_1,\dots,\bar{y}_m;I\left(Y_1\right),\dots,I\left(Y_j\right)\right), \tag{1}$$

with $I(Y)$ strictly increasing in each $I(Y_j)$. It is *additively decomposable* if it can be written:

$$I\left(Y\right)=I_B\left(Y\right)+\textstyle\sum_{j=1}^{m} s_j I(Y_j), \tag{2}$$

where $I_B(Y)$ is between-group inequality (the value of I if there were no inequality within the subgroups) and the weights s_j sum to one. Typically, s_j is either the population share, n_j/n, or the wealth share, $n_j\bar{y}_j/n\bar{y}$.

Some of the most popular inequality measures, for example Atkinson's index and the coefficient of variation (CV), are decomposable and some others, such as Theil's index, are also additively decomposable (Jenkins 1991, Cowell 2011). However, there are at least two popular indexes that are not decomposable: the variance of logarithms, often used in conjunction with earnings or income regressions, and the Gini coefficient.[6]

6 If subgroup wealth distributions are non-overlapping, the Gini coefficient is additively decomposable (see, e.g., Cowell 2011). It is also worth noting that the concept of between-group inequality remains well defined for non-decomposable indexes.

2.1.2. Decomposition by income components

Let wealth have K components, mean wealth of type k be denoted $\bar{y}^k$ and the vector of type k wealth be Y^k, for $k=1,\ldots,K$. Again, I is an inequality index, with overall value $I(Y)$. Inequality of component k is $I(Y^k)$. Decomposing $I(Y)$ in this case means attributing to each component a proportional share, S_k, of $I(Y)$. Shorrocks (1982) showed that even for a single inequality index there are legitimate alternative ways of doing this unless an appropriate symmetry assumption is made. But clear-cut results are obtained if one imposes "two factor symmetry". The latter requires that, if $k=2$, two components should be assigned the same contribution to inequality if (a) the distribution of wealth from both sources is identical and (b) together they make up total wealth. Under this assumption, whatever inequality index is used, S_k is given by the "natural" decomposition of the variance, $V(Y)$, or the square of the CV.

The natural decomposition of $V(Y)$ can be found from:

$$V\left(Y\right)=\sum_k V(Y^k)+\sum_{j\neq k}\sum_k \rho_{jk}[V\left(Y^j\right)V\left(Y^k\right)]^{\frac{1}{2}}. \tag{3}$$

The contribution of component k to the first term is simply $V(Y^k)$, but what is its contribution to the second interaction term? In Shorrocks' analysis, the natural approach is to assign to component k half the value of all the interaction terms involving that factor. If that is done, we get: $S_k=\frac{cov(Y^k,Y)}{V(Y)}$, where $cov(Y^k,Y)$ is the covariance between component k and total wealth. This decomposition can be used with any valid relative inequality index. Thus, there is a unique decomposition of inequality by wealth components—a useful result in applied work.

2.2. Gini coefficient

The most popular inequality index is the Gini coefficient. This may be true partly because it is easy to explain to a general audience, simply because it ranges from 0 to 1. (It is easy to overlook how unusual this property is, but there is no other inequality index in common use that has it.) More sophisticated audiences know that the Gini equals twice the area between the diagonal and the Lorenz curve in the familiar diagram, which is a nice aid to intuition. This is the usual way to explain the Gini coefficient, e.g., in first-year economics textbooks.

Gini (1914) defined his index in terms of the mean difference, that is, the average absolute difference between the incomes of pairs of individuals. This he divided by 2 and normalized by the mean, yielding:

$$G=\frac{1}{2n^2\bar{y}}\sum_{i=1}^n\sum_{j=1}^n\left|y_i-y_j\right|. \tag{4}$$

At bottom, economic inequality is about income differences between people. This formula shows that the Gini coefficient takes into account *every* such difference—a virtue emphasized, e.g., by Sen (1973). It also obeys the principle of transfers. But while the Gini coefficient is attractive in these ways, it also has some less happy properties. One we have seen already is its lack of additive decomposability.

Probably the most important limitation of the Gini coefficient is that its sensitivity to transfers depends only on the amount transferred and the *number of individuals between* the "donor" and "recipient". It does not depend on the difference in wealth between these two individuals, nor does it depend directly on how high their wealth is. The result is that, in practice, the Gini coefficient is most sensitive to transfers in the middle of the distribution. That is because typical distributions of wealth, income, or earnings are unimodal and have high density in the middle. This means, for example, that if \$1,000 is transferred across a wealth gap of \$100,000 in the middle of the distribution, the Gini will change much more than if \$1,000 were transferred across a \$100,000 wealth gap close to the bottom of the distribution or at the very high top, where fewer individuals would be between the donor and recipient.

The fact that the Gini coefficient is more sensitive to transfers in the middle of the distribution than in the bottom end is somewhat troubling. This means that the most popular inequality index is not transfer sensitive, despite the fact that many, if not most, observers apparently feel that inequality is more important lower in the distribution—as revealed by the intense interest in poverty and poverty measurement. The appropriate response is not to stop using the Gini coefficient. When many distributions are being compared it can be supplemented with other measures that are transfer sensitive, for example Atkinson's index or Theil's index (see, e.g., Sen 1973, Jenkins 1991 or Cowell 2011). When a small number of distributions are being compared, as in this paper, there is no need to rely heavily on a summary index, as discussed above.

3. Decomposition methods with covariates

3.1. Decompositions and counterfactuals

Despite the limitations mentioned above, the Gini coefficient remains a widely used measure of inequality. Over the last 15 years, scholars and policy analysts have also increasingly used top income (or wealth) shares as a measure of inequality. This focus was in large part motivated by the tremendous growth in the share of income going to the top 1% in the United States, Canada and many other countries. Top income shares, or other percentile shares, indicate the percentage of income going to different groups depending on their rank in the distribution. P-shares are all simple functions of the Lorenz curve. For instance, the top 10% share of the distribution $F(\cdot)$ is given by $1 - L(F; p_{90})$, where $L(F; p)$ is the Lorenz ordinate evaluated at the p-th percentile.

Like the Gini coefficient, Lorenz ordinates and P-shares are not decomposable measures of inequality, as discussed in section 2. Nor are interquartile differences like the gap between the 90th and 10th percentiles that have been widely used in the labour economics literature on wage inequality. Strictly speaking, this means these inequality measures cannot be explicitly written out as a function of the sample composition, group means and within-group inequality (the terms s_j, $\bar{y}_j$ and $I(Y_j)$ in equations (1) and (2), respectively).

Fortunately, several procedures are now available to carry out informative decompositions using covariates when the inequality measure of interest is not decomposable in the conventional way defined in equations (1) or (2). Going back to the famous Oaxaca–Blinder (OB) decomposition of the mean, the critical issue when performing decompositions is to construct counterfactual values for the measure of interest. For instance, one may be interested in knowing how the mean and other features of the distribution would change if we were to increase the share of university-educated workers by 10% and reduce the share of high-school educated students by the same amount. Since the overall mean can be written as $\bar{y}=\sum_{j=1}^{m} s_j\bar{y}_j$, the counterfactual mean $\bar{y}^C=\sum_{j=1}^{m} s_j^C\bar{y}_j$ is obtained by replacing the shares s_j by counterfactual shares s_j^C. It is easy to show that replacing s_j by s_j^C is equivalent to reweighting each observation by a factor $\hat{\psi}_j=s_j^C/s_j$ when computing the sample mean.[7] The same procedure can then be used to compute the counterfactual value of any other distributional measure of interest (DiNardo et al. 1996). For example, consider the case of the top 10% share, $S(p_{90})=1-L(F;p_{90})$, which is estimated as $\hat{S}(p_{90})=\frac{\sum_{(y_i\geqslant\hat{q}_{90})} y_i}{\sum_{i=1}^{N} y_i}$, where $\hat{q}_{90}$ is the sample estimate of the 90th quantile of the distribution of y. The counterfactual wealth share $\hat{S}^C(p_{90})$ can be computed as:

$$\hat{S}^C(p_{90})=\frac{\sum_{(y_i\geqslant\hat{q}_{90}^C)}\hat{\psi}_j y_i}{\sum_{i=1}^{N}\hat{\psi}_j y_i}. \tag{5}$$

Thus, it is always possible to compute a counterfactual value of an inequality measure where each observation is reweighted by the factor $\hat{\psi}_j=s_j^C/s_j$ regardless of whether the measure is decomposable. Note, however, that counterfactuals obtained by reweighting are partial equilibrium in nature. For example, changing the fraction of individuals with different levels of education may have an impact on their wages and, ultimately, on their wealth level. We abstract from these possible general equilibrium effects in this paper.

In the remainder of the paper, we will work with a set of covariates X that could either capture groups (e.g., if X is a categorical variable indicating the level of completed education), or a more general set of discrete or continuous covariates. In the more general setting, the reweighting factor ψ_j will be replaced with $\psi(X)$.

In a setting with a more general set of covariates, it becomes important to go beyond the simple counterfactual experiment discussed above and consider

7 Since the group means are defined as $\bar{y}_j=(1/N_j)\sum_{i=1}^{N_j} y_{ij}$, substituting this expression into the equation for the counterfactual mean and using the fact that $s_j=N_j/N$ yields:
$\bar{y}^C=\sum_{j=1}^{m} s_j^C(1/N_j)\sum_{i=1}^{N_j} y_{ij}=(1/N)\sum_{j=1}^{m} s_j^C(N/N_j)\sum_{i=1}^{N_j} y_{ij}=(1/N)\sum_{j=1}^{m}\sum_{i=1}^{N_j}(s_j^C/s_j)y_{ij}$.
Technically speaking, $\bar{y}$ denotes a sample average rather than a population mean; we have dispensed with the distinction so far, but we return to a more formal notation below.

alternative counterfactuals where only some elements of X are being manipulated. For instance, we may want to know how much of the increase in the top 1% wealth share can be attributed to observed changes in the distribution of education, holding other factors unchanged. In the case of the mean, this is a well-known problem, typically tackled using an OB decomposition. The decomposition is based on a linear (in the parameters) model, $Y = X\beta + \varepsilon$, where the error ε satisfies the zero conditional mean assumption ($\mathbb{E}(\varepsilon|X) = 0$). Applying the law of iterated expectations, it follows that:

$$\mathbb{E}\left(Y\right) = \mathbb{E}_X\left[\mathbb{E}\left[Y|X\right]\right] = \mathbb{E}\left(X\right)\beta = \beta_0 + \textstyle\sum_{k=1}^{K}\mathbb{E}\left(X_k\right)\beta_k, \tag{6}$$

where K is the number of individual covariates excluding the constant. The sample analog of equation (6) is given by:

$$\bar{Y} = \bar{X}\hat{\beta} = \hat{\beta}_0 + \textstyle\sum_{k=1}^{K}\bar{X}_k\hat{\beta}_k, \tag{7}$$

where $\hat{\beta}$ is the OLS estimate of β. As discussed at the beginning of this section, the OB decomposition is based on a comparison between actual and counterfactual means. Consider an OB decomposition of the change in mean wealth between 1999 and 2012, two of the periods considered in the empirical application in section 4. One interesting counterfactual in this setting is the average wealth that would prevail in 2012 if the distribution of the covariates X had remained as in 1999. Under the linearity and zero conditional mean assumption, the counterfactual average wealth $\bar{Y}^C$ can be computed as:

$$\bar{Y}^C = \bar{X}_{1999}\hat{\beta}_{2012} = \hat{\beta}_{0,2012} + \textstyle\sum_{k=1}^{K}\bar{X}_{k,1999}\hat{\beta}_{k,2012}. \tag{8}$$

Note that the reweighting approach introduced above could also be used to compute the counterfactual mean as: $\bar{Y}^C_{RWT} = (1/N)\sum_{i=i}^{N}\hat{\psi}(X_i)Y_i$.

