Research Paper

Analytical Studies Branch research paper series

Rising Income Inequality in the 1990s: An Exploration of Three Data Sources

By Marc Frenette, David Green and Garnett Picot

Business and Labour Market Analysis Division
24-F, R.H. Coats Building, Ottawa, K1A 0T6

Telephone: 1 800 263-1136

This paper represents the views of the authors and does not necessarily reflect the opinions of Statistics Canada.

Statistics Canada   Statistique Canada

Canada

Rising Income Inequality in the 1990s: An Exploration of Three Data Sources

by

Marc Frenette,* David Green** and Garnett Picot *

Business and Labour Market Analysis
24-F, R.H. Coats Building, Ottawa, K1A 0T6
*Statistics Canada
**UBC, Department of Economics

How to obtain more information :
National inquiries line: 1 800 263-1136
E-Mail inquiries: infostats@statcan.ca

December 2004

Aussi disponible en français

Table of Contents

Abstract

This study investigates trends in family income inequality over the last two decades, with particular attention paid to the largely undocumented 1990s. This study attempts to paint a picture of inequality trends over the 1990s from three distinct data sources: survey data, tax data, and Census data. While all three sources point to an increase in income inequality over the decade, there are differences regarding the magnitude and timing of the increase. Furthermore, there are large differences in the extent of inequality across the three data sources at any point in time, which may result from differences in population coverage.

1. Introduction

The levels and trends in inequality in both pre- and post-tax and transfer income are key inputs into any discussions of equity in a society, how that society's economy functions and how both of these have been changing over time. The comparison of pre- and post-tax and transfer income inequality is particularly useful in considering the role of government redistribution in economic well-being. In Canada's case, several authors document that while earnings inequality rose throughout the 1980s and early 1990s in Canada, after-tax family income inequality changed very little (e.g., Beach and Slotsve, 1994, Morissette, Myles, and Picot, 1994, Jantti, 1997, Rashid, 1998, and Wolfson and Murphy, 1998). Canada's tax and transfer system thus played a strong role in offsetting market driven inequality increases. Less is known about the patterns in the 1990s. The goal of this paper is to expand our knowledge of inequality patterns in the 1990s and compare them with the better known patterns from the 1980s.

Beginning in the mid-1990s, the Canadian tax and transfer system underwent substantial change, including an overhaul of the Unemployment Insurance system (now the Employment Insurance system), substantial reforms to provincial social assistance programs (such as declining benefits and tighter eligibility rules), the introduction of the Child Tax Benefit, and the (modest) beginning of a round of tax cuts at the end of the 1990s. The late 1990s also saw a very strong economic recovery, with prime-aged unemployment and employment rates reaching levels by 2000 not observed since at least 1981. These changes raise several questions. Did earnings inequality fall significantly during the later part of the cycle, given the very strong economic growth, and as witnessed late in the cycle in the U.S. (e.g., Ilg and Haugen, 2000)? This would have contributed to declining family income inequality. Or was the earnings growth primarily concentrated at the top of the earnings distribution, resulting in rising inequality? To the extent that earnings inequality increased, did the tax and transfer system continue to offset the increases, as it did through the 1980s?

Unlike similar analyses for earlier periods, the comparability of data over time is a major issue for the 1990s. Ensuring that the data are appropriate for the analysis is a major component of the research for the decade. Hence, to answer these and related questions, we use three different data sources. We start with survey data covering the period 1980 to 2000, using a combination of the Survey of Consumer Finances (SCF) up to 1996 and the Survey of Labour and Income Dynamics (SLID) (which replaced the SCF as Statistics Canada's main survey) beginning in 1996. We complement our findings with results from taxation data (the T1 Family Files, or T1FF). The T1FF provides a consistent series of income estimates from 1992 to 2000 (thus covering the "seam" created by the transition from the SCF to SLID), and given its large size, does not suffer from the variability of estimates often associated with survey data. Another advantage of the T1FF is the relatively high coverage rate. Finally, we use Census data for the years 1980, 1985, 1990, 1995 and 2000. In particular, we focus on 1980, 1990 and 2000, years that are in roughly comparable positions of the business cycle. This provides another consistent source that covers the last two business cycles. Coverage at the bottom end of the distribution may be better in the Census than in the surveys, because reporting is compulsory, and response rates tend to be much higher than in the income surveys. The disadvantage of the Census is that it lacks information on income taxes paid and can thus only provide post-transfer, pre-tax data. As a result, effects of changes in the tax system on family income inequality are missed.

From this description, one can see there is no perfect data source for the 1990s. The survey data has a discontinuity in 1996, the taxation data do not cover the complete 1990s cycle, let alone allowing comparisons with the 1980s cycle, and the Census data are pre-tax and only occur every five years. We use all three sources to develop a picture of trends in family income inequality during the 1990s. Unfortunately, the survey and tax data provide quite different pictures of the level and trends in family income inequality. The survey data suggests a far more compressed income distribution, and points to an increase in pre-tax and transfer family income inequality generated from larger increases in incomes at the top than at the bottom of the distribution. According to that same source, the tax and transfer system largely, though not completely, offset the increase in market income inequality, resulting in a small rise in after tax and transfer income inequality over the 1990s. Interestingly, the offsetting effects from the tax and transfer system appear to occur entirely in the first half of the decade; adding in tax and transfer effects actually leads to increases in inequality over the second half of the 1990s.

In contrast, the tax data points to much larger increases in pre-tax and transfer inequality driven both by rises in market income at the top of the distribution and a lack of improvement among families at the bottom. Further, the inequality compounding effects from the tax and transfer system in the second half of the decade are just as strong as those observed in the survey data. Thus, according to the tax data, and in contrast to the survey data, the 1990's witnessed very substantial increases in after tax and transfer income inequality. Moreover, the level of inequality is much higher in the tax data than the survey data in each year, due mainly to much lower earnings at the bottom of the distribution.

Given the very different inequality patterns observed in the two types of data, it is important to consider the strengths and weaknesses of each source. To do so, we turn to Census data, which has coverage rates comparable to the tax data but a survey methodology that is closer to the Survey of Consumer Finance. While the Census pre-tax and transfer income inequality levels more closely resemble those observed in tax data, the trends diverge substantially from those observed in both survey and tax data.

Census and survey data both point to a similar increase in income inequality over the 1990s, although the pathways are very different. Survey data suggests that inequality rose moderately in both the first and second halves of the decade. In the first half, this was largely due to a moderate decline in income at the bottom of the distribution, while in the second half it was mainly due to a large improvement at the top of the distribution (incomes improved moderately at the bottom at this time). According to Census data, inequality rose substantially in the early 1990s, due mainly to a large decline in the bottom of the distribution. In the latter half of the decade, inequality remained fairly stable, since there were substantial improvements at the top and bottom of the distribution. In essence, the main difference between Census and survey data lies in what both sources suggest happened at the bottom of the distribution. According to survey data, income fell moderately in the early 1990s, and then rose moderately in the latter half of the decade. Census also suggests a decline followed by an increase, but the fluctuations were much larger. It is possible, although difficult to ascertain with certainty, that this difference is the result of the inherent difficulty in capturing the bottom end of the distribution in non-mandatory surveys.

The income distributions appear to be more similar in Census and tax data, although the trends at the bottom diverge just as much as they do when we look at Census and survey data. In the late 1990s (the only period in which we can compare Census and tax data), both sources point to a large increase at the top (as does survey data), but tax data points to a moderate decline at the bottom. Recall that Census points to a substantial improvement at the bottom. As a result, income inequality rose considerably according to tax data, yet remained fairly stable according to Census data in the late 1990s.

In summary, while all three data sources point to an increase in income inequality over the 1990s, there are substantial differences in the levels of inequality, as well as in the extent and timing of the trends. The main objective of this paper is to document these differences.

The paper proceeds in six sections. In the second section, we provide details on the SCF/SLID and tax data we use for the majority of our analysis. The third section contains a description of our approach in measuring inequality movements. The fourth section provides a brief discussion of cyclical elements in the economy that may affect our results. The fifth section constitutes most of the paper and contains our results. The sixth section contains conclusions.

2. Details of the data

Income inequality studies generally rely on survey data, and the traditional source has been the Survey of Consumer Finances (SCF), which is available up to 1997.[1] The SCF is an annual cross-sectional survey that targets all households in Canada, with the exception of those living in the territories, institutions, or on native reserves. The exceptions account for less than 3% of the Canadian population. The sample of roughly 35,000 households is selected as a supplement to the April Labour Force Survey (LFS). The primary objective of the survey is to provide income estimates by detailed sources, and responses by proxy are allowed (i.e., one person in the household may answer questions about other members).

The SCF was Statistics Canada's official source of income estimates until 1996, when the Survey of Labour and Income Dynamics (SLID) began to be used for this purpose. The SCF and SLID share many common features. Both are annual household surveys that use the LFS as a sampling frame, and have the same target population. Although SLID can be (and is) used for cross-sectional estimates, it is also designed for longitudinal analysis. Panels are interviewed for up to 6 years, with new (and overlapping) panels introduced every 3 years. New panels were introduced in 1993, 1996, and 1999. Each year, a panel is interviewed in January (mainly to collect labour information) and in May (to collect income information). For the income interview, respondents have the option of allowing Statistics Canada to link to their T1 tax files (if possible) in order to collect their income information, thus eliminating the need for an income questionnaire. More than 80% of respondents provide Statistics Canada with the permission to attempt this match, and the income of about 70% of all respondents is obtained from the tax files in this way. Another difference between the two surveys relates to the formation of families. In the SCF, families are derived with respect to the "head" of the family, which gives priority to the husband. In SLID, it is with respect to the "major

---

1. The following information on the SCF and SLID draws heavily from Statistics Canada (2002).

income recipient". Sample surveys must of course deal with sampling error (sampling variability in the estimates) and non-sampling variability, including potential non-response bias.

Response rates in the SCF generally hover around 80%, while in SLID, it is slightly higher (80% to 85%). The cross-sectional sample weights in SCF and SLID are adjusted for non-response to ensure accurate population counts for different province-age-sex groups, as well as by household and family size. The weights are *not* adjusted for income related response bias—i.e., potentially differential response rates at different parts of the income distribution. If important changes are occurring at the top or bottom of the income distribution, the data being used to track changes in inequality must have a good coverage rate in these parts of the distribution, and these are typically the segments that present surveys with the most difficult task regarding response rates and coverage.