As discussed below, in this setting the reweighting factor $\hat{\psi}(X_i)$ represents the estimated probability that an observation with covariates X_i is observed in 2012 instead of 1999. But unlike equation (8), this alternative way of computing the counterfactual mean is a potentially complicated function of X_i.[8] One major advantage of the counterfactual based on the linear model is that it yields a linear closed form solution in the mean value of the covariates. When comparing the actual mean income in 2012, $\bar{Y}_{2012}$, to the counterfactual mean, $\bar{Y}^C$, we get:

$$\bar{Y}_{2012} - \bar{Y}^C = \bar{X}_{2012}\hat{\beta}_{2012} - \bar{X}_{1999}\hat{\beta}_{2012} = \textstyle\sum_{k=1}^{K}\left(\bar{X}_{k,2012} - \bar{X}_{k,1999}\right)\hat{\beta}_{k,1999}. \tag{9}$$

Thus, the difference between the actual and the counterfactual mean is a weighted sum of the 1999 to 2012 difference in the mean value of each covariate, using the OLS coefficients as weights. For example, if the k-th covariate is years of education, $(\bar{X}_{k,2012} - \bar{X}_{k,1999})\hat{\beta}_{k,1999}$ indicates the impact on mean wealth of changing education from its 1999 to its 2012 level. In the OB decomposition, these types of counterfactual experiments are used to compute the "explained" or "composition effect" part of the decomposition.

8 Kline (2011) discusses a special case where the reweighting factor is a linear function of the Xs, in which case the two ways of computing the counterfactual are equivalent.

The OB decomposition is obtained by subtracting and adding the counterfactual mean $\bar{Y}^C$ to the difference in means between 1999 and 2012:

$$\begin{aligned}\bar{Y}_{2012} - \bar{Y}_{1999} = \hat{\Delta}^{\mu}_{OB} &\equiv (\bar{Y}_{2012} - \bar{Y}^C) - (\bar{Y}_{1999} - \bar{Y}^C) \\ &= \sum_{k=1}^{K} \left(\bar{X}_{k,2012} - \bar{X}_{k,1999}\right) \hat{\beta}_{k,2012} + \left(\hat{\beta}_{0,2012} - \hat{\beta}_{0,1999}\right) \\ &\quad + \sum_{k=1}^{K} \bar{X}_{k,1999} \left(\hat{\beta}_{k,2012} - \hat{\beta}_{k,1999}\right).\end{aligned} \tag{10}$$

As just discussed, the first term represents the explained or composition effect. Under the above two assumptions, the last two terms reflect changes in the "wealth structure" as summarized by the regression coefficients, or returns to observable characteristics, $\hat{\beta}$. These last two components are often referred to as the unexplained part of the decomposition.

The OB decomposition is very easy to compute as it simply involves estimating OLS regressions and sample means. Because of the linearity assumption, OB provides a "detailed" decomposition of $\bar{Y}_{2012} - \bar{Y}_{1999}$ in the sense that both the composition effect, $\hat{\Delta}^{\mu}_{OB:X} = \bar{Y}_{2012} - \bar{Y}^C$, and the wealth structure effect, $\hat{\Delta}^{\mu}_{OB:S} = \bar{Y}_{1999} - \bar{Y}^C$, can be divided up in the contribution of each covariate. This is arguably the most important advantage of the OB decomposition over other methods, like reweighting, that can be used when one is interested only in performing an aggregate decomposition, i.e., dividing up $\bar{Y}_{2012} - \bar{Y}_{1999}$ into the two broad components $\bar{Y}_{2012} - \bar{Y}^C$ and $\bar{Y}_{1999} - \bar{Y}^C$. Note, however, that the contribution of each covariate to the wealth structure effect arbitrarily depends on the choice of the base group (Oaxaca and Ransom 1999) and has to be interpreted with caution. The detailed decomposition also critically relies on the assumption that the error term ε satisfies the zero conditional mean assumption. The assumption insures that the estimated effect of the covariates is not confounded by unobserved factors.[9]

Fortin et al. (2011) show that this convenient feature of the OB decomposition can be generalized to arbitrary measures of inequality. They show how to perform a detailed OB-type decomposition of inequality measures using the recentred influence function (RIF) regressions of Firpo et al. (2009). This provides a convenient way of analyzing the source of changes in inequality measures, such as the Gini coefficient or P-shares, despite the fact these measures are not decomposable in the sense defined in section 2. The remainder of this section describes in more detail decomposition methods based on reweighting and RIF regressions, focusing on the case of the Gini coefficient and P-shares.

9 For example, if ε represents cognitive skills that are positively correlated with education, the estimated effect of education will likely be biased because of the usual omitted variable bias problem. Counterfactual experiments based on changes in education will capture both the direct effect of education and the indirect effect of cognitive skills that are correlated with education. Interestingly, as long as the correlation between education and cognitive skills is same across groups and/or periods, the aggregate decomposition will remain valid. See footnote 13 and Fortin et al. (2011) for more details.

3.2. Reweighting

Inequality measures such as the Gini coefficient and P-shares can be represented as real-valued functionals $\nu : F_\nu \to \mathbb{R}$ of the underlying income (or wealth) distribution F_Y. For instance, the p-th Lorenz ordinate, $\nu(F_Y) = L(F_Y; p)$ can be represented as:

$$L\left(F_Y;p\right) = \frac{\displaystyle\int_{\underline{y}}^{q_p} y dF_Y(y)}{\displaystyle\int_{\underline{y}}^{\infty} y dF_Y(y)} = \frac{1}{\mu}\int_{\underline{y}}^{q_p} y dF_Y(y), \tag{11}$$

where $\underline{y}$ is the lower bound of the support of F_Y and μ represents its mean. The P-shares are simply differences of Lorenz ordinates.[10] Likewise, the Gini coefficient can be represented as:

$$G(F_Y) = \frac{1}{\mu}\int_{\underline{y}}^{\infty} F_Y\left(y\right)\left(1 - F_Y\left(y\right)\right) dy. \tag{12}$$

As discussed above, an aggregate decomposition of an inequality measure can be performed by computing a counterfactual value of this measure, which is itself a function of the underlying distribution F_Y. Thus, once we know how to compute the counterfactual distribution F_Y^C, it is straightforward to compute the counterfactual value of the Gini or P-shares.

Using the law of iterated probabilities, the (marginal) distribution of Y at time t can be written as: $F_{Y_t}(y) = \int F_{Y|X_t}(y|X) dF_{X_t}(X)$, where $F_{Y|X_t}$ is the conditional distribution of Y at time t given covariates X and F_{X_t} is the marginal distribution of X at time t.

Now consider the counterfactual distribution that would prevail if the distribution of covariates at time t was replaced by the distribution at another time period r. The resulting counterfactual distribution is:

$$F_{Y_t}^C\left(y\right) = \int F_{Y|X_t}\left(y|x\right) dF_{X_r}(x) \quad = \int F_{Y|X_t}\left(y|x\right)\psi_X(x) dF_{X_t}(x), \tag{13}$$

where the reweighting factor $\psi_X(x)$ is defined as: $\psi_X(x) = dF_{X_r}(x)/dF_{X_t}(x)$.[11] Using Bayes' law (DiNardo et al. 1996), it follows that:

$$\psi_X(x) = \frac{dF_{X_r}(x)}{dF_{X_t}(x)} = \frac{\Pr(X|T=r)}{\Pr(X|T=t)} = \frac{\Pr(T=r|X)}{\Pr(T=r|X)} \Big/ \frac{\Pr(T=t)}{\Pr(T=r)}. \tag{14}$$

$\Pr(T=r|X)$ can be computed by estimating a probit or logit model for the probability of being in period r (in a pooled sample for period r and t data) given X. The sample proportion $\Pr(T=r)$ is computed as the empirical fraction of

10 The Lorenz ordinate is typically computed over positive values in the case of income inequality. In the case of net worth that includes debt, negative values can be included.

11 Equation (13) makes explicit the fact that the assumption of invariance of the conditional distribution is maintained in the construction of counterfactuals. As discussed earlier, it excludes general equilibrium effects.

observations in period r, $\widehat{Pr}(T=r)$ when data from periods r and t are pooled together. The estimated reweighted factor $\hat{\psi}_X(x)$ is then obtained by plugging in the estimates $\widehat{Pr}(T=r|X)$ and $\widehat{Pr}(T=r)$ in equation (14).

In principle, the estimated reweighting factor could be used to construct an estimate $\hat{F}^C_{Y_t}(y)$ of the counterfactual distribution $F^C_{Y_t:X|t=r}(y)$, which, in turn, could be plugged into the equations for the Gini or Lorenz ordinates (equations (11) and (12)). A simpler procedure is to directly compute the distributional measure of interest by reweighting each observations with $\hat{\psi}_X(x)$. For instance, in the case of the p-th Lorenz ordinate, the estimated counterfactual value is: $\hat{L}^C_{RWT}(p)=\sum_{(y_i\leqslant \hat{q}^C_{p,t})}\hat{\psi}(X_i)y_i/\sum_{i=1}^{N_t}\hat{\psi}(X_i)y_i$, where $\hat{q}^C_{p,t}$ is the counterfactual p-th quantile.[12] This formula generalizes the simpler formula for counterfactual P-shares (or Lorenz ordinates) presented in Section 3.1.

Once an estimate of the counterfactual inequality measure is available, it is straightforward to compute an aggregate decomposition of changes in that measure over time. For example, using $\hat{L}_t(p)$ as a short for $\hat{L}(F_{Y_t};p)$ the change in the p-th Lorenz ordinate between 1999 and 2012 can be written as:

$$\begin{aligned}\hat{L}_{2012}(p)-\hat{L}_{1999}(p)=\hat{\Delta}^{L(p)}_{RWT}&\equiv\left(\hat{L}_{2012}(p)-\hat{L}^C_{2012}(p)\right)\\&+\left(\hat{L}^C_{2012}(p)-\hat{L}_{1999}(p)\right),\end{aligned}\tag{15}$$

where the first term on the right hand side represents the composition (or explained) effect, while the second term represents the wealth structure (or unexplained) effect. Intuitively, since $\hat{L}^C_{2012}(p)$ is obtained by replacing the 2012 distribution of Xs by the one in 1999, $\hat{\Delta}^{L(p)}_{RWT,X}=\hat{L}_{2012}(p)-\hat{L}^C_{2012}(p)$ should reflect solely changes in the distribution of covariates, i.e., composition effects. The remaining "unexplained" change, $\hat{\Delta}^{L(p)}_{RWT,S}=\hat{L}^C_{2012}(p)-\hat{L}_{1999}(p)$ depends on changes in the way covariates X map into Y.

When looking only at means under the assumption that $Y=X\beta+\varepsilon$, the parameters β summarize all the required information about the relationship between Y and X. Fortin et al. (2011) discuss in detail how the same rationale applies in the case of distributional statistics besides the mean. They consider a general case where Y depends in a fairly arbitrary way on X and ε through a general function: $Y=m(X,\varepsilon)$. Fortin et al. (2011) show that under the assumption that the time period indicator T is conditionally independent of ε given X, or $T\perp\varepsilon|X$ (also known as the ignorability assumption), the composition effect depends solely on changes in the distribution of X and ε, while the wealth structure effect in equation (15), $\hat{\Delta}^{L(p)}_{RWT,S}=\hat{L}^C_{2012}(p)-\hat{L}_{1999}(p)$, depends only on changes in the functions $m(.,.)$. In other words, although decompositions are often viewed as

12 The counterfactual quantile $\hat{q}^C_{p,t}$ can be computed using any statistical software (like Stata) that supports the use of weights in the computation of quantiles. Reweighted terms are computed by multiplying the sample weights by $\hat{\psi}_X(x)$

simple accounting exercises, they can be given more of a structural interpretation under the conditional independence assumption.[13]

3.3. Detailed decompositions based on RIF regressions

Reweighting methods provide a convenient way of computing counterfactuals based on secular changes in the distribution of all covariates X. But as discussed in section 3.1, we are often interested in computing counterfactuals linked to specific covariates such as educational achievement, family composition, etc. These covariate-specific counterfactuals are the building blocks of detailed decompositions à la Oaxaca–Blinder.

Firpo et al. (2007, 2009) propose to estimate recentred influence function (RIF) regressions as a way of estimating these counterfactuals. Influence functions, also known as Gâteaux (1913) derivatives, were introduced by Hampel (1974) as a tool for robustness analysis. The influence function $IF(y;\nu)$ of a distributional statistic ν evaluated at $Y=y$ indicates by how much ν changes when there is a small increase in the fraction of the distribution F_Y concentrated at $Y=y$. More formally, $IF(y;\nu)$ is a directional derivative indicating by how much $\nu(F_Y)$ changes when an (infinitesimally) small step is taken in the direction of a mass point distribution centred at $Y=y$.[14]

To provide intuition on how the influence function can be used to compute counterfactuals, consider what happens when we increase (by a small amount) the share of university relative to high school educated individuals. As the average influence function among university and high school educated students is $\mathbb{E}[IF(y;\nu)|\text{Univ}]$ and $\mathbb{E}[IF(y;\nu)|\text{HS}]$, respectively, the effect of changing the share of university-educated individuals is given by $\mathbb{E}[IF(y;\nu)|\text{Univ}]-\mathbb{E}[IF(y;\nu)|\text{HS}]$.[15] Ignoring other covariates, this difference corresponds to the coefficient in a bivariate regression of the influence function on a dummy variable for university education (with high school as the base group). Thus, as in a standard OB decomposition, regression methods can be used to compute counterfactuals when the regressand is an influence function instead of y.

13 The conditional independence assumption is slightly weaker than the independence assumption as it allows X and ε to be correlated, provided that the correlation does not change over time. In addition to conditional independence, Fortin et al. (2011) show that three other conditions must hold for the structural interpretation to be valid. The first condition requires two mutually exclusive groups—which correspond to the two time periods in this case. Second, the counterfactual must be simple, in that it refers to the wealth structure of one group or the other. Third, the support of the distribution of the two groups must be overlapping in its entirety.

14 To measure the influence of a particular point y of the distribution, the idea is to construct a mixture of the actual distribution F and a contamination of F at point y: $T=(1-\varepsilon)F+\varepsilon\delta_y$, where δ_y is a degenerate distribution with mass of 1 at point y. The influence function of the distributional statistic $\nu(F)$ is then obtained as the directional derivative of $\nu(T)$ as ε goes to zero: $IF(y;\nu)=\lim_{\varepsilon\to 0}[\nu(T)-\nu(F)]/\varepsilon$

15 Firpo et al. (2009) provide a formal derivation of how to use the influence function to compute the effect of a small change in the distribution of covariates. Applied to this specific example, their theorem 1 implies that the effect of a small change Δs in the share of university-educated individuals on the distributional statistic ν is given by $\Delta s\cdot(\mathbb{E}[IF(y;\nu)|\text{Univ}]-\mathbb{E}[IF(y;\nu)|\text{HS}])$.

For the purpose of performing decompositions, it is more convenient to work with the recentred influence function that is obtained by adding the distributional statistic to the *IF*. This insures that the change in the average value of the *RIF* over time is equal to the change in the distributional statistics.[16] Note that in the case of the mean, the influence function, $IF(y;\mu)$, is $y-\mu$ and $RIF(y;\mu)=IF(y;\mu)+\mu=y$. Thus, in that simple case, counterfactuals can be computed by running a standard OLS regression of the *RIF*—y in this case—on the covariates and using the estimated regression coefficients $\hat{\beta}$ to compute covariate-specific counterfactuals. Firpo et al. (2009) show that the same regression approach can be used for other distributional statistics when y is replaced by the relevant *RIF*. One potential limitation is that for statistics other than the mean, the $RIF(y;\nu)$ is based on a first-order approximation of the impact of y on the distributional statistic. For this reason, Fortin et al. (2011) propose a decomposition procedure that combines an OB decomposition with reweighting and ensures that the decomposition separates composition and wealth structure effects even if the first-order approximation is valid only locally.

While Firpo et al. (2009) focus mostly on the case of quantiles, the *RIF* can be readily computed for other distributional statistics such as the Gini (Monti 1991) and Lorenz ordinates (Essama-Nssah and Lambert 2012). The *RIF* for the Gini is given by:

$$RIF\left(y;G\right)=2\frac{y}{\mu}G+1-\frac{y}{\mu}+\frac{2}{\mu}\int_0^y F(z)dz, \tag{16}$$

while the *RIF* for the p-th Lorenz ordinate is:

$$RIF\left(y;L\left(p\right)\right)=\begin{cases}\dfrac{y-\left(1-p\right)q_p}{\mu}+L\left(p\right)\cdot\left(1-\dfrac{y}{\mu}\right) & \textit{if } y<q_p\\ \dfrac{pq_p}{\mu}+L\left(p\right)\cdot\left(1-\dfrac{y}{\mu}\right) & \textit{if } y\geqslant q_p.\end{cases} \tag{17}$$

By design, influence functions integrate to zero and recentred influence functions integrate to the distributional statistic $\nu(F)$. Thus, using the law of iterated expectations, we can write: $\nu(F_Y)=\mathbb{E}_X[\mathbb{E}[RIF(y;\nu)|X]]$. If we assume that, as in the case of the mean, the conditional expectation of the RIF can be represented as a linear function, $\mathbb{E}[RIF(y;\nu)|X]=X\gamma$, where γ represent the parameters of the RIF regression, it follows that:

$$\nu\left(F_Y\right)=\mathbb{E}\left[X\right]\gamma. \tag{18}$$

Each element of the parameter vector γ indicates by how much the distributional statistic $\nu(F_Y)$ changes in response to a change in the mean value of the corresponding element of the covariate vector X. In other words, γ can be used to compute covariate-specific counterfactuals and form the basis of an

16 Since recentring the influence function involves only adding a constant to the *IF*, the estimated coefficients from a regression of the *IF* or *RIF* on the covariates are identical except for the constant.

OB-type decomposition of $\nu(F_Y)$, as in the case of the counterfactual experiment discussed above (increase in the share of university-educated individuals). After plugging in the sample estimates into equation (18), we have: $\hat{\nu} = \bar{X}\hat{\gamma}$, where $\hat{\gamma}$ is estimated by running an OLS regression of $RIF(y;\nu)$ on X. It follows that an OB-type decomposition for the distributional statistic ν can be written down as: $\hat{\nu}_t - \hat{\nu}_r = \hat{\Delta}^{\nu}_{OB} \equiv (\bar{X}_t - \bar{X}_r)\hat{\gamma}_t + \bar{X}_r(\hat{\gamma}_t - \hat{\gamma}_r)$.

As noted earlier, one concern with the approach based on RIF regressions is that it may not provide an accurate estimate of counterfactuals when changes in X are large, given that the *RIF* is based on a linear approximation. The importance of this problem can be assessed by comparing RIF- and reweighted-based counterfactuals. Consider the counterfactual ν^C where the distribution of covariates at time t is replaced by the distribution at time r.[17] When using RIF regressions, the counterfactual is estimated as $\hat{\nu}^C_{RIF} = \bar{X}_r\hat{\gamma}_t$.

By definition, the average value of the *RIF* in the reweighted sample is equal to the reweighted estimate of the distributional statistic. Using again the assumption that the *RIF* is linear in X, we can write the counterfactual estimate under reweighting as $\hat{\nu}^C_{RWT} = \bar{X}^C_t\hat{\gamma}^C_t$, where $\bar{X}^C_t$ is the reweighted average of X obtained using the reweighting factor $\hat{\psi}_X$, while $\hat{\gamma}^C_t$ is the OLS estimate from a regression of *RIF* on X in the reweighted sample.[18]

As discussed in the empirical application of section 4, the reweighted average $\bar{X}^C_t$ tends to be very close to $\bar{X}_r$ in practice. The main source of discrepancy between $\hat{\nu}^C_{RIF}$ and $\hat{\nu}^C_{RWT}$ is, therefore, potential differences between the OLS estimates obtained with $(\hat{\gamma}^C_t)$ and without $(\hat{\gamma}_t)$ reweighting. Fortin et al. (2011) discuss how this difference is linked to specification errors in the linear regression equation. They show that while differences between $\hat{\gamma}_r$ and $\hat{\gamma}_t$ may either be due to changes in the wealth structure (the underlying $m(.,.)$ functions) or specification errors, the difference between $\hat{\gamma}_r$ and $\hat{\gamma}^C_t$ reflects solely differences in the wealth structure between the two periods. They address this issue by adding and subtracting alternative counterfactuals, $\bar{X}^C_{2012}\hat{\gamma}^C_{2012}$ and $\bar{X}_{1999}\hat{\gamma}^C_{2012}$. After a few re-arrangements, this yields the alternative OB decomposition with reweighting (OBR) of changes in the distributional statistic between 1999 and 2012:

$$\begin{aligned}\hat{\nu}_{2012} - \hat{\nu}_{1999} = \hat{\Delta}^{\nu}_{OBR} \equiv & \left(\bar{X}_{2012} - \bar{X}^C_{2012}\right)\hat{\gamma}_{2012} \;+\; \bar{X}_{1999}\left(\hat{\gamma}^C_{2012} - \hat{\gamma}_{1999}\right) \\ & + \left(\bar{X}^C_{2012} - \bar{X}_{1999}\right)\hat{\gamma}^C_{2012} + \bar{X}^C_{2012}\left(\hat{\gamma}_{2012} - \hat{\gamma}^C_{2012}\right).\end{aligned} \tag{19}$$

The first two terms in equation (19) are the adjusted estimates of the composition and wealth structure effects. The third term reflects possible reweighting errors, while the fourth term represents the specification error, just discussed. Finding a small specification error suggests that the RIF regressions provide an accurate way of computing counterfactuals. Likewise, a small reweighting error

17 Fortin et al. (2011) show that reweighting provides a consistent estimate of the counterfactual provided that the logit or probit used for computing the reweighting factor is non-parametrically estimated (i.e., flexible enough in X).

18 See Fortin et al. (2011) for more details.

indicates that the probit or logit model being estimated captures well the changes in the distribution of X over time.

4. Empirical evidence

4.1. Data and descriptive statistics

We investigate changes in wealth inequality using data from the four wealth surveys conducted by Statistics Canada over the last 35 years: the Assets and Debts module (ADS) of the 1984 Survey of Consumer Finances (SCF) and the 1999, 2005 and 2012 Survey of Financial Security (SFS). These specialized modules and surveys collect information from a relatively small sample (from 9,000 to 24,000) of Canadian families on their assets, debts, employment, income and education and in the SFSs, include employer-sponsored pension plans valued on a termination basis. Appendix table A1 provides some descriptive statistics on the socio-demographic variables that we utilize. Some differences between these data sources are worth noting. A fundamental difference between income and wealth inequality is that while the first is measured at the individual level, the latter is usually available only at the family level.[19] Thus as with changes in family earnings inequality (Fortin and Schirle 2006), changes in family formation figure prominently as a driving force in the evolution of wealth inequality.[20] In turn, changes in family formation are arguably driven less by economic forces than socio-demographic changes, which are less likely influenced by public policy. The ADS is based on information from individual surveys of family members aged 15 plus on assets (except housing) and debts, which is then aggregated to the family level. In contrast, the SFSs collect this information directly at the family level and use a supplementary "high-income" sample to improve the quality of wealth estimates.

The range of assets surveyed differs substantially between the ADS and the SFS, as the latter includes information on employer-sponsored retirement plans.[21] Because of the importance of this source of wealth for families at the lower end of the distribution, we will focus mostly on the evolution (from 1999 to 2012) of the more complete measure of wealth—net worth with pensions. It is defined as the difference between total assets and total debts, but where total assets include

19 In Canada, where couples file their income tax separately, it would be in principle possible to know the extent to which wealth is divided unequally within couples. For the United States, Saez and Zucman (2016) make the assumption that wealth is divided equally within couples.

20 In the analysis of family earnings inequality, it is common practice because of economies of scale in consumption to use family equivalent scales by dividing the total family income by the square of the number of family members for example (OECD 2013). That case is less clear in terms of wealth, which can be cast in terms of future consumption or bequest motives. We refer the reader to Cowell and Van Kerm (2015) for a complete discussion of this issue.

21 In addition, SFS assets include contents of the home, collectibles and valuables, annuities and registered retirement income funds (RRIFs). Also the ADS does not include mortgages on real estate other than the primary residence, but it includes in assets a variable called "cash on hand" not covered in the SFS.