Official income estimates used the SCF up to and including 1995; since then, SLID has become the official Statistics Canada source. Since two data sources are used to create one time-series, there is the potential for the "seam" problem: a discontinuity in the series that is related to the change in data sources. In particular, the (partial) use of tax data in SLID may help reduce response errors (relative to SCF).

Coverage of particular income components is also an issue in surveys. Generally speaking, coverage of the earnings income component has been good, but that of transfer components such as employment insurance benefits and social assistance receipts less so (Kapsalis, 2001). Changes in these coverage rates can influence the outcomes of analyses such as the one conducted here, particularly for the bottom end of the distribution. In particular, the movement towards partial tax data in SLID (compared to full survey data in SCF) may have changed the coverage of transfer income.

We complement our survey findings with taxation data. The Canada Customs and Revenue Agency (CCRA) collects personal income tax forms (the T1s) from all tax filers in Canada. From the T1 file, Statistics Canada creates the T1FF (the T1 Family File), which attempts to reconstruct census families by imputing the presence of non-filing children and spouses. The T1FF is a census of all individuals who file taxes, or whose social insurance number (SIN) appears on another family members' tax file. Non-filing children (and those with a SIN that does not appear on another family members' tax file) are imputed. Unlike survey data, T1FF may contain records of people living in the territories, institutions, or on native reserves. At best, we can only identify people living in the territories with a reasonable amount of certainty; to be as consistent as possible, we have deleted those records.

The T1FF has been particularly well-suited for estimates of income at the lower end of the distribution since 1992 given the large number of incentives for lower income families to file taxes.[2] On the other hand, the creation of those incentives implies that tax data are unlikely to be based on a

---

2.   The Child Tax Credit was available as a non-refundable tax credit for families with children prior to 1993, which did nothing for families with earnings below the taxable threshold. The Child Tax Credit was replaced by the Child Tax Benefit in 1993, and this provided an incentive for families with children to file taxes even if they had no earnings. Finally, tax filers could apply for the Goods and Services Tax (GST) Credit beginning in 1989, although no payments were made until December 1990.

consistent sample of observations before versus after 1992. Indeed, given changes from the Child Tax Credit to the Child Tax Benefit in 1993, the 1992 data may also not be comparable to subsequent years.[3] This, plus the fact that transfer income was not reported at all before 1989 and not reported consistently before 1992 implies that we cannot use tax data before 1992 or, possibly, 1993. We present results including both 1992 and 1993 tax data but emphasize 1993 as a starting point in establishing trends.

Tax data offer three distinct advantages over SCF/SLID. First, the population coverage rate has been over 95% since 1992 in T1FF, which fares quite favourably relative to survey data (generally around 80%). Secondly, T1FF gives us a consistent time-series over the SCF/SLID seam (including fairly consistent estimates of transfer income). And finally, sampling variability is minimized in tax data given the much larger sample sizes.

One drawback of using tax data is the lack of consistent income estimates prior to 1992, especially at the bottom of the distribution. Another issue is that census families, rather than the more commonly used economic families, are created in the tax data. A census family is defined as a now-married couple (with or without children), a common-law couple (same) or a lone-parent with a child who is under the age of 25 and who does not have his or her own spouse or child living in the household. In our analysis, unattached individuals are also considered census families so that the complete population is covered. An economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law or adoption. Once again, in this analysis unattached individuals are included when the economic family concept is employed. By definition, all persons who are members of a census family are also members of an economic family.[4] Economic families are generally the preferred family unit of analysis for studies of economic well-being. However, to facilitate comparisons between survey and tax data, we often turn to the census family concept. The move from economic to census families has virtually no effect on our results; the same general conclusions hold, no matter which concept is used.

3. Approach

In this study, the focus is on after-tax family income (except with the census data, which do not have information on taxes paid). This is total income (including transfers) minus taxes paid.[5] Total income consists of market income plus transfer income. Market income is comprised of employment earnings plus other (non-transfer) income (such as investment and pension income).

---

3. The overall population coverage rate in T1FF (relative to Census) went from 95.12% in 1992 to 96.15% in 1993, possibly because the Child Tax Benefit gave some families an incentive to file taxes. Its coverage rate generally hovered around 96% in subsequent years, with no sudden substantial changes.

4 . Examples of the broader concept of economic family include the following: two co-resident census families who are related to one another are considered one economic family, and two co-resident siblings who are not members of a census family are considered an economic family.

5 . Note that taxes paid include only the personal income tax (federal and provincial). We do not include EI and CPP/QPP premiums.

And finally, employment earnings consist of paid earnings plus net self-employment income and other employment income. The following equations summarize these definitions:

after-tax income = total income – income taxes

market income + transfer income

employment earnings + other market income

paid earnings + net self-employment income + other employment earnings
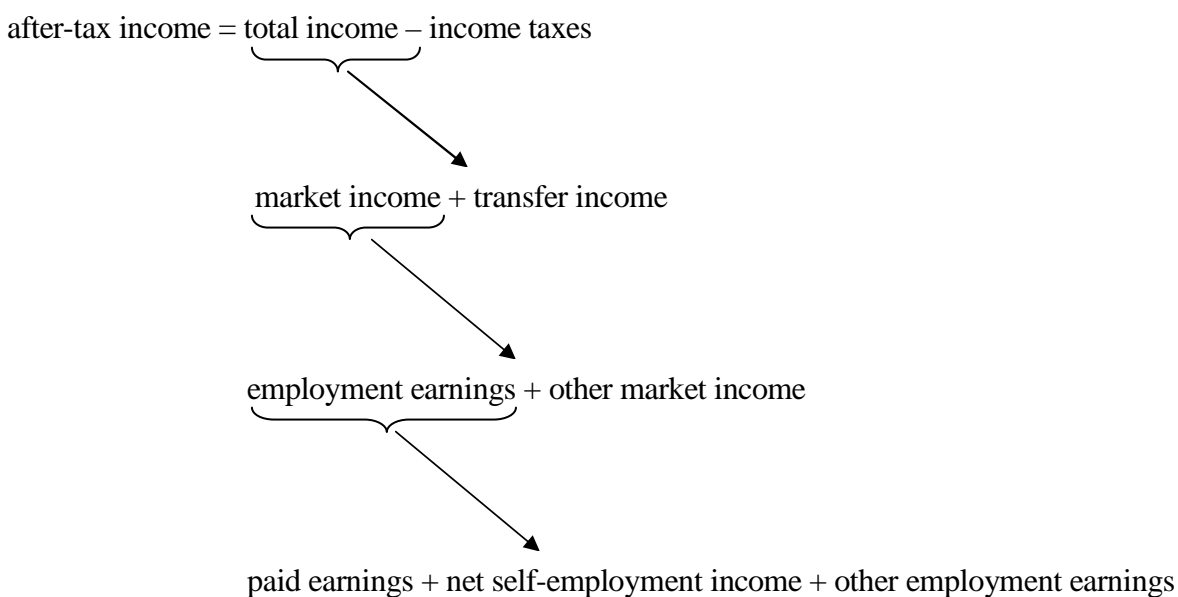
Other market income includes income from investments, (actual, not taxable) dividends, net rental income, pension income, and alimony received; capital gains are not included. In the SCF, net partnership income is included as part of net self-employment income, while in SLID and T1FF, it is included as part of investment income. Note that starting in 1997, new alimony arrangements were no longer tax deductible (for payments made) nor taxable (for payments received).[6] Thus, there may be a discontinuity in the measurement of other income in tax data (T1FF and to a lesser extent SLID). Since only new arrangements are affected, however, the break should be minimal. This was confirmed by re-running all estimates without alimony received. It had no significant effect on the findings.

To measure inequality and its evolution through time, we rely on a series of indices that are sensitive to changes at various points of the income distribution. The Gini coefficient (G) is perhaps the most commonly used inequality index.[7] The Gini coefficient is sensitive to changes in the middle of the distribution, which renders it less than ideal in detecting changes over time when these are driven by events at the top and/or bottom. To provide a more complete picture, we also look at

---

6. Alimony paid is not available in SLID and the Census, and thus, could not be included in our definition of after-tax income. The results were not sensitive to this limitation, however, as we generated similar numbers from T1FF when alimony paid was subtracted from after-tax income.

7. The Gini is based on the Lorenz curve, which is a mapping of the functional relationship between the cumulative percentage of income held by the population and the cumulative percentage of the population. Perfect equality is achieved if the Lorenz curve is a straight line—each member of the population holds an equal share of the total income in the economy. The Gini measures the ratio of the area between the line of perfect equality and the Lorenz curve to the area between line of perfect equality and the segment of lines under perfect inequality (all income is held by one member of the economy).

the exponential measure (Exp),[8] which is bottom sensitive, and the coefficient of variation squared (CV^2), which is top-sensitive. The formula for each appears below:

$$(1) Exp = \sum_{i=1}^{n} p_i \exp(-y_i / \overline{y})$$

$$(2) G = 2\sigma_{y,F(y)} / \overline{y}$$

$$(3) CV^2 = \left(\sigma_y / \overline{y}\right)^2$$

Note that "p_i" denotes the share of the size of family ("i") in the sample of "n" families, "y" denotes income, σ_y denotes the standard deviation of income, and $\sigma_{y,F(y)}$ denotes the covariance between income and its cumulative distribution—F(y). The Exp and CV^2 measures may be heavily influenced by outlier incomes, which may be relatively more prominent in smaller data sets such as SCF and SLID. To reduce this undesirable impact, we have dropped the top and bottom 0.1% of the income distribution in each year for all calculations (about 60 families per year). [9]

For significance tests, we rely on the work of Kovacevic and Binder (1997), who used the 1991 SCF to study variance estimation of earnings inequality measures. After accounting for the complex survey design (clustered sampling), they conclude that the coefficients of variation of the exponential, the Gini, and the CV^2 are 0.0026, 0.0066, and 0.0564, respectively. Assuming homoscedastic variances and equal population sizes over time, significant results (at 1%) are achieved with minimum movements of 0.005 (Exp), 0.01 (G), and 0.16 (CV^2) in absolute terms. We apply these thresholds throughout the paper. We also apply the same statistical significance criteria in tax data as we do in survey data, although one could relax the criteria substantially given the much larger sample sizes.