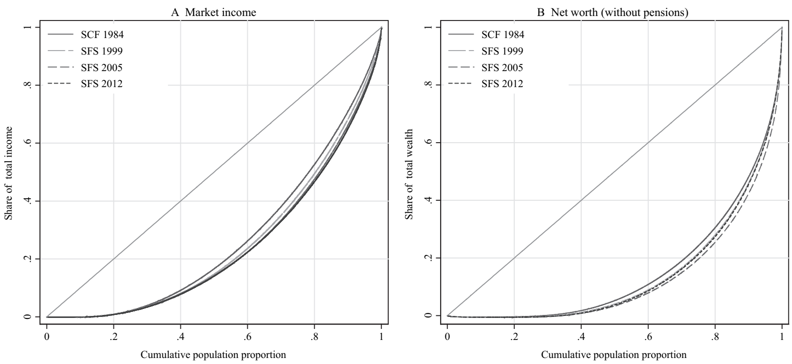

FIGURE 1 Lorenz curves for market income and net worth (without pensions)

the value of employer-sponsored pensions. We direct the reader to Morissette et al. (2006) and Morissette and Zhang (2007) for an exhaustive analysis of net worth without pensions, which makes the ADS and SFS more comparable. As noted in the introduction, issues surrounding the coverage of the top end of the wealth distribution and top coding are important caveats to take into account in the interpretation of the results. The combination of small sample sizes and sporadic temporal coverage has also influenced the methodologies used to describe the evolution of wealth inequality, until recently.[22]

We begin in figure 1 by displaying the Lorenz curve for market income in panel A and net worth (without pensions) in panel B.[23] Because debts can exceed assets, the measure of net worth can take on negative values. But as long as the mean net worth is positive, the wealth shares, Lorenz curve and Gini coefficient will be well defined. Also displayed in the figure is the 45-degree line; the area between this line of equality and the Lorenz curve corresponds to half of the Gini coefficient.

Figure 1 illustrates two well-known stylized facts about income and wealth inequality in industrialized economies (Davies and Shorrocks 2000, Cowell and Van Kerm 2015). First, wealth inequality is much higher than income inequality. As can be seen by looking carefully at the graph, panel A shows that the bottom decile in 2012 (most outward dashed line) has a negative share or null total income share while the top decile gets 35% of total income. On the other hand, panel B shows much fatter tails for wealth, with the bottom 35% holding negative or zero wealth and the top decile holding 54% of total wealth. The

22 See Morissette and Zhang (2007) on the consequences of changes in interviewing techniques on the ability to capture high net worth individuals across survey waves. They also note that the degree of truncation may have changed over time. These authors compare all measures to those that exclude families in the top 1% and top 5% to assess the impact of the changes.

23 Recall that the Lorenz curve is the graph, $\{(p, L(F;p)) : 0 \leqslant p \leqslant 1\}$, of the p-th wealth percentile and the wealth share (Lorenz ordinate) $L(F;p) = (1/\mu)\int_{\underline{y}}^{q_p} y dF(y)$, where $q_p = Q(F;p) = \inf(y|F(y) \geqslant p)$ and $\underline{y}$ is the lower bound of the support of F, the wealth distribution, which can be negative.

TABLE 1
Gini coefficient by assets components

Year	(1) 1984	(2) 1999	(3) 2005	(4) 2012
Market income	0.472	0.503	0.519	0.526
	(0.003)	(0.004)	(0.007)	(0.005)
Wealth without pensions	0.694	0.718	0.741	0.721
	(0.006)	(0.005)	(0.010)	(0.005)
Wealth with pensions		0.665	0.685	0.672
		(0.004)	(0.009)	(0.005)
Total assets without pensions	0.644	0.645	0.671	0.653
	(0.006)	(0.005)	(0.010)	(0.006)
Total assets with pensions		0.614	0.634	0.623
		(0.004)	(0.009)	(0.005)
Housing	0.633	0.616	0.646	0.630
	(0.004)	(0.004)	(0.012)	(0.006)
Debt	0.764	0.719	0.720	0.730
	(0.003)	(0.003)	(0.007)	(0.005)

NOTES: Negative and null values of assets are included in the computation using the *sgini* (Van Kerm 2009) Stata routine. Standard errors are computed using the jackknife procedure.

above discussion shows how difficult it is to quantify those changes in terms of "what happens where" in the distribution using Lorenz curves. Another important stylized fact—increasing inequality over time—is illustrated by the fact that Lorenz curves for both market income and net worth (without pensions) are becoming more convex over time. However, the changes in Lorenz ordinates are not statistically significantly different, except for market income between 1984 and 1999.[24] We discuss next the fact that changes over time in the corresponding Gini coefficients are rarely statistically significant.

Table 1 reports the Gini coefficients by asset classes to illustrate in which classes there is more inequality. Recall that the formula for the Gini coefficient can be written in terms of the Lorenz curve $G(F) = 1 - 2\int_0^1 L(F;p)dp$, where $F(\cdot)$ is the distribution of the asset class. In these small samples, there are significant increases from 1984 to 1999 in the Gini coefficients for market income and wealth without pensions, but significant decreases for housing and debt. From 1999 to 2005, we can see statistically significant increases in all inequality components, except debt. The changes from 2005 to 2012 correspond to decreases in inequality, except for debt, but these changes are not statistically significant. Only market income shows continuous increases over the entire period, whereas there are no significant increases from 1999 to 2012 in any asset component.

We turn next to the modern way (Piketty and Saez 2003, 2013) of describing changes in inequality in terms of P-shares. The P-share, $S(p_1, p_2) = L(F;p_2) - L(F;p_1)$ with $p_1 \leqslant p_2$, corresponds to the proportion of total wealth that falls into the interval $[p_1, p_2]$. Because pensions are an important asset for households

24 For clarity, we do not display the confidence bands in figure 1, but they are available upon request.

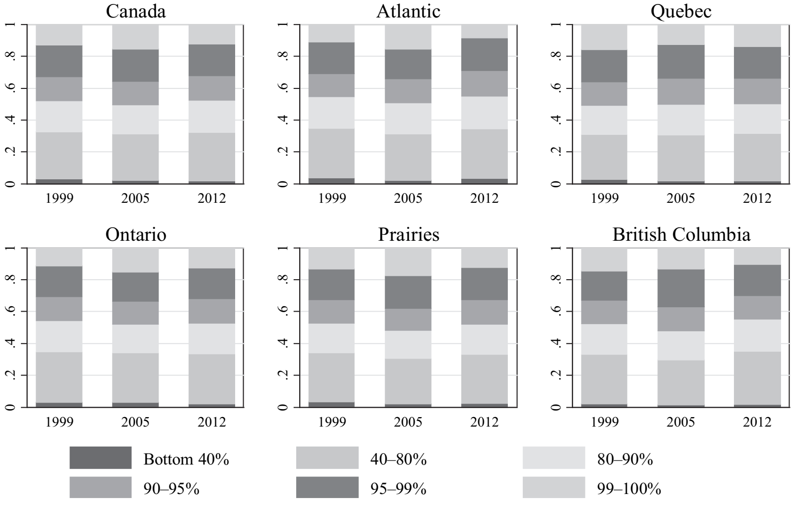

FIGURE 2 P-shares of net worth (with pensions)

at the lower end of the distribution, we focus on net worth including pensions. Thus we omit the 1984 ADS when we display the changes over time in the shares accruing to six percentile groupings, which clearly separates the middle quintiles (P40–80) from other quintiles and provides a detailed partition of the upper decile.[25]

Figure 2 shows that for Canada as a whole, the shares of total wealth including pensions accruing to the top centile increased from 13.2% to 15.5% from 1999 to 2005, but following the GFC of 2007/08 decreased to 12.5% in 2012.[26] There were barely any significant changes for the upper-middle class (P80–99). The share of total wealth accruing to the middle class (P40–80), around 30%, saw swings similar to those of the top 1% but about half in magnitude and less statistically significant. The two bottom quintiles saw statistically significant declines in their shares from about 3% to 2% over the entire period. This qualitative description applies to each region, with differences in the magnitude of the rebound of the top centile. In Quebec, the middle class experienced continued increases at the expense of the upper middle class. In British Columbia, the top centile share experienced continued decreases and the middle class experienced a greater rebound in 2012 ending up with a 33% share instead of 30% in Canada.

25 See Bonesmo Fredriksen (2012) for a comparison of similar P-shares between Canada and seven other industrialized countries, computed with the Luxembourg Income Study data circa 2000.

26 Because of the smaller number of observations in 2005, the changes from 1999 to 2005 in the wealth of the top centile is significant only at the 10% level. The decline in the share of the two bottom quintiles is significant at the 5% level. Changes for the other P-shares are not statistically significant. Appendix table A2 displays the numbers behind figure 2 with standard errors computed using Jann (2016) Stata routine.

TABLE 2
Decomposition of Gini coefficient by family types

Year	1984	1999	2005	2012
Aggregate wealth inequality	0.694	0.665	0.685	0.672
	(0.006)	(0.004)	(0.009)	(0.005)
Between-family types	0.239	0.265	0.258	0.303
	(0.010)	(0.008)	(0.037)	(0.008)
Within-family types	0.155	0.135	0.142	0.139
	(0.002)	(0.002)	(0.003)	(0.002)
Overlap	0.300	0.265	0.285	0.230
	(0.009)	(0.007)	(0.037)	(0.006)

NOTES: The measure of wealth is net worth with pensions, except for 1984, when the employer sponsored pensions are unavailable. Decomposition by sub-groups performed using the *ginidesc* Stata routine (Aliaga and Montoya 1999). Standard errors are computed using the jackknife procedure.

4.2. Decompositions by family types and by reweighting

Among the most important explanatory factors behind changes in wealth inequality, previous research has identified changes in family types and life-cycle patterns as unavoidable ones (Pendakur 1998, Milligan 2005, Morissette and Zhang 2006). We regroup the family types into six categories to allow comparisons across survey waves, distinguishing elderly and non-elderly households. Among the non-elderly, there are four groups: unattached individuals, lone parent families, other families with children and families without children. As shown in table A1, there has been an increase in the fraction of households without children (either single or couples) and a decrease in families with children. Among the elderly, we distinguish single individuals and couples. In table 2, we provide a classic decomposition of the Gini coefficient by sub-groups, in this case, family types.

The overall Gini coefficient can be decomposed into a between-family types component plus the sum of Gini coefficients within each family type weighted by the population and wealth share of each family type as:

$$G\left(F\right) = G_B(F) + \textstyle\sum_{j=1}^{J} s_j \pi_j G\left(F_j\right) + R, \tag{20}$$

where $G_B(F)$ is the "between-group" Gini coefficient, F_j is the wealth distribution within family type j, s_j and π_j are respectively the population share and the total wealth share of each family type j. In this case, the between-group Gini coefficient is derived by assigning to each individual the mean wealth within their family group. Finally, R captures the degree of overlap between the wealth distributions of the different family groups.

Table 2 reports the results of this decomposition for the family types described above. The precise evolution of the shares of family types is reported in appendix table A1. It shows the decreasing importance of families with children (both lone parent and couples) whose share decreased by 12.4 percentage points and represented less than a quarter of households in 2012. The growing importance of households without children among the non-elderly (both unattached individuals and couples), whose share increased by 8.5 percentage points over the period, is substantial; these households constituted more than half of all households in 2012.[27] The growth in the share of elderly households (both unattached individuals and couples) of 3.8 percentage points implies that elderly households represented more than 20% of households in 2012.[28] Given the greater wealth shares of the families without children and elderly couples, the *between* component always dominates the *within* component. In addition, reflecting the fact that the Gini coefficient is not a fully decomposable index, the overlap dominates the within changes and often the between component. Despite the fact that changes over time in the components of wealth inequality are generally not statistically significant, the results of table 2 highlight the important role of family formation and population aging in changes in wealth inequality.

Our next exercise is thus to construct counterfactuals similar to those presented in Morissette and Zhang (2007) to assess the impact of changes in socio-demographic characteristics on wealth inequality. Here we additionally distinguish the role of family types from that of other covariates. We construct counterfactuals that show what the distribution of wealth would have been in the absence of changes in these covariates, such as educational improvement, population aging, or changes in family formation, using a reweighting factor $\psi_X(x)$ defined in equation (14) under the assumption of invariance of the conditional distribution. For example, we ask what would be the distribution of wealth in 2012 if the above covariates had stayed at their 1999 level.

To estimate $\Pr(T=2012|X)$, we pool data from the two waves and estimate a logit for the probability of being in year 2012 as a function of the Xs. We use the simple logit formulation to simplify the exposition, though most empirical studies use a more flexible functional form to ensure that the estimated model fits well the odds ratio. A flexible specification involving a large number of interactions between covariates will yield a more accurate prediction, but one has to be careful that no single value of a particular variable or interaction becomes a perfect predictor, as this would violate the assumption of overlapping or common support.