If the Exp, G, and CV^2 move in the same direction, then it is likely that the old and new Lorenz curves do not cross, which is indicative of an unambiguous change in inequality (based on the Lorenz ordering). If, however, the three measures do not agree in the trends over time, then the Lorenz curves definitely cross, resulting in an ambiguous change in inequality.[10]

These three inequality measures (Exp, G, and CV^2) provide a fairly robust set of summary inequality measures, but to acquire a better sense of what part of the income distribution changes over time, we also turn to an analysis of changes in mean income by vingtiles (ordered groups of 5%), and related top/bottom ratios. [11]

---

8. Unlike other bottom sensitive measures (e.g., Theil-Entropy or Theil-Bernoulli), the exponential measure is well defined for zero or negative incomes. Since we include business and self-employment income, we expect to have some negative incomes.

9. Although dropping the top and bottom 0.1% of the income distribution in each year is really only necessary in survey data to control effects of large outliers, to be consistent, we applied this measure to tax data as well.

10. See Cowell (1977) or Wolfson (1986) for more details.

11. Note that Wolfson (1997) illustrates empirically that the set of inequality measures proposed here (Exp, G, and

We use the SCF/SLID series in a manner intended to account for the seam, or break in the series (between 1995 and 1996). In particular, rather than calculating change over the seam (1995 to 1996) using data from two different and possibly not comparable surveys, SCF and SLID, we sum the *changes* from each of the two data sources to minimize the impact of the seam. For example, the total change between 1989 and 2000 would be the change in the SCF measure between 1989 and 1996, plus the change in the SLID measure between 1996 and 2000. This is possible because two sets of income data are available for 1996. Of course, the SCF and SLID numbers may not just be different in terms of levels, but also in terms of change over time. The incremental approach suggested here only accounts for the difference in levels, which is likely the more prominent discontinuity.

4. Income inequality in context: The role of the economic cycle

When comparing inequality measures over time, it is vital to place the results within context of the economic cycle. A long standing line of research in economics investigates whether hourly wages move pro- or counter-cyclically. Much of that investigation focuses on accounting for sample selection effects generated from the fact that lay-offs do not occur randomly within the wage distribution and so, wages may appear to rise in a recession when in reality we are simply observing the effects of the lower wage workers being laid off first. We do not face this type of composition effect because we keep all sample observations, including those with zero market income, throughout our period. Our market income measure will still reflect the cycle, though, because spells of unemployment will generate substantial reductions in annual earnings. Further, if returns to assets of various types fall during recessions this will also cause drops in annual market income. We view the fact that our measure will capture all of these types of effects as a good feature since movements in annual income generated from moving in and out of the labour market definitely affect individual well-being. The first of these effects (the lay-off effect) will likely cause inequality to increase during recessions since we expect lay-offs to fall disproportionately on the lower part of the market income distribution. This pattern should reverse itself and lead to larger increases in earnings at the bottom than the top of the distribution during periods of expansion. On the other hand, we expect the second effect (the reduction in returns to assets effect) to reduce inequality in a recession, especially if we focus on non-retired individuals. On balance, we expect the first effect to be stronger and so predict that we will observe a counter-cyclical movement in market income inequality. Since the tax and transfer system may partially or fully cushion these fluctuations, after-tax income inequality should be much less variable over a business cycle, and generally speaking, over time.

The Canadian economy saw two full business cycles between 1980 and 2000 (Chart 1). The economy was essentially at a business cycle peak in 1980 or 1981 (with a prime-aged male unemployment rate of around 5.1%), 1989 (6.3%), and 2000 (5.7%). Cyclical troughs were observed in 1983 (9.9%), and 1992/1993 (10.7% and 10.6%, respectively).

CV^2) isn't always reliable in detecting crossing Lorenz curves. Consequently, it may be preferable to examine the actual Lorenz curves for this purpose. Since our main goal is not to determine whether inequality changed in the strict (or unambiguous) sense, but rather in a more general sense, we opt to simply look at the income distribution as a complement to our summary measures.

Chart 1: Annual national unemployment rate, males 25-54
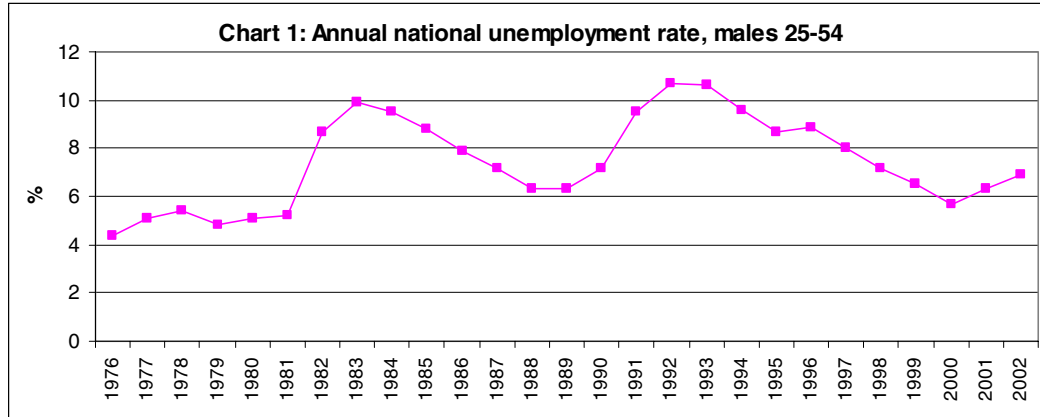
In our analysis, we focus on two "peak-to-peak" year comparisons: 1980 and 1989, and 1989 and 2000. When we turn to tax data, only the years 1992 to 2000 are suitable for income estimates, and 1993 to 2000 are best suited for this purpose. Since this period does not cover a full business cycle, we can not compare similar years in the cycle. Consequently, we place the tax results in context of the economic recovery during this period. When using the census data, we focus on changes between 1980 and 1990, and 1990 to 2000. These years are at (or very near) business cycle peaks.

Table 1: Market and after-tax income inequality, SCF/SLID - economic families*

Data source	Year	Market income			After-tax income		
		EXP	G	CV^2	EXP	G	CV^2
SCF	1980	0.4446	0.3687	0.4690	0.4141	0.2849	0.2855
	1981	0.4443	0.3684	0.4658	0.4134	0.2837	0.2801
	1982	0.4514	0.3867	0.5199	0.4143	0.2864	0.2889
	1983	0.4573	0.4015	0.5735	0.4168	0.2944	0.3128
	1984	0.4575	0.4010	0.5756	0.4164	0.2923	0.3131
	1985	0.4544	0.3936	0.5537	0.4151	0.2884	0.3062
	1986	0.4542	0.3932	0.5531	0.4147	0.2874	0.3012
	1987	0.4536	0.3921	0.5404	0.4139	0.2856	0.2910
	1988	0.4532	0.3907	0.5315	0.4125	0.2811	0.2779
	1989	0.4508	0.3849	0.5318	0.4118	0.2783	0.2779
	1990	0.4545	0.3945	0.5443	0.4124	0.2806	0.2752
	1991	0.4619	0.4123	0.6241	0.4146	0.2873	0.3022
	1992	0.4626	0.4140	0.5958	0.4131	0.2832	0.2807
	1993	0.4659	0.4213	0.6254	0.4139	0.2858	0.2899
	1994	0.4653	0.4200	0.6248	0.4132	0.2834	0.2837
	1995	0.4654	0.4204	0.6414	0.4147	0.2878	0.2997
	1996	0.4656	0.4211	0.6392	0.4158	0.2914	0.3053
SLID	1996	0.4679	0.4263	0.6786	0.4174	0.2962	0.3195
	1997	0.4681	0.4269	0.6957	0.4189	0.3003	0.3377
	1998	0.4672	0.4247	0.7097	0.4189	0.2998	0.3439
	1999	0.4635	0.4163	0.6730	0.4182	0.2978	0.3322
	2000	0.4625	0.4140	0.6802	0.4192	0.3009	0.3467

Absolute growth in inequality**:		EXP	G	CV^2	EXP	G	CV^2
Peak of the cycle comparisons:	1980-1989	0.0062	0.0162	0.0629	-0.0023	-0.0065	-0.0076
	1989-2000	0.0094	0.0239	0.1090	0.0059	0.0178	0.0546
The 1989-2000 period in detail:	1989-1993	0.0151	0.0365	0.0936	0.0021	0.0074	0.0120
	1993-1996	-0.0002	-0.0002	0.0138	0.0019	0.0056	0.0155
	1996-2000	-0.0054	-0.0124	0.0016	0.0018	0.0048	0.0272

* Income is measured at the economic family level, but the the unit of analysis is the individual.
Income is divided by the number of "adult equivalents" in the family (see text for more details).
** To partially account for the introduction of SLID in 1996, the absolute growth in inequality is the sum of the absolute growth in SCF up to 1996, plus the absolute growth in SLID from 1996 onwards.
Shaded results are significant at 1% (see text for more details).

5. Results

5.1 Evidence from survey data

In Table 1, three measures of income inequality are shown for the SCF (1980 to 1996) and the SLID (1996 to 2000): a bottom-sensitive (the exponential measure—Exp), middle-sensitive (the Gini coefficient—G), and top-sensitive (the coefficient of variation squared—CV^2) measure. Market and after-tax income is measured at the economic family level, but the unit of analysis is the individual. Thus, the results relate to income inequality among all individuals, based on their economic family income. The income data are "adult equivalent adjusted" to account for economies of scale in larger families.[12]

Our thresholds for statistically significant changes are: 0.005 for the exponential, 0.01 for the Gini, and 0.16 for the CV^2. Given the difference in these thresholds, one must be careful in comparing the magnitude of the changes over time among inequality measures.

Not surprisingly, the level of market income inequality is higher than after-tax income inequality, as the tax and transfer system offsets market income inequality to some extent (Charts 2a through 2c).