To isolate the effect of one particular covariate, let's say family type, U, from the set of all covariates, $X=\{U,Z\}$ we compute a counterfactual that keeps that factor at the 2012 level by excluding that covariate from the reweighting:

$$F^C_{Y_{2012}:U|T=2012,Z|T=1999} = \int F_{Y_t|U,Z}(y|u,z)\psi_{U|Z}(u,z)dF_{U_t}(u|z)dF_{Z_t}(z), \qquad (21)$$

27 Because of the smaller number of observations in 2005, the changes from 1999 to 2005 in the wealth of the top centile is significant only at the 10% level. The decline in the share of the two bottom quintiles is significant at the 5% level. Changes for the other P-shares are not statistically significant. Appendix table A2 displays the numbers behind figure 2, with standard errors computed using Jann (2016) Stata routine.

28 We note that part of the change may be linked to the life cycle of baby boomers that have become empty nesters over the period.

TABLE 3
Gini and P-shares under alternative counterfactuals obtained by reweighting

Year/ counterfactual	1999 (1)	2005 (2)	2005 with 1999 Xs (3)	2005 with 1999 Xs w/o FT (4)	2012 (5)	2012 with 1999 Xs (6)	2012 with 1999 Xs w/o FT (7)
1: Gini	0.665	0.685	0.693	0.693	0.672	0.692	0.695
	(0.004)	(0.009)	(0.010)	(0.010)	(0.005)	(0.005)	(0.005)
%Δ from 1999		2.9%**	4.2%**	4.2%**	1.1%	4.0%**	4.5%***
2: P0–40	0.028	0.022	0.019	0.020	0.020	0.015	0.015
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
%Δ from 1999		−21.4%***	−32.1%***	−28.6%***	−28.6%***	−46.4%***	−46.4%***
3: P40–80	0.300	0.289	0.279	0.282	0.304	0.289	0.285
	(0.004)	(0.008)	(0.010)	(0.010)	(0.005)	(0.005)	(0.005)
%Δ from 1999		−3.7%	−7.0%**	−6.0%*	1.3%	−3.7%*	−5.0%*
4: P80–90	0.193	0.185	0.185	0.183	0.197	0.200	0.197
	(0.002)	(0.005)	(0.006)	(0.006)	(0.002)	(0.003)	(0.003)
%Δ from 1999		−4.1%	−4.1%	−5.2%*	2.1%	3.6%**	2.1%
5: P90–95	0.148	0.146	0.146	0.144	0.153	0.153	0.152
	(0.002)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
%Δ from 1999		−1.4%	−1.4%	−2.7%	3.4%*	3.4%*	2.7%
6: P95–99	0.199	0.203	0.212	0.207	0.200	0.210	0.213
	(0.003)	(0.006)	(0.008)	(0.008)	(0.004)	(0.004)	(0.005)
%Δ from 1999		2.0%	6.5%	4.0%	0.5%	5.5%**	7.0%**
7: P99–100	0.132	0.155	0.159	0.164	0.125	0.133	0.138
	(0.005)	(0.013)	(0.017)	(0.017)	(0.005)	(0.005)	(0.005)
%Δ from 1999		17.4%	20.5%	24.2%	−5.3%	0.8%	4.5%

NOTES: The measure of wealth is net worth with pensions. Counterfactuals are computed using reweighting on the available covariates: head's age, head's education (4 categories), family size (5 categories), family types (6), regions (5) in columns (3) and (6). In columns (4) and (7), family types (FTs) are omitted thereby revealing the impact of family types. ***p< 0.01, **p< 0.05, *p< 0.10.

where $\psi_{U|Z}(u,z) = \frac{dF_{U|T=1999}(u|z)}{dF_{U|T=2012}(u|z)} = \frac{\psi_{UZ}(u,z)}{\psi_Z(z)}$. The difference between the two counterfactuals $F^C_{Y_{2012}:X|T=1999}$ and $F^C_{Y_{2012}:U|T=2012,Z|T=1999}$ will give the effects of U.

The results are presented in table 3, which displays the Gini coefficient for net worth with pensions for the years 1999, 2005 and 2012, as well as the P-shares for the six intervals, illustrated in figure 2, in columns (1), (2) and (5). The counterfactuals that bring back all covariates to their 1999 levels are presented in columns (3) and (6). Those that leave family formation at its contemporaneous level are shown in columns (4) and (7). As explained above, wealth inequality as measured by the Gini coefficient increased from 1999 to 2005, but in 2012 was back at a level not statistically significantly different from 1999. Interestingly, if these changes in socio-demographic characteristics had not taken place, wealth inequality would have been even significantly larger in 2005 and 2012; without changes in family formation between 2005 and 2012, inequality would have remained as high as in 2005.

Changes in P-shares are more informative about the winners and losers of increasing wealth inequality. While the Gini shows no changes in inequality from 1999 to 2012, the P-shares in panel 2 show that the shares of the wealth accruing to the two bottom deciles decreased by 0.7 percentage points, a 25% decline.[29] That decline would have been larger by 0.3 to 0.5 percentage points if not for the changes in socio-demographic characteristics, although changes in family formation were conducive to a lower decline.

Contrary to what some might fear given the meagre market income increases experienced by the middle class over the period, the next four deciles (P40–80, the wider middle class) in panel 3 did not see much decline in their share of total wealth, with initial declines largely compensated by later gains. The middle class's share of total wealth stood at 30% in 1999. After going down to 28.9% in 2005, it was back up to 30.4% in 2012. The second upper decile experienced a similar larger decline from 1999 to 2005, going from 19.3% to 18.5%, but rebounded to 19.7% in 2012. The wealth shares of the next 5% and next 4% show similar largely non-significant changes of about half a percentage point. The top 1% experienced a larger increase in percentage terms followed by a sharp decrease. Because these changes are largely not statistically significant, the appropriate interpretation of these numbers is that not much changed in the wealth distribution.

Our reweighting exercise shows that the socio-demographic changes provided a larger buffer against increasing wealth inequality for the four lowest quintiles (P0–40 and P40–80). Reweighting the samples to make them look like 1999 reduces their share of wealth even more, yielding statistically significant declines. The socio-demographic changes had almost no impact on the second upper decile (P80–90). For the P95–99 and top 1%, these changes appear to reduce the shares accruing to these groups; reweighting the samples to make them look like 1999 increases their share of wealth even more, although there is a lack of statistical significance.

We turn next to a decomposition exercise that allows us to perform a detailed decomposition of the Gini indexes and the related P-shares to better assess how the socio-demographic factors interact.

4.3. Decomposition using recentred influence functions

An important advantage of using RIF regressions is that it allows the decomposition of any distributional statistic for which the influence function exists. Firpo et al. (2007) initially focused only on quantiles, the variance of logs and the Gini. Several papers have begun to use the RIF–Gini regressions to explore changes in income inequality (Choe and Van Kerm 2014, Gradin 2016). Carpentier et al. (2017) have used the approach to assess the effects of caps on loan-to-value (LTV) ratios on net wealth inequality across several EU countries.[30] Cowell et al. (2017)

29 A similar decline in the lower two quintiles has been reported in Uppal and LaRochelle-Côté (2015).

30 They find that among households with active mortgages, those with higher LTV ratios tend to be in tails of the distribution thereby increasing wealth inequality, but there are some offsetting effects of increases in house prices in the middle of the distribution.

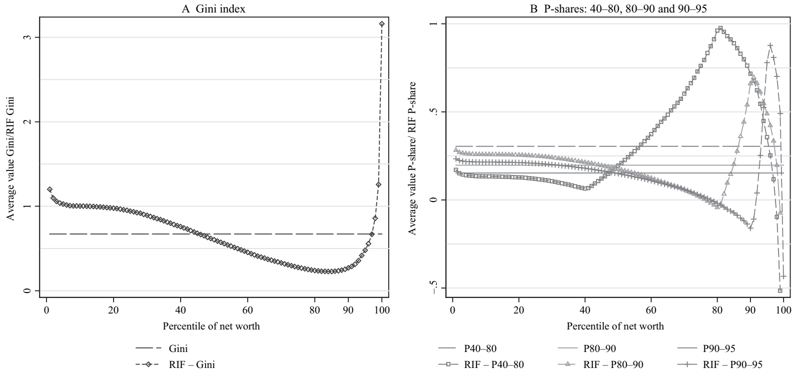

FIGURE 3 Gini index and P-shares vs. respective RIF by percentiles of net worth (SFS 2012)

use RIF–Gini regressions to study the role of inheritance in wealth inequality across several OECD countries. Here we also estimate RIF-P-shares regressions, which given the close link between the Gini index and Lorenz ordinates, give us the opportunity to shed new light on the empirical contribution of different P-shares to the Gini index.

Integrating by parts equation (16) leads to the following more intuitive expression for the recentred influence function of the Gini index, G:

$$RIF\left(y;G\right)=2\frac{y}{\mu}\left[F\left(y\right)-\frac{1+G}{2}\right]+2\left[\frac{1-G}{2}-L\left(F(y)\right)\right]+G, \tag{22}$$

where $L(F(y))$ is the Lorenz ordinate at $p=F(y)$, $\frac{1+G}{2}$ and $\frac{1-G}{2}$ corresponds respectively to the areas above and below the Lorenz curve. As pointed out by Monti (1991), the first term is unbounded because it increases by the factor y/μ, while the second is bounded between $G-1$ and $1+G$. Thus the $RIF(y;G)$ is continuous and convex in y; its first derivative is equal to $\frac{2}{\mu}[F(y)-\frac{1+G}{2}]$, and it reaches its minimum when $F(y)=\frac{1+G}{2}$.[31] Given the range of the Gini index of wealth in our samples (around 0.67), this minimum should be reached around the 84$^{\text{th}}$ percentile. The function is theoretically unbounded from above, but in practice it reaches its maximum at the upper bound of the empirical support of the distribution. This implies that the Gini index is not robust to measurement error in high incomes, as pointed out by Cowell and Victoria-Fesser (1996).

Figure 3a illustrates, for 2012, the average value of the $RIF(y;G)$ by percentile of net worth in comparison with its mean value, which is equal to the Gini index by construction.[32] It shows that, as expected, the values of the $RIF(y;G)$ are higher than the Gini both at the bottom and very top of the wealth distribution. However, the proportion of households with values above the mean

31 The second derivative of $RIF(y;G)$ is $\frac{2}{\mu}\frac{dF(y)}{dy}=\frac{2}{\mu}f(y)\geqslant 0$.

32 Figures for the other years are similar with slight differences at the very top.

TABLE 4
Oaxaca–Blinder decomposition using RIF regressions

Inequality measure:	Gini	P0–40	P40–80	P80–90	P90–95	P95–99	P99–100
A: Changes from 1999 to 2005 (X 100)	1.957*	−0.579***	−1.092	−0.773	−0.211	0.344	2.311
Composition effects attributable to							
Family types	0.007	−0.098*	0.156	0.209*	0.090	−0.179*	−0.178
Family size	0.105	−0.030*	−0.093	0.014	0.049	0.067	−0.008
Head's age	−0.457*	0.152*	0.330*	−0.062	−0.073*	−0.199*	−0.148
Head's education level	−0.657***	0.185***	0.471***	0.117	0.089	−0.282**	−0.580**
Regions	−0.018	0.005	0.018	−0.006	−0.008	−0.013	0.003
Total composition effect	−1.021***	0.216**	0.882***	0.271**	0.147	−0.606***	−0.910***
Total unexplained	2.978***	−0.795***	−1.973*	−1.044*	−0.359	0.950	3.221*
B: Changes from 1999 to 2012 (X 100)	0.701	−0.738***	0.467	0.379	0.508*	0.081	−0.697
Composition effects attributable to							
Family types	0.709***	−0.176***	−0.592***	−0.136*	0.142**	0.412***	0.350*
Family size	0.289***	−0.052***	−0.281***	−0.099*	0.007	0.238***	0.187
Head's age	−1.844***	0.504***	1.460***	−0.110	−0.221***	−0.960***	−0.673***
Head's education level	−1.182***	0.341***	0.931***	−0.081	−0.043	−0.665***	−0.484*
Regions	−0.018	0.004	0.022	−0.015	−0.014	0.003	0.001
Total composition effect	−2.046***	0.621***	1.540***	−0.440***	−0.129	−0.973***	−0.619*
Total unexplained	2.747***	−1.359***	−1.073*	0.820**	0.637*	1.053**	−0.078

NOTES: The composition effects are the time (t_1-t_0) differences in the explanatory variables multiplied by the corresponding t_1 coefficients. All variables, except age entered as quadratic, are categorical. There are 6 family types, 4 family sizes, 6 education classes and 5 regions. ***p< 0.01, **p< 0.05, *p< 0.1.

is much larger in the bottom two quintiles than at the top.[33] The numbers behind the graph reveal that the 40% of poorest households account for about 20% of the total Gini, while the richest percentile account for 4.7%. This implies that the contribution of the top percentile to the Gini is lower than its contribution to the mean wealth of 12.5% in 2012.

As discussed in section 2, the impact of a transfer between two households on the Gini coefficient depends only on the proportion of households between the donor and the recipient household. The same result can be shown in the case of the $RIF(y;G)$. Since the $RIF(y;G)$ is a linear function in $F(y)$, it follows that the impact of a transfer Δ from a donor with wealth y' to a recipient with wealth y is given by $(\frac{1}{N})\Delta\frac{2}{\mu}[F(y')-F(y)]$, where $F(y')-F(y)$ is the proportion of observations between the recipient and the donor.[34]

That said, plotting the $RIF(y;G)$ provides a more nuanced interpretation of the notion that the Gini coefficient is particularly sensitive to what happens in the middle of the distribution. For instance, figure 3a suggests that increasing the proportion of a group concentrated in the upper middle (e.g., university graduates) relative to the lower middle (e.g., high school graduates) of the distribution would likely reduce the Gini coefficient. As we will see below, this is consistent with the findings in tables 4 and 5 that secular increases in the proportion of university graduates contributed to a reduction of the Gini coefficient over time.

For the P-share, $S(p_1,p_2)=L(p_2)-L(p_1)$, where $L(p)$ is the Lorenz ordinate, differencing the RIF for Lorenz ordinates in equation (17) yields the piece-wise linear function:

$$RIF\left(y;S\left(p_1,p_2\right)\right)=\begin{cases}\frac{(1-p_1)q_1-(1-p_2)q_2}{\mu}+S\left(p_1,p_2\right)\cdot\left(1-\frac{y}{\mu}\right) & if\ y<q_1\\ \frac{y-(1-p_2)q_2-p_1q_1}{\mu}+S\left(p_1,p_2\right)\cdot\left(1-\frac{y}{\mu}\right) & if\ q_1\leqslant y<q_2,\\ \frac{p_2q_2-p_1q_1}{\mu}+S\left(p_1,p_2\right)\cdot\left(1-\frac{y}{\mu}\right) & if\ y\geqslant q_2\end{cases} \quad (23)$$

where $q_1=Q(F;p_1)$ and $q_2=Q(F;p_2)$.[35] For values below the lower bound of the interval of interest, the slope of the RIF, $-(1/\mu)S(p_1,p_2)$ is negative. It turns positive in the middle segment, $(1/\mu)(1-S(p_1,p_2))$ before becoming negative again, $-(1/\mu)S(p_1,p_2)$ in the upper segment. This change of slope reflects two offsetting effects on the numerator, $\int_{q_1}^{q_2} ydF_Y(y)$, and denominator, $\int_{y}^{\infty} ydF_Y(y)$, of the P-share $S(p_1,p_2)$. Recall that the RIF indicates by how much the P-share increases when the distribution of Y moves in the direction of a mass point at $Y=y$. A higher value of y below the lower bound (q_1) increases the denominator

33 Gradin (2016) finds similar effects for the income distribution in Spain and Germany in 2012.

34 Consider a small transfer Δ from a donor with wealth y' to a recipient with wealth y, where $y<y'$. Making a transfer Δ to a household with wealth y amounts to replacing it with a household with wealth $y+\Delta$. The impact on the Gini is $\Delta Gini=(1/N)(RIF(y+\Delta;G)-RIF(y;G))$. For a small Δ we have $\Delta Gini\approx(\frac{1}{N})\Delta\frac{dRIF(y;G)}{dy}=(\frac{1}{N})\Delta(\frac{2}{\mu})[F(y)-\frac{1+G}{2}]$. Adding the impact on the Gini for the two households (wealth change of -Δ and Δ for the donor and recipient, respectively), we get a total effect of $(\frac{1}{N})\Delta(\frac{2}{\mu})[F(y')-F(y)]$.

35 See footnote 23.

of the P-share without affecting the numerator. As a result, the P-share decreases. The same phenomenon occurs above the upper bound (q_2). It is only for observations in the middle segment (between q_1 and q_2) that a higher value of y has a larger impact on the numerator relative to the denominator.

A more subtle effect in equation (23) is linked to the fact that adding observations at different points of the distribution changes the value of the quantiles q_1 and q_2, which are the bounds between the three segments. For instance, adding an observation with a value of y just above q_2 increases the mass in the upper part of the distribution. The quantile q_2 (and q_1) has to increase accordingly to ensure that we still have a fraction of observations $1-p_2$ above q_2. Moving the interval $[q_1, q_2]$ up in the distribution has a positive effect on the P-share, as it increases average wealth in the interval $[q_1, q_2]$, holding the average in the whole sample (the denominator in the P-share) constant.

Figure 3b plots the value of the $RIF(y; S(p_1, p_2))$ in 2012 by percentile of net worth for three middle P-shares (P40–80, P80–90, P90–95). One important message from figure 3b is that the P-shares do not depend in a simple way on the value of y because of the factors mentioned above. For example, in the case of P40–80, observations just above the 40th percentile decrease the value of the P-share (the *RIF* is below its mean indicated by P40–80 horizontal line), while observations just above the 80th percentile increase it. In the former case, the dominant effect is that adding an observation at the lower range of the $[q_{40}, q_{80}]$ interval reduces average wealth between the 40th and 80th percentile and has a negative impact on the P-share. In the latter case, the dominant effect is that adding an observation just above the 80th percentile shifts up the $[q_{40}, q_{80}]$ interval, which increases the P-share. As we consider observations higher and higher up in the distribution (e.g., above the 95th percentile), the effect eventually turns negative because of the large increase in the denominator of the P-share.

By analogy with the OB decomposition, we can study the changes in inequality measures by constructing, for example, the counterfactual Gini of net worth that would have prevailed if households' characteristics had remained as in 1999, but were valued as in 2012, using the $RIF(y; G)$ coefficients $\hat{\gamma}^G_{2012} : G^C_{RIF} = \bar{X}_{1999}\hat{\gamma}^G_{2012}$. As shown in table A1, our limited set of covariates includes six family types, four family sizes, the education level (four categories) and age of the head of household, and the five regions.[36]

The results of this decomposition for 1999–2005 are presented in table 4, panel A, along with the results for the 1999–2012 comparison, in panel B. The changes in inequality measures have been multiplied by 100 for ease of display. But going back to table 3, one should remember that many of these changes are small. Thus it is not surprising to see that most changes, except for P0–40, are not statistically significant.[37] There is nevertheless something to learn from the decomposi-

36 The base household in our decomposition is a non-elderly one-person household without children with a head with a high school education living in Ontario.

37 We note that our limited set of covariates do not allow us to account for this decline, which remained largely unexplained in table 4.

TABLE 5
Reweighted decomposition using RIF regressions from 1999 to 2005

Inequality measure:	Gini	P0–40	P40–80	P80–90	P90–95	P95–99	P99–100
Unadjusted change (X 100)	1.957^{*}	-0.579^{***}	−1.092	−0.773	−0.211	0.344	2.311
Composition effects attributable to							
Family types	−0.018	0.098	−0.137	-0.214^{*}	−0.098	0.164	0.188
Family size	−0.107	0.028^{*}	0.089	0.013	−0.030	−0.056	−0.045
Head's age	0.411	−0.137	−0.297	0.056	0.066	0.179	0.133
Head's education level	0.663^{***}	-0.187^{****}	-0.477^{***}	−0.115	−0.085	0.288^{**}	0.576^{**}
Regions	0.032	−0.008	−0.028	0.001	0.006	0.022	0.008
Total composition effect	0.981^{**}	-0.207^{*}	-0.851^{***}	-0.259^{*}	−0.140	0.597^{***}	0.860^{**}
Specification error	−0.157	0.093	0.046	−0.018	−0.106	0.005	−0.020
Wealth structure effects attributable to							
Family types	−3.037	1.085^{**}	1.045	1.738	1.666	1.892	-7.425^{*}
Family size	-5.823^{**}	1.006^{**}	4.209^{*}	3.601^{**}	2.019^{*}	1.254	-12.089^{***}
Head's age	−2.261	−0.764	4.499	−0.602	−0.861	1.137	−3.408
Head's education level	1.001	−0.284	−0.913	0.014	0.074	1.583	−0.474
Regions	0.992	-0.599^{**}	−0.353	0.589	0.866	0.688	−1.191
Constant	11.959^{*}	−1.146	−10.425	-6.406^{*}	−4.232	−5.599	27.808^{**}
Total structural effect	2.832^{***}	-0.702^{***}	-1.938^{*}	-1.065^{*}	−0.468	0.954	3.220^{*}
Reweighting error	−0.051	0.010	0.042	0.015	0.010	−0.008	−0.069

NOTES: The composition effects are the time (t_1–t_{10}) differences in the explanatory variables multiplied by the corresponding t_{10} coefficients, where t_{10} refers to the means and coefficients of t_1 reweighted to look like t_0. The wealth structure effects are the difference between the t_{10} and t_0 coefficients evaluated at the means in t_0. All variables, except age entered as quadratic, are categorical. There are 6 family types, 4 family sizes, 6 education classes and 5 regions. ***$p< 0.01$, **$p< 0.05$, *$p< 0.1$.

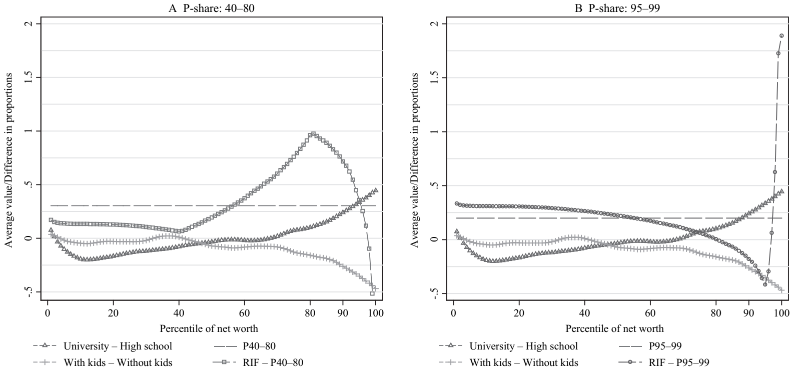

FIGURE 4 Impact of family types and education differences on RIF P-shares (SFS 2012)

tion; it shows which forces have been at play to yield this null change. In particular, changes in family types and size generally account for a decrease in the bottom P-shares and an increase in the top P-shares. This should lead to an increase in overall inequality as measured by the Gini. Changes in head's educational attainment and age account for an increase in bottom P-shares and a decrease in very top P-shares, as shown above. This should lead to a decrease in overall inequality as measured by the Gini index, given the relative importance of those very top P-shares, illustrated in figure 3. Indeed, these effects dominate in the explained part of the decomposition, leaving a total unexplained component that is larger and more significant than the raw changes.

To better understand how changes in covariates affect the P-shares, we illustrate in figure 4 the differences in the proportion of households with a university degree and those with a high school degree and the differences between the proportion of non-elderly non-lone parent households with children and those without children by percentile of net worth in the SFS 2012 (line with triangles). In panel A, the relationship between the education differential and the $RIF(y; S(40, 80))$ appears positively correlated until percentile 80. After that, the two lines go in opposite directions, but the positive relationship dominates. This is consistent with the results in table 4 showing that head's education positively contributes to increases in the P40–80 share.[38] The negative effect of education on the P95–99 share reported in table 4 is less intuitive. Therefore, panel B of figure 4 is helpful in showing the largely negative relationship between the

38 As educational achievement increases over time, this means that the RIF-regression coefficient for university education is positive, thus generating a positive composition effect. Ignoring other covariates, the estimated coefficient on the dummy variable for university education indicates the difference between the average value of the *RIF* for this group relative to the average value for the omitted group (high school graduates). This difference in averages can be expressed as a weighted average of the *RIF* at each percentile of the distribution, using the difference in the proportion of university and high school graduates as weights. Thus, the sign of the RIF-regression coefficient depends on the sign of the correlation between the *RIF* and the difference in proportion. Figure 4 provides a graphical illustration of what drives the sign of this correlation.

education differential and the $RIF(y; S(95,99))$. Observations close, but below the 95th percentile contribute negatively to the P95–99 share, while the education differential continues to rise and contributes to a negative relationship between the two curves. The $RIF(y; S(95,99))$ turns around and increases rapidly after the 95th percentile, but this is not enough to generate an overall positive relationship with the education differential. The overall relationship between the two curves remains negative and contributes to the negative terms found in the decomposition in table 4 for the head's education level.