Chart 2a: Exponential, SCF/SLID - economic families

---
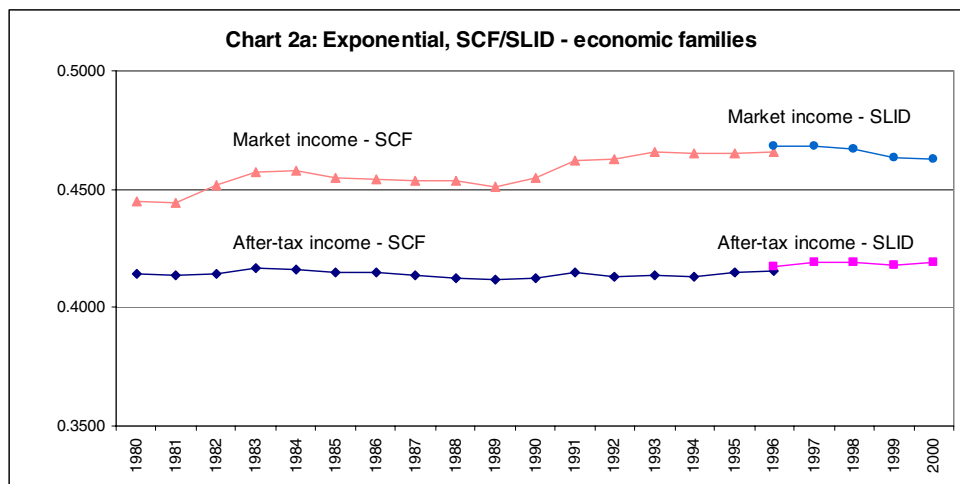
12. Specifically, family income is divided by the square root of the family size (the number of members in the family)—in essence, the needs of each additional member in the family can be met at lower cost due to economies of scale.

Chart 2b: Gini coefficient, SCF/SLID - economic families

Chart 2c: Coefficient of variation squared, SCF/SLID - economic families
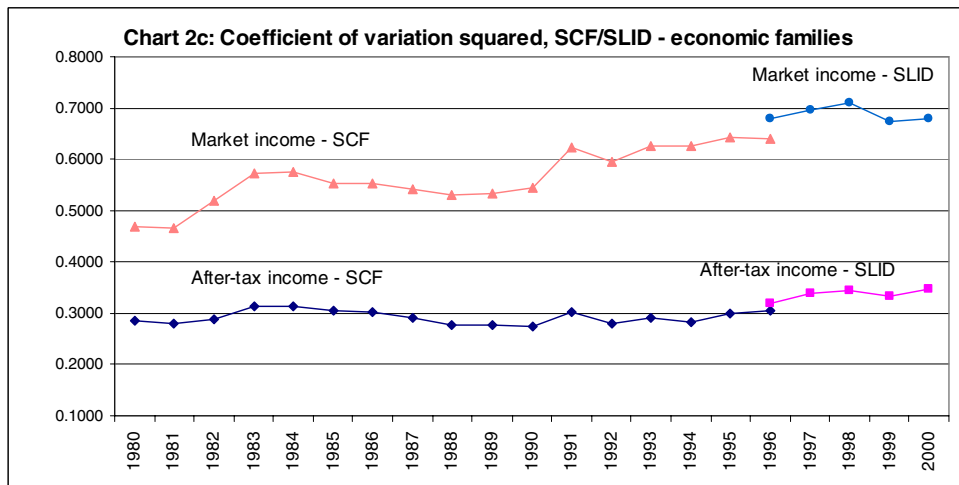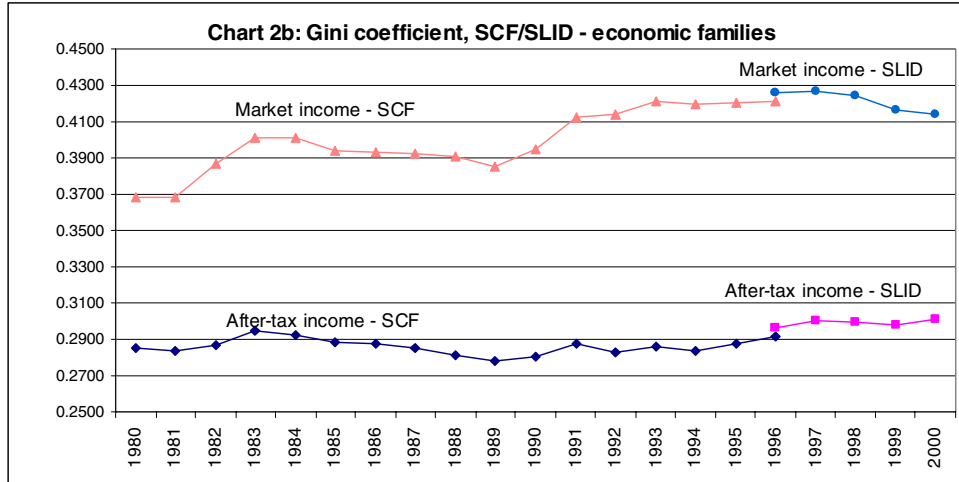
Our results are consistent with the well known fact that earnings inequality rose during the 1980s and early 1990s, even though our focus is different in several ways. First, we examine total market income (which includes income from sources such as investment and pensions) rather than earnings alone. Our analysis is also at the family level, whereas earnings studies are normally conducted on individuals. Finally, it is worth noting that in earnings studies, the sample is normally limited to employed individuals; we include all families in our analysis (including those with no earnings). These measures were taken since we want to compare inequality before and after accounting for the tax/transfer system, which can potentially affect all families.

Between 1980 and 1989 (peak-to-peak), market income inequality rose by 0.0162 based on the Gini coefficient, and by 0.0062 based on the Exp measure, both of which are well above the significance threshold. The CV^2 also rises substantially over this period (in fact, more so in percentage terms than the other two measures) but fails to exceed the 1% significance threshold. Thus, it is difficult to be sure from these measures whether the inequality increases are being disproportionately driven by one part of the distribution. Over this same period (1980 to 1989), after-tax income inequality appears to have declined very moderately by all measures, although the changes are not statistically significant. In essence, as market income inequality was rising in the 1980s, the tax and transfer

system offset this trend. This is not surprising, as the transfer system was becoming increasingly generous at the time. For example, social assistance benefits were rising during that period, especially in Ontario (National Council of Welfare, 1997).

In the 1990s, market income inequality continued to rise as a result of the recession in the early part of the decade and the "weak recovery" that followed. The largest increases were registered between 1989 and 1993, a transition from the peak of the cycle to the trough of the recession. Once again, all three indexes show strong increases but we are unable to make clear statistical statements based on the CV^{2}. The tax and transfer system reduced the magnitude of the inequality increase but did not fully offset it, as after tax and transfer income inequality also rises in the heart of the recession. However, the same was true of the recession of the early 1980s: the tax and transfer system did not fully offset market income inequality increases then either. For example, the Gini for market income rose by 0.033 from 1980 to 1983 while the after tax and transfer income Gini rose by 0.010, changes very similar to those recorded for the 1989-93 period at the bottom of Table 1. None of the changes in after tax and transfer income inequality in either period are statistically significant at the 1% level.

Between 1993 and 1996, family market income inequality was relatively stable with changes that were both economically insubstantial and statistically insignificant. The same is true of the after tax and transfer income inequality, though the indexes all register slightly higher increases in inequality for this measure than for market income. The lack of a clear trend is perhaps not surprising in a period that is a mixed bag in terms of labour market trends and, as described below, changes in the tax and transfer system.

The late 1990s witnessed a fall in market income inequality, though not by nearly enough to offset the increases during the previous recession. The same ratcheting-up of inequality occurred in the 1980s business cycle. Indeed, the recessionary increases and expansionary declines in the inequality indices were very similar in the 1980s and 1990s.[13] However, the economic expansions in the two decades were dramatically different in the roles played by the tax and transfer system. In the 1980s expansion, the inequality indices for after tax and transfer inequality fell by almost exactly the same amount as the indices for market income. In contrast, in the 1990s expansion (late 1990s), declines in market income inequality were observed, but after tax and transfer income inequality rose marginally. However, the increases in the inequality indexes for after tax and transfer income in the 1996-2000 period are not statistically significant in their own right, although the differences between the increase in the after-tax and transfer indexes and the declines in the corresponding market income inequality indexes are statistically significant for both the Exp and Gini indexes.

How do these trends concur with intuition? During recessionary periods, people at the bottom end of the earnings distribution are normally the ones who are hardest hit. Young workers who were recently hired may be the first to experience a lay-off, while potential new entrants to the labour market may find it difficult to land a job. As a result, one would expect earnings to fall at the bottom end. In other words, market income inequality is expected to rise. In expansionary periods, firms need to hire more workers to meet demand. Improvements are thus expected to occur mainly at the bottom end of the distribution, meaning that market income inequality is expected to fall during

---

13. For example, the Gini increased by 0.033 from 1980 to 1983, and by 0.036 from 1989 to 1993. It then declined by 0.017 in the expansions from 1983 to 1989 and by 0.012 from 1996 to 2000.

economic recoveries. The tax and transfer system generally plays a mitigating role in this respect: laid-off individuals are often eligible for employment insurance benefits and, failing that, social assistance benefits; income taxes may also smooth out increases registered at the top.[14]

The outcome of this discussion is that we would expect market income and after-tax and transfer income inequality indexes to move relative to one another in roughly the manner they did in the 1980s. The only difference between what happened in the 1980s and our description of the expected scenario is that after-tax and transfer income inequality declined as much as market income inequality in the 1980s expansion. Similarly, the various indexes moved in the predicted manner in the recession of the early 1990s. However, the marginal increase in after-tax and transfer inequality that accompanied the decline in market income inequality in the late 1990s may be related to changes in the transfer system in the last half of the 1990s. This issue is being addressed more fully in a subsequent paper.

To summarize, survey data suggest that family market income inequality rose in the 1980s, but after-tax income inequality remained relatively stable. In the 1990s, these data suggest that family market income inequality rose more than it did in the 1980s. After-tax income inequality also rose marginally. It appears that the tax and transfer system offset only part of the rise in family market increase inequality over the 1990s. This is in contrast with the 1980s cycle, when the tax/transfer system more than offset the rise in market inequality. In all cases, the exponential and the Gini show significant changes, while the CV^2 does not (although the change is in the same direction).