Opposite effects are found for the differences in family types. Figure 4 illustrates the difference between the proportion of couples with and without children (line with plus signs).[39] This difference is lower than average at the lower end (up to the 35th percentile) and upper end (starting at the 90th percentile) of the distribution of net worth, but it is relatively stable at around 18 percentage points in between. Because it does not increase in the part of the distribution where the $RIF(y; S(40,80))$ strongly increases (between the 40th and the 80th percentile), we expect to see a negative relationship, which is consistent with the results reported in table 4 (panel B). On the other hand, figure 4b shows a more positive relationship between the two curves in the case of the P95–99, a finding confirmed by the results of table 4 (panel B).

We attempt to dig a little deeper into the unexplained component of the traditional OB decomposition by performing a reweighted decomposition. As explained in Fortin et al. (2011), the problematic interpretation of the unexplained part of the OB decomposition arises from the fact that the linearity assumption may not hold. Estimating a reweighted decomposition allows us to assess the importance of this potential departure from the OB assumptions. The idea is to use the average of the reweighted sample, $\bar{X}^C_{2012} = \sum_i \hat{\psi}_X X_{i,2012}$, to construct the counterfactual $\hat{\nu}^C_{RWT} = \bar{X}^C_{2012}\hat{\gamma}^{C,\nu}_{2012}$, where $\hat{\gamma}^{C,\nu}_{2012}$ are the coefficients estimated in the 2012 sample where the covariates are reweighted to look like 1999. Adding and subtracting that second counterfactual term, the reweighted decomposition now comprises four terms as in equation (19). The first term, $\hat{\Delta}^{\nu}_{OBR,X} \equiv (\bar{X}_{2012} - \bar{X}^C_{2012})\hat{\gamma}^{\nu}_{2012}$, corresponds to a "pure" composition effect attributable to changes in characteristics. The second term, $\hat{\Delta}^{\nu}_{OBR,S} \equiv \bar{X}_{1999}(\hat{\gamma}^{C,\nu}_{2012} - \hat{\gamma}^{\nu}_{1999})$, corresponds to the wealth structure effect, that is the difference in the impact of the explanatory variables on the statistic of interest evaluated at the 1999 means. The third term, $\hat{\Delta}^{\nu}_{OBR,RE} \equiv (\bar{X}^C_{2012} - \bar{X}_{1999})\hat{\gamma}^{C,\nu}_{2012}$, is the reweighting error and goes to zero in a large sample when the logit model is well specified. Finally, the fourth term, $\hat{\Delta}^{\nu}_{OBR,SE} \equiv \bar{X}^C_{2012}(\hat{\gamma}^{\nu}_{2012} - \hat{\gamma}^{C,\nu}_{2012})$ is the specification error that corresponds to the difference in the composition effects estimated by reweighting and RIF regressions.

These components are easily obtained by running two OB decompositions on $RIF(y;\nu)$. First, perform the decomposition using the 2012 sample and the 2012

39 Important changes in the proportion of these types of households are reported in table A1. Essentially, the proportion of couples with children decreases over time while those without children increase. However, these changes are not uniform across the net worth distribution.

TABLE 6
Reweighted decomposition using RIF regressions from 1999 to 2012

Inequality measure:	Gini	P0–40	P40–80	P80–90	P90–95	P95–99	P99–100
Unadjusted change (X 100)	0.701	−0.738***	0.467	0.379	0.508*	0.081	−0.697
Composition effects attributable to							
Family types	0.765***	−0.198***	−0.618***	−0.124	0.137**	0.435***	0.369*
Family size	0.156	−0.039**	−0.139	−0.044	0.020	0.124	0.078
Head's age	−1.768***	0.487***	1.394***	−0.118	−0.219***	−0.915***	−0.628***
Head's education level	−1.235***	0.343***	0.989***	−0.042	−0.028	−0.698***	−0.565**
Regions	−0.013	0.001	0.020	−0.008	−0.009	−0.002	−0.001
Total composition effect	−2.095***	0.593***	1.646***	−0.337***	−0.099	−1.056***	−0.747**
Specification error	−0.056	−0.124	0.230	0.367	0.190	−0.204	−0.458
Wealth structure effects attributable to							
Family types	−8.195***	0.525	9.847***	4.164***	1.084	−6.751***	−8.869***
Family size	−6.378***	0.401	7.573***	2.627***	0.598	−3.495***	−7.704***
Head's age	12.079*	−6.868***	−4.987	4.072	−0.006	7.756	0.032
Head's education level	−1.860*	0.391*	1.227	0.949	−0.083	−0.660	−1.824
Regions	−0.954	0.250	0.664	−0.007	0.626	0.095	−1.628
Constant	8.240	4.060***	−15.769*	−11.425**	−1.781	4.408	20.506*
Total structural effect	2.932***	−1.241***	−1.444**	0.380	0.438	1.353**	0.515
Reweighting error	−0.080	0.034	0.036	−0.030	−0.021	−0.012	−0.006

NOTES: The composition effects are the time (t_1–t_{10}) differences in the explanatory variables multiplied by the corresponding t_{10} coefficients, where t_{10} refers to the means and coefficients of t_1 reweighted to look like t_0. The wealth structure effects are the difference between the t_{10} and t_0 coefficients evaluated at the means in t_0. All variables, except age entered as quadratic, are categorical. There are 6 family types, 4 family sizes, 6 education classes and 5 regions. ***p< 0.01, **p< 0.05, *p< 0.1.

sample reweighted to be as 1999 to get the pure composition effect, $\hat{\Delta}^{\nu}_{OBR,X}$. The total unexplained in this decomposition corresponds to the specification error $\hat{\Delta}^{\nu}_{OBR,SE}$ and allows one to assess the importance of departures from the linearity assumption. Second, perform the decomposition using the 1999 sample and the 2012 sample reweighted to be as 1999 to get the pure wealth structure effect $\hat{\Delta}^{\nu}_{OBR,S}$ in the unexplained part of the decomposition. The reweighting error $\hat{\Delta}^{\nu}_{OBR,RE}$ is given by the total explained effect in this decomposition. It provides an easy way of assessing the quality of the reweighting.

Tables 5 and 6 report the results of the decompositions that reweight the 2005 sample to look like 1999 and the 2012 sample to look like 1999, respectively, as in table 3. The first observation is that the composition effects from these decompositions are very similar to those of table 4 and the specification and reweighting errors are small and not statistically significant, with a small exception for the P90–95 in table 5. This is reassuring about the validity of the traditional OB decomposition.

The next question is whether we can attribute changes in wealth inequality to the changes in the wealth structure, i.e., the changes in the returns to the covariates. The analogy with income inequality is that increases in the returns to education are understood to have played a leading role in increases in earning inequality (Lemieux 2006). For the Gini index, no explanatory variables seemed to have played a similar role. For the P0–40, the only P-share where we have significant changes over time, the impact of changing age wealth effects is found to be significantly negative in both comparisons and is of sufficient size to yield a negative wealth structure effect. Further investigations may want to explore which components of wealth, likely debt in this case, have changed over time for these poorer households.

5. Conclusion

This article provides an overview of some of the main issues linked to the analysis of changes in wealth inequality and uses recent advances in decomposition methods to further our understanding of these changes. One novel contribution is to analyze how changes in the distribution of socio-economic factors such as age, education and family type have contributed to the changes in the concentration of wealth at different points of the distribution. In particular, we show how recentred influence function (RIF) regressions can be used to study the determinants of percentile shares, including the famous top 1% share. Notwithstanding the shortcomings of survey data on wealth, the findings reveal the compensating role of family formation and increases in human capital in mitigating increases in wealth inequality in Canada.

As noted in many studies (e.g., Davies 1993, Morissette and Zhang 2007), the precise measurement of wealth in the upper tail is challenging because of both sampling and non-sampling error. Random sampling would only occasionally select billionaire families to interview, for example, so they are unlikely to be included even without non-sampling issues. But two forms of non-sampling error also affect the upper tail—differentially low response rates among wealthier families and under-reporting of their assets. The SFS places the wealth share of the top 1% in Canada at the implausibly low level of 12.5% in 2012. Much higher estimates are obtained taking evidence from additional, independent sources into account.[40] However, the survey data used here comprise a set of covariates that include educational attainment, which is not the case in alternative data sources such as income tax records. Exploiting these advantages while acknowledging the shortcomings of survey data, we have focused on measurement and decomposition methods that are robust in the face of difficulties in measuring the extremes of the distribution.

Appendix

TABLE A1
Descriptives statistics – Proportions and means

Variables	1984	1999	2005	2012
Family type:				
Unattached non-elderly	21.53	23.55	25.48	25.29
Lone parent families	5.05	4.85	4.46	3.84
Other families with children non-elderly	31.82	26.78	23.10	20.67
Families without children non-elderly	24.75	26.56	28.96	29.49
Older unattached individuals	8.05	8.61	8.20	9.31
Older head families	8.80	9.66	9.79	11.41
Family size:				
1 person	29.58	32.16	33.68	34.59
2 persons	26.08	27.97	29.63	31.36
3 persons	16.30	16.17	15.46	12.23
4 persons	17.18	15.42	14.21	13.21
5 or more persons	10.87	8.29	7.02	8.61
Head's age	45.32	46.94	47.53	49.45
Head's education level:				
Less than high school	38.96	26.86	20.96	16.20
High school diploma	36.34	23.27	26.20	25.70
Postsecondary certificate or diploma	11.47	28.25	27.89	28.32
University degree or certificate	13.23	21.26	24.58	29.26
Not stated	0	0.35	0.36	0.52
Regions				
Atlantic provinces	0.079	0.076	0.074	0.071
Quebec	0.257	0.255	0.252	0.246
Ontario	0.361	0.367	0.372	0.374
Prairie provinces	0.179	0.164	0.166	0.174
British Columbia	0.125	0.138	0.137	0.136
No. of observations	14,029	15,933	5,267	12,003

NOTE: Sample weighted.

40 Using the Forbes billionaire data to adjust the upper tail and aligning the SFS data with national balance sheet totals, Davies et al. (2016, table 6–5) estimates the share of the top 1% in household wealth in Canada at 25.6%. See also Davies et al. (2017).

TABLE A2
Percentile shares of total net worth (with pensions)

Year	Canada	Atlantic provinces	Quebec	Ontario	Prairie provinces	British Columbia
1999						
P0–40	0.028	0.036	0.026	0.030	0.034	0.020
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
P40–80	0.300	0.314	0.281	0.315	0.306	0.309
	(0.004)	(0.008)	(0.009)	(0.007)	(0.007)	(0.011)
P80–90	0.193	0.196	0.186	0.194	0.188	0.194
	(0.002)	(0.004)	(0.005)	(0.003)	(0.004)	(0.006)
P90–95	0.148	0.146	0.145	0.151	0.143	0.145
	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
P95–99	0.199	0.197	0.205	0.195	0.196	0.187
	(0.003)	(0.007)	(0.011)	(0.005)	(0.006)	(0.006)
P99–100	0.132	0.111	0.157	0.114	0.134	0.145
	(0.005)	(0.008)	(0.010)	(0.007)	(0.011)	(0.017)
2005						
P0–40	0.022	0.023	0.019	0.030	0.021	0.015
	(0.002)	(0.010)	(0.003)	(0.004)	(0.004)	(0.005)
P40–80	0.289	0.288	0.285	0.308	0.282	0.282
	(0.008)	(0.020)	(0.014)	(0.014)	(0.023)	(0.019)
P80–90	0.185	0.194	0.196	0.182	0.179	0.183
	(0.005)	(0.013)	(0.008)	(0.008)	(0.013)	(0.010)
P90–95	0.146	0.150	0.161	0.140	0.138	0.151
	(0.004)	(0.010)	(0.007)	(0.006)	(0.009)	(0.008)
P95–99	0.203	0.190	0.211	0.188	0.201	0.235
	(0.006)	(0.012)	(0.009)	(0.011)	(0.012)	(0.031)
P99–100	0.155	0.154	0.128	0.151	0.178	0.134
	(0.013)	(0.037)	(0.014)	(0.024)	(0.041)	(0.011)
2012						
P0–40	0.020	0.031	0.020	0.020	0.026	0.017
	(0.001)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)
P40–80	0.304	0.314	0.293	0.311	0.304	0.330
	(0.005)	(0.008)	(0.010)	(0.009)	(0.009)	(0.011)
P80–90	0.197	0.202	0.191	0.196	0.191	0.202
	(0.002)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)
P90–95	0.153	0.165	0.154	0.150	0.151	0.149
	(0.002)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)
P95–99	0.200	0.201	0.205	0.196	0.206	0.198
	(0.004)	(0.007)	(0.006)	(0.006)	(0.006)	(0.014)
P99–100	0.125	0.087	0.137	0.127	0.123	0.103
	(0.005)	(0.005)	(0.012)	(0.009)	(0.011)	(0.006)

NOTE: Percentile shares and standard errors computed with Jann (2016) P-share Stata routine.