5.2 Evidence from tax data

The taxation data provide an alternate source for examining income inequality trends. We focus on the years 1993 to 2000 since this period yields the most reliable and consistent source of income estimates in the tax data, particularly for low-income families. We also show results for 1992 onwards, but it is clear from the tables that follow that the data improved in 1993 with the introduction of the Child Tax Benefit and a sudden increase in lower-income filers. Only census families are formed in the tax data, as opposed to economic families. In Table 2, we show estimates from our three inequality measures (Exp, G, and CV^2) for both survey and tax data, all measured on a census family basis; results also appear in Charts 3a to 3c. Note that to reduce processing time, a 2% random sample of the tax data was used in these calculations.
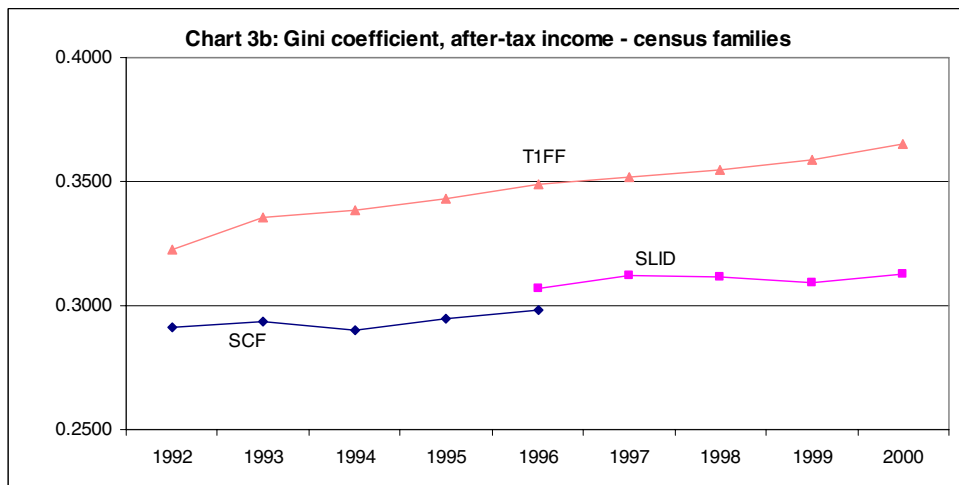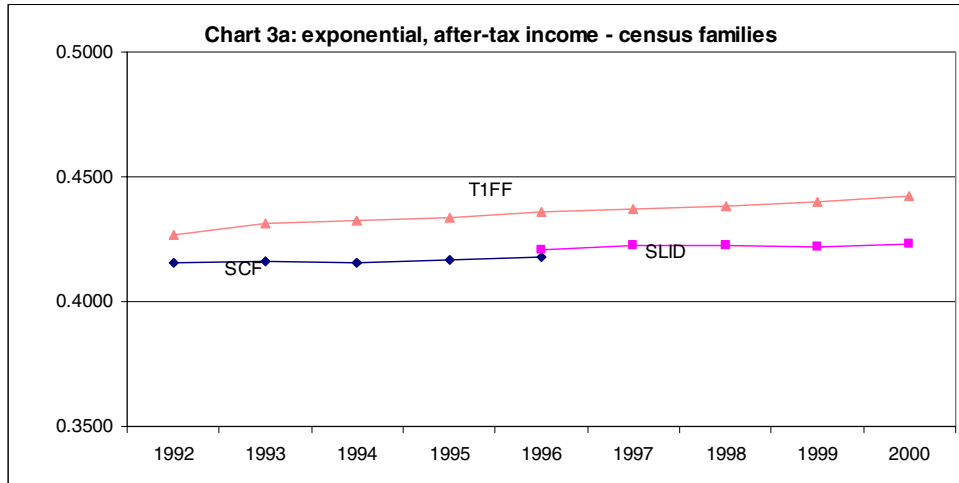
---

14. See Beach and Slotsve (1994) for more discussion on the cyclical aspects of the earnings distribution.

Table 2: Market and after-tax income inequality, SCF/SLID and T1FF - census families*

		Survey data						Tax data					
		Market income			After-tax income			Market income			After-tax income		
Data source	Year	Exp	G	CV²	Exp	G	CV²	Exp	G	CV²	Exp	G	CV²
Survey data: SCF Tax data: T1FF	1992	0.4671	0.4242	0.6264	0.4157	0.2912	0.2971	0.4837	0.4627	0.8508	0.4267	0.3228	0.4024
	1993	0.4703	0.4313	0.6521	0.4163	0.2934	0.3057	0.4905	0.4770	0.9194	0.4314	0.3357	0.4360
	1994	0.4695	0.4296	0.6551	0.4154	0.2904	0.2998	0.4920	0.4798	0.9290	0.4323	0.3385	0.4401
	1995	0.4695	0.4298	0.6708	0.4170	0.2950	0.3157	0.4923	0.4807	0.9420	0.4339	0.3428	0.4562
	1996	0.4700	0.4310	0.6710	0.4181	0.2985	0.3214	0.4951	0.4865	1.0185	0.4363	0.3490	0.4873
Survey data: SLID Tax data: T1FF	1996	0.4755	0.4435	0.7407	0.4211	0.3070	0.3475						
	1997	0.4763	0.4451	0.7621	0.4229	0.3120	0.3694	0.4950	0.4865	1.0152	0.4373	0.3518	0.4950
	1998	0.4752	0.4426	0.7734	0.4229	0.3117	0.3756	0.4963	0.4893	1.0912	0.4386	0.3548	0.5317
	1999	0.4712	0.4338	0.7307	0.4220	0.3090	0.3610	0.4954	0.4875	1.0846	0.4400	0.3585	0.5426
	2000	0.4706	0.4326	0.7496	0.4233	0.3129	0.3803	0.4965	0.4898	1.1433	0.4427	0.3653	0.5835
Absolute growth in inequality**:		Exp	G	CV²	Exp	G	CV²	Exp	G	CV²	Exp	G	CV²
The 1993-2000 recovery	1993-2000	-0.0052	-0.0112	0.0279	0.0040	0.0109	0.0485	0.0060	0.0129	0.2239	0.0113	0.0295	0.1476
	1993-1996	-0.0003	-0.0003	0.0189	0.0018	0.0050	0.0157	0.0046	0.0096	0.0992	0.0049	0.0133	0.0513
	1996-2000	-0.0050	-0.0110	0.0089	0.0023	0.0059	0.0327	0.0014	0.0033	0.1248	0.0064	0.0162	0.0963

* Income is measured at the census family level, but the the unit of analysis is the individual. Income is divided by the number of adult equivalents in the family (see text for more details).
** To partially account for the introduction of SLID in 1996, the absolute growth in inequality is the sum of the absolute growth in SCF up to 1996, plus the absolute growth in SLID from 1996 onwards. Shaded results are significant at 1% (see text for more details).

Chart 3a: exponential, after-tax income - census families

Chart 3b: Gini coefficient, after-tax income - census families

Chart 3c: Coefficient of variation squared, after-tax income - census families
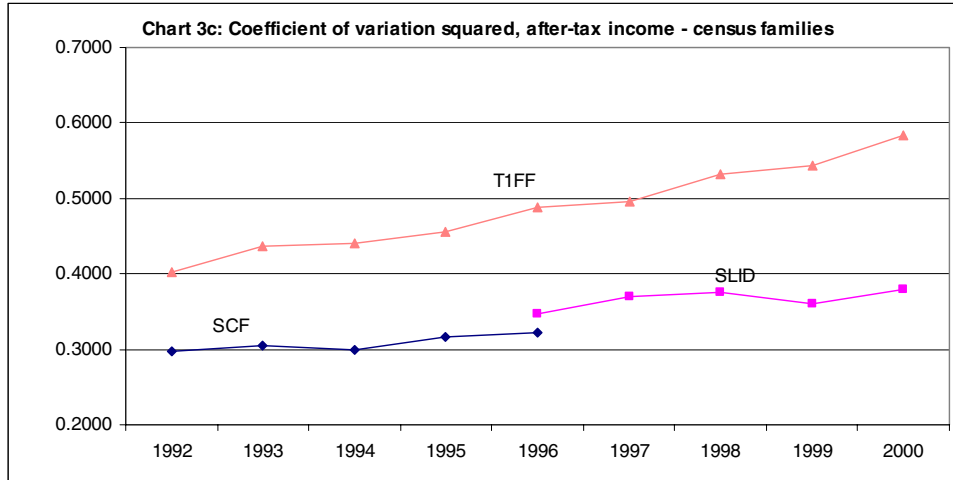
The broad inequality trends in survey data are virtually identical whether we use census or economic families between 1992 and 2000 (compare the census family results from survey data in Table 2 to economic family results from survey data in Table 1). The change in the family concept associated with the move to tax data should thus have no substantial impact on the results. As in our discussions based on economic families, the survey data grouped by census family shows no substantial movements in market income inequality between 1993 and 1996, but does show declines in inequality in the 1996-2000 period. The declines in the latter period are statistically significant for both the exponential and Gini indexes. Also as in the earlier discussion, although family market income inequality declines during the expansion, after-tax and transfer income inequality increases over the same period. The difference in the trends in inequality based on the two income measures is again statistically significant for the exponential and Gini indexes.

The trends in market income inequality using tax data are substantially different from those observed in the survey data. In particular, in the tax data, there is evidence of a substantial (though not quite statistically significant) increase in market income inequality even in the 1993-96 period. Further, in contrast to the survey data, the expansion period witnessed additional increases in market income inequality. The sum of the effects for the two sub-periods is substantial and statistically significant. As in the survey data, the tax and transfer system then has the effect of increasing inequality in final (after-tax and transfer) income relative to market income. Indeed, the net result is very large and often statistically significant increases in after tax and transfer income inequality within both sub-periods and for the 1993-2000 period as a whole. In the end, the tax data point to much stronger increases in inequality and no mediating role for the tax and transfer system in the period following 1993. In fact, changes in taxes paid and transfer benefits received tended to increase (not reduce) inequality over this period.

It is important to keep in mind that the significance thresholds applied to the tax data are based on the work of Kovacevic and Binder (1997), which used SCF data. The samples we use in the tax data are much larger, and would thus yield smaller significance thresholds (if such work was carried out). In other words, our significance tests on tax data understate the true level of significance.

5.3 Results by vingtile

We turn now to examining the distribution of income using mean incomes by vingtile (groups of observations containing 5% of the population ordered by income level). This allows us to investigate the form and movements of the income distributions in more detail than is possible using the summary measures we have employed up to this point. Further, it will help us shed more light on the differences in inequality movements between survey and tax data.

Mean market income and mean after-tax income by vingtile are shown in Tables 3a and 3b, respectively, for survey and tax data, based on census families. The values are in 2000 constant dollars, and are adjusted for family size in order to create a per capita level of income, as well as to account for economies of scale in larger families (i.e., the sharing of goods). Specifically, income is divided by the number of "adult-equivalents" in the family, which is equal to the square root of the number of people in the family.

Table 3a: Mean adult-equivalent market income by vingtile ($2000) - census families*

Data source	Year	Vingtile										
		Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top
SCF	1992	-93	456	3,418	6,823	19,036	24,273	29,573	47,130	53,409	62,250	88,834
	1993	-13	219	2,710	6,058	18,036	23,409	28,794	45,823	51,893	60,621	87,762
	1994	-20	263	2,978	6,374	18,685	23,950	29,296	46,902	53,092	62,041	90,019
	1995	-26	432	3,311	6,652	18,802	23,992	29,403	46,605	52,753	62,118	93,616
	1996	-9	314	3,175	6,752	18,852	24,179	29,470	47,455	53,395	62,926	93,882
SLID	1996	-67	226	2,701	5,848	17,341	23,134	28,566	46,121	52,154	60,812	94,035
	1997	-71	217	2,677	5,986	17,883	23,564	29,141	47,090	53,278	62,656	98,076
	1998	-144	325	3,141	6,599	19,018	24,752	30,352	49,206	55,583	64,841	103,544
	1999	-49	709	3,888	7,290	19,705	25,486	31,163	49,783	56,204	65,913	103,901
	2000	-97	958	4,471	7,892	20,554	26,418	32,190	51,253	57,909	68,433	110,245
T1FF	1992	-251	309	2,634	5,324	16,020	21,523	27,054	45,473	51,812	61,678	98,768
	1993	-182	38	1,550	4,300	15,073	20,630	26,264	44,754	51,159	61,151	99,524
	1994	-209	3	1,262	4,101	15,122	20,823	26,616	45,408	52,017	62,109	100,716
	1995	-197	13	1,382	4,110	15,296	20,933	26,779	45,772	52,444	62,765	102,774
	1996	-201	10	1,287	4,022	15,226	20,980	26,779	46,083	52,789	63,349	107,513
	1997	-187	14	1,411	4,175	15,486	21,288	27,221	46,818	53,797	64,641	109,801
	1998	-163	28	1,561	4,443	16,089	21,918	27,991	48,364	55,526	66,841	117,913
	1999	-148	33	1,718	4,779	16,665	22,660	28,830	49,555	56,992	68,741	121,520
	2000	-118	46	1,770	4,915	17,110	23,173	29,495	51,083	58,684	70,807	127,182
Absolute growth: (1993-2000)		Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top
SCF/SLID**		-25	827	2,235	2,739	4,029	4,054	4,301	6,764	7,258	9,926	22,330
Tax data		64	8	219	614	2,037	2,543	3,231	6,329	7,525	9,656	27,658

* Income is measured at the census family level, but the the unit of analysis is the individual. Income is divided by the number of adult equivalents in the family (see text for more details).
** To partially account for the introduction of SLID in 1996, the absolute growth in inequality is the sum of the absolute growth in SCF up to 1996, plus the absolute growth in SLID from 1996 onwards. The percentage growth is the absolute growth divided by the base (or earlier) year of analysis, expressed in percentage terms.

Table 3b: Mean adult-equivalent after-tax income by vingtile ($2000) - census families*

Data source	Year	Vingtile											Top decile/ bottom decile
		Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	
SCF	1992	5,065	9,648	12,008	13,941	20,278	23,312	26,682	38,317	42,478	48,545	65,257	7.7
	1993	5,229	9,500	11,709	13,497	19,495	22,618	26,017	37,400	41,303	47,456	64,802	7.6
	1994	5,354	9,599	11,935	13,888	19,998	23,096	26,211	37,779	41,827	47,917	65,231	7.6
	1995	5,253	9,463	11,834	13,717	19,784	22,847	26,275	37,681	41,740	47,896	67,354	7.8
	1996	4,981	9,113	11,576	13,630	19,740	22,851	26,168	37,994	41,944	48,256	67,262	8.2
SLID	1996	4,221	8,795	11,010	12,944	19,032	22,320	25,773	37,414	41,421	47,166	67,368	8.8
	1997	4,058	8,835	11,168	13,137	19,271	22,618	26,176	38,107	42,355	48,578	70,913	9.3
	1998	4,299	9,263	11,734	13,637	19,991	23,393	26,920	39,392	43,638	50,327	74,046	9.2
	1999	4,351	9,292	11,979	14,006	20,624	24,000	27,666	40,313	44,715	51,278	74,237	9.2
	2000	4,632	9,589	12,303	14,264	21,030	24,560	28,257	41,471	45,904	52,890	78,476	9.2
T1FF	1992	3,520	8,199	10,423	12,361	18,701	21,814	25,192	37,501	41,878	48,560	71,728	10.3
	1993	2,088	7,352	9,655	11,561	17,998	21,192	24,632	36,834	41,204	47,958	71,747	12.7
	1994	1,913	7,225	9,684	11,531	18,069	21,391	24,881	37,342	41,825	48,710	72,359	13.2
	1995	1,866	7,028	9,439	11,362	17,998	21,284	24,834	37,432	41,968	48,958	73,616	13.8
	1996	1,780	6,820	9,230	11,094	17,883	21,260	24,841	37,661	42,271	49,420	75,827	14.6
	1997	1,825	6,694	9,130	11,079	17,957	21,340	24,954	38,028	42,798	50,090	77,023	14.9
	1998	1,964	7,029	9,552	11,647	18,381	21,879	25,664	39,382	44,252	51,878	82,456	14.9
	1999	1,692	6,914	9,621	11,826	18,841	22,421	26,301	40,417	45,533	53,545	85,315	16.1
	2000	1,479	6,629	9,469	11,788	19,167	22,905	26,921	41,622	46,897	55,108	89,039	17.8

Absolute growth: (1993-2000)	Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	Top decile/ bottom decile
SCF/SLID**	163	408	1,159	1,454	2,243	2,473	2,635	4,651	5,124	6,524	13,569	1.0
Tax data	-609	-723	-185	227	1,169	1,714	2,289	4,788	5,693	7,150	17,292	5.1

Percentage growth: (1993-2000)	Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	Top decile/ bottom decile
SCF/SLID**	3.1%	4.3%	9.9%	10.8%	11.5%	10.9%	10.1%	12.4%	12.4%	13.7%	20.9%	13.3%
Tax data	-29.2%	-9.8%	-1.9%	2.0%	6.5%	8.1%	9.3%	13.0%	13.8%	14.9%	24.1%	40.2%

* Income is measured at the census family level, but the the unit of analysis is the individual. Income is divided by the number of adult equivalents in the family (see text for more details).
** To partially account for the introduction of SLID in 1996, the absolute growth in inequality is the sum of the absolute growth in SCF up to 1996, plus the absolute growth in SLID from 1996 onwards. The percentage growth is the absolute growth divided by the base (or earlier) year of analysis, expressed in percentage terms.

Survey data suggest that after-tax income rose across all of the distribution, with gains ranging from 10% to 14% for most between 1993 and 2000. The ratio of the mean income in the top to bottom deciles rose by 13.3% over the period. In the top vingtile, after-tax income rose by 20.9%, while in the bottom vingtile, there was a much smaller increase of 3.1%. This implies that the rising inequality detected over this period was the consequence of above average gains at the top, and a *lack* of relative improvements at the bottom.

Tax data paint a somewhat different picture. As is the case with survey data, the rising inequality seems to have resulted in part from an improvement in the top vingtile. Indeed, tax and survey data register quite similar increases in average incomes over about the top 40% of the distribution. In contrast with survey data, however, we find a sharp decline in after-tax income at the bottom. In the bottom vingtile, after-tax income declined by 29.2% between 1993 and 2000. As a result of the large gains at the top, and the very large declines at the bottom, the ratio of the mean income in the top to bottom deciles rose by more than 40% between 1993 and 2000. This is a direct reflection of the trends in market income portrayed in Table 3a. While the survey data show increases in mean market income across most of the distribution, suggesting that almost everyone benefited from the

1990s expansion (though to varying degrees), the tax data on market income indicates that people in the bottom 20% of the distribution saw no real improvement.[15]

In addition to the differences in trends in income between the two data sources, especially at the bottom of the income distribution, there are also large differences in the levels of income, once again especially at the bottom end. In a very broad sense, the distribution of income generated from survey and tax data might differ for two reasons: either the respondents/tax filers report different levels of income, or the population covered is different. It is of course likely that income from some sources is not reported in tax data. It is also conceivable that people at the bottom of the income distribution under-report to a greater extent, at least expressed as a proportion of their income. However, survey respondents may also not include income from some sources. The gap in the bottom vingtile between tax and survey data (SCF, a file without any tax information) is generally about $3,000 in adult-equivalent 2000 dollars. For a family of four, this difference is about $6,000 in unadjusted 2000 dollars. This represents a difference in the order of 200% of the incomes reported in the tax data. The move from the SCF survey data to (mainly) tax data in the SLID series did not result in a significant narrowing of this gap. This suggests that the difference may lie more in coverage than in differential reporting between survey and tax data and whether survey data (both SCF and SLID) cover a different population than tax data is a very difficult question to answer, and is beyond the scope of this study.[16]

In the next section, we discuss how the Census can be used to provide a third portrait of inequality trends, and assess if they more closely resemble the tax or survey data results.

5.4 Evidence from Census data

If the gap between survey and tax data is indeed caused by a difference in the way income is reported, then we would expect the gap to remain even if somehow the coverage in survey data was as high as in tax data. Alternatively, if the difference in coverage is at the heart of this gap then we would expect the gap to disappear if the survey data coverage were to match that of tax data.

One way of testing these competing hypotheses is to turn to Census data. The Census is a survey aimed at collecting information on the entire population, and is conducted every five years shortly after the tax season (sometime in May in recent years). By law, response is mandatory, and every

---

15. All along, we've described trends in market income inequality by referring to earnings (the main component of market income). The appendix divides the sample in the tax data by vingtile of market income, and looks at the mean of the three components of market income: paid earnings, net self-employment income, and other market income. Clearly, the overall trend in market income inequality is being driven by paid earnings.