References

Aliaga, R., and S. Montoya (1999) "GINIDESC: Stata module to compute Gini index with within- and between-group inequality decomposition," Statistical Software Components S372901, Boston College Department of Economics

Atkinson A. B., and A. J. Harrison (1978) *The Distribution of Personal Wealth in Britain*. Cambridge: Cambridge University Press

Beach, C. M., and R. Davidson (1983) "Distribution-free statistical inference with Lorenz curves and income shares," *The Review of Economic Studies*, 50(4), 723–35

Beach, C. M., and J. Richmond (1985) "Joint confidence intervals for income shares and Lorenz curves," *International Economic Review* 26(6), 439–50
Bishop, J. A., S. Chakraborti, and P. D. Thistle (1989) "Relative deprivation and economic welfare: A statistical investigation with Gini-based welfare indices," *Scandinavian Journal of Economics* 93(3), 421–37
Bishop, J. A., J. P. Formby, and W. J. Smith (1991) "International comparisons of income inequality: Tests for Lorenz dominance across nine countries," *Economica* 58(232), 461–77
Bonesmo Fredriksen, K. (2012) "Less income inequality and more growth – Are they compatible? Part 6. The distribution of wealth," OECD Economics Department working paper no. 929
Bricker, J., A. Henriques, J. Krimmel, and J. Sabelhaus (2016) "Measuring income and wealth at the top using administrative and survey data," *Brookings Papers on Economic Activity* 2016(1), 261–331
Brzozowski, M., M. Gervais, P. Klein, and M. Suzuki (2010) "Consumption, income and wealth inequality in Canada," *Review of Economic Dynamics* 13(1), 52–75
Carpentier, J.-F., J. Olivera, and P. Van Kerm (2017) "Macroprudential policy and household wealth inequality," LISER Working Paper Series, no. 2017-09
Choe, C., and P. Van Kerm (2014) "Foreign workers and the wage distribution: Where do they fit in?," LISER technical report no. 2014-02
Cowell, F. A. (2011) *Measuring Inequality*, 3rd edition. Oxford: Oxford University Press
Cowell, F. A., and F. Mehta (1982) "The estimation and interpolation of inequality measures," *The Review of Economic Studies* 49(2), 273–90
Cowell, F. A., B. Nolan, J. Olivera, and P. Van Kerm (2017) "Wealth, top incomes and inequality," In K. Hamilton and C. Hepburn, eds., *Wealth: Economics and Policy*. Oxford: Oxford University Press, forthcoming
Cowell, F. A., and P. Van Kerm (2015) "Wealth inequality: A survey," *Journal of Economic Surveys* 29(4), 671–710
Cowell, F. A., and M.-P. Victoria-Feser (1996) "Robustness properties of inequality measures," *Econometrica* 64(1), 77–101
——— (2008) "Modelling Lorenz curves: Robust and semi-parametric issues." In D. Chotikapanich, ed., *Modeling Income Distributions and Lorenz Curves*, pp. 241–53. New York: Springer
Dalton, H. (1920) "The measurement of inequality of incomes," *The Economic Journal* 30(119), 348–61
Davidson, R., and J.-Y. Duclos (2000) "Statistical inference for stochastic dominance and for the measurement of poverty and inequality," *Econometrica* 68(6), 1435–64
Davies, J. B. (1979) "On the size distribution of wealth in Canada," *Review of Income and Wealth* 25(3), 237–59
——— (1993) "The distribution of wealth in Canada." In E. N. Wolff, ed., *Research on Economic Inequality*, Vol. 4, pp. 159–80. Greenwich, CT: JAI Press
Davies, J. B., and L. Di Matteo (2017) "Filling the gap: Examining Canadian wealth inequality over the long run," presented at the 51st Annual Conference of the Canadian Economics Association
Davies, J. B., and M. Hoy (1995) "Making inequality comparisons when Lorenz curves intersect," *The American Economic Review* 85(4), 980–86
Davies, J. B., R. Lluberas, and A. F. Shorrocks (2016) *Global Wealth Databook 2016*, Credit Suisse Research Institute, Zurich. Available at publications.credit-suisse.com/tasks/render/file/index.cfm?fileid=AD6F2B43-B17B-345E-E20A1A254A3E24A5
——— (2017) "Estimating the level and distribution of global wealth, 2000–2014," *Review of Income and Wealth* 63(4), 731–59

Davies, J. B., and A. F. Shorrocks (2000) “The distribution of wealth.” In A. B. Atkinson and F. Bourguignon, eds., *Handbook of Income Distribution*, pp. 605–75, Vol. I. North-Holland Elsevier

Di Matteo, L. (1997) “The determinants of wealth and asset holding in nineteenth century Canada: Evidence from microdata,” *Journal of Economic History* 57(4), 907–34

——— (2016) “Wealth distribution and the Canadian middle class: Historical evidence and policy implications,” *Canadian Public Policy* 42(2), 132–51

DiNardo, J., N. M. Fortin, and T. Lemieux (1996) “Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach,” *Econometrica* 64(5), 1001–44

Essama-Nssah, B., and P. J. Lambert (2012) “Influence functions for policy impact analysis.” In J. A. Bishop and R. Salas, eds., *Inequality, mobility and segregation: Essays in honor of Jacques Silber*, Chap. 6, pp. 135–59. Bingley, UK: Emerald Group Publishing Limited

Firpo, S., N. M. Fortin, and T. Lemieux (2007) “Decomposing wage distributions using recentered influence function regressions,” unpublished manuscript, University of British Columbia

——— (2009) “Unconditional quantile regressions,” *Econometrica* 77(3), 953–73

Fortin, N. M. and T. Lemieux (2015) “Changes in wage inequality in Canada: An interprovincial perspective,” *Canadian Journal of Economics* 48(2), 682–713

Fortin, N. M., T. Lemieux, and S. Firpo (2011) “Decomposition methods.” In O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 4A, Chap. 1, pp. 1–102. Amsterdam: North-Holland

Fortin, N. M., and T. Schirle (2006) “Gender dimensions of changes in earnings inequality in Canada.” In D. Green and J. Kesselman, eds., *Dimensions of Inequality in Canada*, pp. 307–46, Vancouver, BC: UBC Press

Gâteaux, R. (1913) “Sur les fonctionnelles continues et les fonctionnelles analytiques,” *Comptes rendus de l’Académie des sciences – Série I – Mathématique* 157, 325–27

Gini, C. (1914) “On the measurement and variability of characters,” *METRON – International Journal of Statistics* 63(1), 3–38

Gradín, C. (2016) “Why is income inequality so high in Spain?” In L. Cappellari, S. W. Polachek, and K. Tatsiramos, eds., *Income Inequality Around the World (Research in Labor Economics, Volume 44)*, pp. 109–77. Cheltenham, UK: Emerald Group Publishing Limited

Green, D. A., W. C. Riddell, and F. St-Hilaire (2016) *The Art of the State Volume V. Income Inequality: The Canadian Story*, D. A. Green, W. C. Riddell, and F. St-Hilaire, eds. Montreal: IRPP

Hampel, F. R. (1974) “The influence curve and its role in robust estimation,” *Journal of the American Statistical Association* 69(346), 383–93

Harrison, A. (1980) “The distribution of wealth in Canada, the UK and US.” In *Reflections on Canadian Incomes*. Ottawa: Economic Council of Canada

Jann, B. (2016) “Assessing inequality using percentile shares,” *The Stata Journal* 16(2), 264–300

Jenkins, S. (1991) “The Measurement of Income Inequality.” In L. Osberg, ed., *Economic Inequality and Poverty: International Perspectives*. Armonk, NY, and London: M.E. Sharpe

Kline, P. (2011) “Oaxaca-Blinder as a reweighting estimator,” *The American Economic Review* 101(3), 532–37

Kolm, S.-C. (1976a) “Unequal inequalities I,” *Journal of Economic Theory* 12(3), 416–42

——— (1976b) “Unequal inequalities II,” *Journal of Economic Theory* 13(1), 82–111

Kopczuk, W. (2015) "What do we know about the evolution of top wealth shares in the United States?," *Journal of Economic Perspectives* 29(1), 47–66
Lemieux, T. (2006) "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill," *The American Economic Review* 96(3), 461–98
Milligan, K. (2005) "Life-cycle asset accumulation and allocation in Canada," *Canadian Journal of Economics* 38(3), 1057–106
Monti, A. C. (1991) "The study of the Gini concentration ratio by means of the influence function," *Statistica* 51(4), 561–80
Morissette, R., and X. Zhang (2007) "Revisiting wealth inequality," *Perspectives on Labour and Income* 19(1), 1–6
Morissette, R., X. Zhang, and M. Drolet (2006) "The evolution of wealth inequality in Canada, 1984–1999." In E. N. Wolff, ed., *International Perspectives on Household Wealth*, Chapter 5. Cheltenham, UK: Edward Elgar Publishing
Oaxaca, R., and M. Ransom (1999) "Identification in detailed wage decompositions," *The Review of Economics and Statistics* 81(1), 154–57
OECD (2013) *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. OECD Publishing
Oja, G. (1983) "The distribution of wealth in Canada," *Review of Income and Wealth* 29(2), 161–73
Osberg, L., and F. Siddiq (1988) "The inequality of wealth in Britain's North American colonies: The importance of the relatively poor," *Review of Income and Wealth* 34(2), 143–63
Pendakur, K. (1998) "Changes in Canadian family income and family consumption inequality between 1978 and 1992," *Review of Income and Wealth* 44(2), 259–82
Piketty, T., and E. Saez (2003) "Income Inequality in the United States, 1913–1998," *The Quarterly Journal of Economics* 118(1), 1–41
——— (2013) "Top incomes and the great recession: Recent evolutions and policy implications," *IMF Economic Review* 61(3), 456–78
Podoluk, J. R. (1974) "Measurement of the distribution of wealth in Canada," *Review of Income and Wealth* 20, 203–16
Saez, E., and M. R. Veall (2005) "The evolution of high incomes in Northern America: Lessons from Canadian evidence," *The American Economic Review* 95(3), 831–49
Saez, E., and G. Zucman (2016) "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *The Quarterly Journal of Economics* 131(2), 519–78
Sen, A. (1973) *On economic inequality*, Oxford University Press (Clarendon)
Shorrocks, A. F. (1982) "Inequality decomposition by factor components," *Econometrica* 50(1), 193–211
Shorrocks, A. F., and J. E. Foster (1987) "Transfer sensitive inequality measures," *The Review of Economic Studies* 54(3), 485–97
Siddiq, F. K., and C. M. Beach (1995) "Characterizing life-cycle wealth distributions using statistical inference and dominance criteria," *Empirical Economics* 20(4), 551–75
Uppal, S., and S. LaRochelle-Côté (2015) *Changes in wealth across the income distribution, 1999 to 2012*. Statistics Canada
Van Kerm, P. (2009) Generalized Gini and concentration coefficients (with factor decomposition) in Stata, CEPS/INSTEAD working paper
Veall, M. R. (2012) "Top income shares in Canada: Recent trends and policy implications," *Canadian Journal of Economics* 45(4), 1247–72