16. Note that coverage rate in T1FF is about 96% to 97% for the period of study. The response rate is between 80% and 85% in SLID, and slightly lower in SCF. As mentioned in the data section, the weights in the survey data are adjusted to make the sample more representative of various province-age-sex groups, as well as household and family sizes. However, no specific adjustments are made to correct the non-representativeness of the sample by income level that may be introduced by non-response. In particular, it would be difficult to automatically dismiss the possibility that lower-income families are under-represented in survey data. For example, we do know that estimates of the immigrant population are considerably lower in survey data and immigrants are more concentrated at the bottom end of the income distribution (Hou and Picot, 2003). Of course, other possibilities would also be difficult to dismiss. (In 1996, immigrants account for 21.5% of individuals 15 years or older in the Census, compared to 19.2% in SLID—authors calculation).

fifth household has to fill out the "long form", which includes detailed questions on the amounts of various income sources from the previous year. Thus, the Census provides a convenient vehicle for examining the importance of the extent of coverage: it is a survey like the SCF, but has coverage rates that are closer to those in the tax data. One drawback of the Census is that it does not contain information on taxes paid. We could impute taxes but instead of doing so, we focus on differences in total, pre-tax income (i.e., the sum of market and transfer income). This still provides a good basis for comparison among the datasets, especially at the bottom of the distribution, where taxes are less important.

The distribution of pre-tax (i.e., market plus transfer) income of census families in adult-equivalent 2000 dollars is shown below in Table 4 for SCF, Census, and T1FF. We focus on a comparison in 1995 to maximize the comparability between the survey and Census data. If we were to focus on 2000, our "survey" data would come from SLID, in which most of the income data actually comes from tax files.

Table 4: Mean adult-equivalent pre-tax income ($2000) by vingtile, SCF, Census, and T1FF - census families*

Data source	Year	Vingtile											Top decile/ bottom decile
		Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	
SCF	1980	4,441	9,259	11,833	14,397	23,705	27,672	31,791	46,191	51,497	59,749	84,383	10.5
	1985	4,833	9,095	11,822	14,050	22,685	26,893	31,166	46,563	51,793	59,700	86,674	10.5
	1990	5,556	10,129	12,902	15,340	24,598	29,196	33,874	49,991	55,569	64,302	93,656	10.1
	1995	5,279	9,589	12,185	14,417	22,919	27,355	32,210	48,506	54,407	63,574	95,041	10.7
	1996	5,051	9,231	11,906	14,338	23,010	27,490	32,252	49,206	55,107	64,415	95,204	11.2
SLID	1996	4,332	8,911	11,274	13,540	21,848	26,493	31,370	47,956	53,701	62,262	95,542	11.9
	2000	4,814	9,742	12,670	14,943	24,238	29,200	34,526	52,708	59,140	69,394	111,212	12.4
Census	1980	2,588	8,389	11,408	13,884	22,982	27,127	31,447	46,858	52,417	61,236	93,229	14.1
	1985	2,373	7,871	11,019	13,463	22,297	26,560	30,959	46,828	52,585	61,675	94,301	15.2
	1990	3,005	8,979	12,315	14,900	24,464	29,016	33,812	51,118	57,393	67,378	104,864	14.4
	1995	2,262	7,586	10,800	13,398	22,639	27,353	32,240	49,449	55,684	65,542	102,365	17.0
	2000	3,104	8,757	12,480	15,227	24,939	29,970	35,356	54,574	61,663	73,157	121,260	16.4
T1FF	1995	1,972	7,123	9,593	11,670	20,273	25,003	30,102	47,867	54,279	64,371	104,274	18.5
	2000	1,734	6,784	9,654	12,146	21,587	26,862	32,459	52,909	60,325	72,212	128,590	23.6

1995 comparisons:	Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	Top decile/ bottom decile
SCF/Census	2.33	1.26	1.13	1.08	1.01	1.00	1.00	0.98	0.98	0.97	0.93	0.63
T1FF/Census	0.87	0.94	0.89	0.87	0.90	0.91	0.93	0.97	0.97	0.98	1.02	1.09

Percentage Growth:	Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top	Top decile/ bottom decile
1980-1990												
SCF	25.1%	9.4%	9.0%	6.6%	3.8%	5.5%	6.6%	8.2%	7.9%	7.6%	11.0%	-4.3%
Census	16.1%	7.0%	7.9%	7.3%	6.4%	7.0%	7.5%	9.1%	9.5%	10.0%	12.5%	2.1%
1990-2000												
SCF/SLID	-0.4%	-0.7%	3.1%	2.6%	3.3%	3.4%	4.5%	7.9%	9.0%	11.3%	18.4%	15.9%
Census	3.3%	-2.5%	1.3%	2.2%	1.9%	3.3%	4.6%	6.8%	7.4%	8.6%	15.6%	14.0%
1995-2000												
SCF/SLID	4.8%	4.9%	9.2%	9.2%	10.8%	10.4%	9.9%	11.2%	11.3%	12.5%	16.7%	9.4%
Census	37.2%	15.4%	15.6%	13.6%	10.2%	9.6%	9.7%	10.4%	10.7%	11.6%	18.5%	-3.9%
T1FF	-12.1%	-4.8%	0.6%	4.1%	6.5%	7.4%	7.8%	10.5%	11.1%	12.2%	23.3%	27.1%

* Income is measured at the census family level, but the the unit of analysis is the individual. Income is divided by the number of adult equivalents in the family (see text for more details).
** To partially account for the introduction of SLID in 1996, the absolute growth in inequality is the sum of the absolute growth in SCF up to 1996, plus the absolute growth in SLID from 1996 onwards. The percentage growth is the absolute growth divided by the base (or earlier) year of analysis expressed in percentage terms.

Based on pre-tax income, the bottom of the distribution in Census data looks more similar to the tax data, especially at the top and bottom of the distribution. In 1995, mean pre-tax adult-equivalent income in the bottom vingtile was around $2,300 in the Census and $2,000 in T1FF, compared to $5,300 in SCF. In the second vingtile, Census and T1FF are only about $500 apart, while Census and SCF are about $2,000 apart; in relative terms, the gap between SCF and Census is non-negligible.

At the bottom of the table, the mean incomes by vingtile in the SCF and T1FF data are compared to the Census numbers in the same vingtile. The SCF data exhibits much larger differences relative to the Census, particularly at the bottom of the distribution. In the bottom vingtile, the SCF value is 2.3 times the Census value. This is followed by the second vingtile in SCF, which is 1.3 times the Census value. The differences between the SCF and Census become smaller and smaller at successive vingtiles until the top vingtile where the SCF value is 93% of the Census value. In contrast, the values taken from tax data tend to be a relatively even proportion below those from the Census with no discernable pattern across vingtiles.

Although it is quite possible that income is under-reported in T1FF relative to surveys (SCF), this does not appear to be the major cause of the difference between tax data and SCF at the bottom of the income distribution. If reporting of income components was the major cause of the difference, this should be detected when moving from purely survey (SCF) to largely tax (SLID) reporting, given that these two data sources have the same coverage issues; however, very little change is observed. Given the similarities between tax and Census results (which differ in the method of income reporting but not very much in coverage), another possible hypothesis is relative under-coverage at the very bottom of the income distribution in SCF (and SLID). Further analysis would be required to reach a more definitive conclusion in this regard.

While Census and T1FF yield similar distributions in 1995, this is not the case in 2000. Specifically, improvements at the bottom of the distribution in Census simply don't register in T1FF, although in levels, incomes in Census are still quite far from those registered in SLID. It is unclear why Census and T1FF do not yield similar stories at the bottom of the distribution in the late 1990s. In particular, the dichotomy does not seem to be generated by a difference in any one particular income component. Census shows improvements across a broad range of components, while no changes or slight declines are registered in T1FF. It is not clear why this dichotomy exists, but it is worth noting that in tax data, census families must be imputed, and some evidence suggests that important differences in family type exist at the bottom of the income distribution. For example, the number of couples with less than $10,000 in total income is 46% greater in tax data than in Census, despite the fact that there is virtually no difference across the entire income distribution[17]. Furthermore, average family size is markedly higher in the bottom of the T1FF distribution, particularly in 2000. In 1995, average family size in the bottom vingtile is 7% higher in T1FF. By 2000, this gap had risen to 12%. More work would be needed to determine the cause of these differences in family structure at the bottom of the income distribution, and whether it can explain why income trends are so different in the late 1990s.

---

17. These numbers were produced by the Small Area and Administrative Data Division at Statistics Canada, where the T1FF file and other files derived from it are housed.

Let us now look at the trends in income inequality in the 1980s and 1990s with survey data and Census, the two data sources available for such an analysis. Although Census data can not be used to study trends in after-tax income inequality (because income taxes paid are not collected), one can nevertheless compare pre-tax income inequality trends between the Census and SCF/SLID. In the 1980s, inequality remained fairly stable over the entire period according to both data sources, although a slight increase is observed in the early 1980s in Census (but not in SCF). In the 1990s, both data sources suggest that inequality rose. In SCF/SLID, the top/bottom decile ratio rose by 15.9%, compared to a 14% increase in Census. When one separates the early and late 1990s, however, the data sources tend to disagree. According to SCF, inequality rose moderately in the early 1990s because of a moderate decline at the bottom and no change at the top. In the late 1990s, SCF/SLID suggests that inequality continued to rise moderately, although this was the result of a substantially rise at the top that tended to counterbalance the moderate improvement at the bottom. Census data paints a very different picture. In the early 1990s, inequality increased substantially because of a large decline at the bottom and no change at the top. In the late 1990s, inequality actually remained fairly stable, since income rose substantially at both the top and the bottom. In essence, the two data sources disagree in the extent to which the bottom of the income distribution was affected by the recession of the early 1990s and the recovery in the latter half of the decade. This is consistent with the notion that Census has a greater tendency to accurately represent families in the bottom of the income distribution, and it is this segment of the population that saw the largest fluctuations in income over the 1990s. Of course, additional research would be needed to ensure that this is indeed the case. Whether this is true or not doesn't affect the main story that can be drawn from both data sources: income inequality rose in the 1990s as the result of large improvements at the top of the income distribution and no change at the bottom.

6. Conclusion

This study examines trends in after-tax family income inequality over the 1980s and 1990s, a period covering two full economic cycles. To this end, we begin our investigation with survey data (SCF and SLID), but quickly turn to tax data (T1FF) for three reasons. First, survey data has a break in the series in the mid to late 1990s. Second, the response rate in survey data is generally around 80% to 85%, and no adjustments are made to make the samples representative of the Canadian population in various income groups. Tax data is less likely to suffer from income-related response bias since the coverage rate is above 95% (since 1992). And finally, given the larger sample sizes in tax data, inequality estimates are less susceptible to high levels of sampling variability.

Using survey data, we confirm the findings of others for the 1980s: family market income inequality rose over the decade,[18] but the tax and transfer system offset this trend, resulting in no significant change in after-tax income inequality.

Survey data suggest a small statistically significant increase in after-tax income inequality in the 1990s (1989 to 2000). Over the whole decade, this again reflects an increase in market income inequality offset by the actions of the tax and transfer system. However, on closer examination, the inequality reducing effects of the tax and transfer system occur entirely in the recessionary period before 1993. In the expansionary period following 1996, market income inequality falls but

---

18. Our calculations differ from those typically reported for the 1980s, in that we include people with zero earnings in the sample; most studies focus on individuals with positive earnings.

inequality in after-tax and transfer income inequality rises. It is possible that this reflects changes in the tax and transfer systems that begin in the middle of the decade, including reduced social assistance benefits and changes in unemployment insurance. However, we do not try to establish the effects of program changes on inequality in this paper.

The tax data we use points to much larger increases in market income inequality driven both by rises at the top of the distribution and very substantial falls at the bottom. According to the tax data, and in contrast to the survey data, significant increases in after-tax and transfer income inequality were witnessed in the 1990s. Moreover, the level of inequality is much higher in the tax data than the survey data in each year, due mainly to much lower earnings at the bottom of the distribution.

In essence, survey and tax data vary in their estimates of the extent to which inequality rose during the recovery of the 1990s. This variation was largely due to the fact that after-tax income at the bottom of the distribution fell substantially according to tax data, while in survey data it simply failed to increase at the same pace as at the top of the distribution.

Aside from differences in trends, the level of income received by families at the bottom of the income distribution is considerably higher in survey data than it is in tax data. We put forth two hypotheses that could explain the difference. First, survey respondents may be more likely to report certain (small) income amounts than tax filers. These small income amounts might be more prominent at the bottom of the distribution. However, the move from pure survey data (in SCF) towards partial tax data (in SLID) does not result in any substantial decline in income at the bottom of the distribution, suggesting this is not likely the primary explanation.

A second hypothesis relates to the possible under-coverage of low-income individuals in survey data. If the gap between survey and tax data is indeed caused by a difference in the way income is reported, then we would expect the gap to still be present even if somehow the coverage in survey data was as high as in tax data. Alternatively, if the difference in coverage is at the heart of this gap, then we would expect the gap to disappear if the coverage in survey data were to match that of tax data.

To shed some light on this issue, we compare the pre-tax income distribution from survey and tax data with that of Census data. The Census collects income data in much the same was as the SCF did, yet the coverage rate is much higher (as in the tax data).

We find that the bottom end of the income distribution in Census data more closely resembles the tax data than the survey data, both in terms of levels and trends. Income at the bottom end of the distribution is always higher in survey data than in Census and tax data. However, the trend in income inequality suggested by tax data is much different than what we observe in Census. In the late 1990s, tax data point to a large increase in income inequality as the result of a large rise in income at the top, and no change at the bottom. Census data also points to a large increase at the top, but this is counterbalanced by an equally large rise at the bottom (in relative terms). Although it is not clear what is driving this difference, we note that tax data tends to have substantially more couples with very low incomes compared to Census, despite the fact that the overall number of couples is virtually identical.

If we focus on survey and Census data, we find very little difference in the trends over the 1980s. Essentially, both data sources suggest that inequality remained stable. Over the entire period covered by the 1990s, we also get similar results in both data sources, although this time, income inequality rose substantially. This agreement in the trends for the 1990s masks important differences observed in the experiences of families in the bottom of the distribution over different parts of the decade. Specifically, survey data suggests that these families saw moderate declines during the recession of the early 1990s, and moderate improvements during the recovery of the late 1990s. Census data suggests that they were harder hit by the recession, but that greater gains were registered later in the decade.

Appendix

Mean adult-equivalent paid earnings, net self-employment income, and other market income by vingtile of market income ($2000) - census families*

Data source	Year	Vingtile										
		Bottom	2nd	3rd	4th...	...8th...	...10th...	...12th...	...17th	18th	19th	Top
Paid earnings	1992	191	162	1,381	2,832	11,839	17,363	23,003	40,427	45,982	53,952	70,531
	1993	130	25	786	2,219	10,842	16,619	22,105	40,003	45,386	53,266	71,310
	1994	135	2	652	2,162	10,734	16,423	22,448	40,096	46,067	54,093	72,591
	1995	125	10	680	2,067	11,033	16,371	22,541	40,169	46,227	53,955	72,807
	1996	153	3	627	2,011	10,937	16,493	22,122	40,383	45,803	54,044	74,733
	1997	148	9	697	2,120	10,951	16,750	22,540	41,077	47,132	55,396	76,887
	1998	115	24	754	2,304	11,605	17,205	23,378	42,220	48,742	57,641	82,713
	1999	105	16	851	2,542	12,154	17,772	23,873	43,411	49,816	58,828	85,111
	2000	84	19	884	2,609	12,529	18,339	24,435	44,557	51,292	60,894	90,412
Net self-employment income	1992	-406	-60	26	343	934	960	948	1,271	1,636	2,350	13,360
	1993	-287	-14	-22	202	962	980	1,074	1,332	1,640	2,790	13,914
	1994	-305	-2	-51	175	1,045	1,166	1,143	1,596	1,711	2,592	14,367
	1995	-334	-11	-70	226	989	1,160	1,122	1,567	1,774	2,789	15,012
	1996	-395	-5	-39	229	1,037	1,207	1,276	1,448	1,958	2,808	16,570
	1997	-323	-11	-27	260	1,228	1,117	1,334	1,740	1,959	3,279	17,116
	1998	-284	-11	-4	305	1,103	1,260	1,245	1,815	2,041	3,279	17,888
	1999	-271	-8	-3	361	1,140	1,260	1,349	1,817	2,071	3,474	19,134
	2000	-235	-13	-19	373	1,130	1,230	1,277	2,007	2,205	3,339	19,362
Other market income	1992	-36	206	1,227	2,150	3,247	3,200	3,103	3,774	4,194	5,376	14,877
	1993	-25	27	787	1,879	3,270	3,030	3,086	3,419	4,133	5,096	14,299
	1994	-39	4	661	1,764	3,343	3,235	3,026	3,716	4,239	5,424	13,757
	1995	12	14	772	1,818	3,273	3,402	3,116	4,037	4,443	6,021	14,955
	1996	40	12	699	1,782	3,253	3,281	3,382	4,253	5,028	6,497	16,211
	1997	-12	15	741	1,795	3,307	3,421	3,348	4,001	4,706	5,966	15,797
	1998	6	15	811	1,834	3,381	3,453	3,367	4,329	4,743	5,921	17,313
	1999	18	26	870	1,875	3,372	3,628	3,608	4,327	5,106	6,440	17,276
	2000	33	40	905	1,933	3,452	3,605	3,783	4,519	5,187	6,574	17,409

* Income is measured at the census family level, but the the unit of analysis is the individual. Income is divided by the number of adult equivalents in the family (see text for more details).

Bibliography

Beach, C.M. and G.A. Slotsve. 1994. "Polarization of earnings in the Canadian labour market." Bell Canada Paper on Economic and Public Policy.

Cowell, F.A. 1977. *Measuring Inequality.* Oxford: Phillip Allan Publishers.

Frenette, M. and G. Picot. 2003. "Life after welfare: The economic well-being of welfare leavers in Canada during the 1990s." Statistics Canada, Analytical Studies Research Paper Series no. 192. Catalogue No. 11F0019MIE2003192.

Hou, F. and G. Picot. 2003. "The rise in low-income rates among immigrants in Canada." Statistics Canada, Analytical Studies Research Paper Series no. 198. Catalogue No. 11F0019MIE2003198.

Jantti, M. 1997. "Inequality in five countries in the 1980s: the role of demographic shifts, markets and government." *Economica*, 64 (255): 415-440.

Kapsalis, C. 2001. "An assessment of EI and SA reporting in SLID", Statistics Canada, Analytical Studies Research Paper Series no. 166. Catalogue No. 11F0019MIE2001166.

Kovacevic, M.S. and D.A. Binder. 1997. "Variance estimation for measures of income inequality and polarization: The estimating equations approach." *Journal of Official Statistics, 13* (1): 41-58.

Llg, R.E. and S.E. Haugen. 2000. "Earnings and employment trends in the 1990s." *Monthly Labor Review*, March: 21-33.

Morissette, R., J. Myles, and G. Picot. 1994. "Earnings inequality and the distribution of working time in Canada." *Canadian Business Economics, 2* (3): 3-16.

National Council of Welfare. 1997. *Another Look at Welfare Reform.* Minister of Public Works and Government Services, Canada.

Rashid, A. 1998. "Family income inequality, 1970-1995." *Perspectives on Labour and Income*, 10 (4): 11-17.

Sargent, T.C. 1998. "The BU ratio: Prospect and retrospect." Economic and Fiscal Policy Branch, Department of Finance, Working paper 98-09.

Statistics Canada. 2002. *Income in Canada 2000*. Statistics Canada Catalogue no. 75-202-XIE.

Wolfson, M. 1986. "Stasis amid change: Income inequality in Canada, 1965-1983." *Review of Income and Wealth*, 32 (4): 337-369.

Wolfson, M. 1997. "Divergent inequalities: Theory and empirical results", *Review of Income and Wealth*, 43 (4): 401-421.

Wolfson, M. and B. Murphy. 1998. "New views on inequality trends in Canada and the United States", *Monthly Labor Review*, April: 3-23.