

Paper presented at the John Deutsch Institute Conference, “Retirement Policy Issues in Canada”, October 26-27, 2007 at Queen’s University. The final version is published in a book entitled *Retirement Policy Issues in Canada*, edited by Michael G. Abbott, Charles M. Beach, Robin W. Boadway and James G. MacKinnon, 2009 (Kingston: John Deutsch Institute, Queen’s University). Published in cooperation with McGill-Queen’s University Press and available at: <http://mqup.mcgill.ca> .

The Distribution of Seniors’ Income using the Statistics Canada

Longitudinal Administrative Database, 1992-2005

DRAFT prepared for Conference on
Retirement Policy Issues in Canada
Queen’s University,
October 26-27, 2007

Michael R. Veall
Department of Economics
McMaster University
Hamilton, ON Canada L8S 4M4

veall@mcmaster.ca

Abstract: Gini coefficients and other summary income distribution measures are calculated for recent years for the population of Canada 66 years of age and over using the Statistics Canada Longitudinal Administrative Database (LAD), based upon income tax returns. The period beginning 1992 is emphasized because 1992 is the first year that the data include the Guaranteed Income Supplement (GIS) and Spousal Allowance (SpA) and because in 1992, the coverage rate of the LAD sampling frame climbed sharply to over 95% and has maintained that level since. Market income, total income (including transfers) and after-tax income are studied. An adjustment is made for family size. The inequality measures for all three types of income increase between 1992 and 2005, with the biggest jump between 1995 and 2000. A very high degree of concentration is found at the upper end of the income distribution, particularly with respect to employment income. It is noted that this has implications for discussions of the empirical effects of Old Age Security (OAS) and GIS clawback on labour supply by the older population. The lower end includes individuals with surprisingly low incomes given government transfer programs. While some of these individuals are immigrants who do not qualify for these programs, many are not. Members of the low end of the distribution are disproportionately unattached (single, divorced or separated) and female.

Acknowledgements: I am grateful for the support of the research program on the Social and Economic Dimensions of an Aging Population (SEDAP) and of the Social Sciences and Humanities Research Council of Canada and for the assistance of Deb Fretz and Simo Goshev, as well as Tom Swoger, André Bernard and Habib Saani of Statistics Canada.

1. Introduction

The income distribution of the older population in Canada might be of special policy interest for two key reasons. First, some of Canada's most important transfer programs, Old Age Security and the Guaranteed Income Supplement, are targeted at this population, as of course is the Canada Pension Plan/Quebec Pension Plan. At the same time, the older population has a number of special tax measures, from the clawback of Old Age Security to the recent allowance of pension income splitting for tax purposes. Second, and perhaps related to the first reason, labour supply by the older population is smaller, and hence the efficiency/equity tradeoff with respect to labour income may be arguably less important for that group, leading to a greater focus within that group on equity or at least on poverty reduction. (However, note that a number of policy proposals designed to reduce the public burden of income support for seniors focus on postponing retirement and hence increasing labour supply by the older population.)

In an important study of recent trends in Canada's all-ages family income distribution, Frenette, Green and Picot (2004) emphasize the use of taxfiler data as opposed to survey data (the Survey of Consumer Finances and its successor the Survey of Labour and Income Dynamics). They point out that since 1992, the coverage rate of the tax data has been 95% or better, because low-income individuals who would not pay tax increasingly have found filing advantageous because of refundable tax credits. These high coverage rates compare to coverage of about 80% for the corresponding surveys, where much of the coverage gap appears to be in the lower tail. Hence inequality measures using survey data are biased downwards.¹

Saez and Veall (2005, 2007) also used taxfiler data to study the Canadian all-ages individual income distribution, following other studies for example by Piketty and Saez (2003) for the United States and Atkinson (2007) for the United Kingdom by focusing on the very top end, that is the percentage of pre-tax income received by the top 5%, 1%, 0.1% and 0.01% individuals in the population. One of the findings of Saez and Veall is that there has been a recent surge in these top shares. For example between 1980 and 2000, the share of the pre-tax income received by the top 5% recipients rose from 22.7% to 29.0%, the top 1% share rose from 8.1% to 13.6%, the top 0.1% share rose from 2.0% to 5.2% and the top 0.01% share rose from 0.5% to 1.9%. While many measures of inequality (such as the log of the ratio of the 90th percentile of income to the 10th percentile of income) are not affected by these top shares and others such as the Gini coefficient² are relatively insensitive to changes in the tails, these top shares might nonetheless be of interest to students of the income distribution.

¹ Survey data may also undersample those at the top end because such individuals may successfully avoid participation. It seems imaginable that such individuals may underreport by not including less transparent sources of income. There may also be topcoding.

² Indeed Frenette, Green and Picot (2004) exclude the top and bottom 0.1% of incomes from all of their calculations. This is often done with survey analysis to exclude large outliers; they make the same exclusions for taxfiler data for comparability.

The current study therefore uses Canadian taxfiler data to study the income distribution of the older population in recent years, with some additional discussion of the bottom and top ends of the income distribution. This discussion includes reference to such variables as immigration status, age, gender and marital status.

Section 2 will briefly describe the data. Section 3 will give the results. Section 4 concludes.

2. The Data

The data source is the Longitudinal Administrative Databank (LAD), which is a replenished, longitudinal sample for 20% of all taxfilers in Canada (so that for 2005 it has over 4 million records). The variables available are confined to those available from tax records, so that there is no information on health or education (except as those variables may affect certain tax deductions or credits). However Statistics Canada, using information provided directly on the form as well as other methods such as address matching, does link individuals into families.

The LAD begins in 1982. However the analysis here does not begin until 1992. One reason is that it was not until 1992 that the LAD included data for the Guaranteed Income Supplement (GIS) and the Spousal Allowance (SpA). There would be little point in proceeding without these important sources of income included.

It might be possible to impute GIS and SpA for earlier years, but there is another reason for beginning the analysis in 1992. Since 1992 LAD coverage rates have been 95% or higher, suggesting that the selection bias associated with income distribution bias is probably small. In the early 1990s, refundable tax credits began to bring the lower tail of the income distribution into the filing population. Of these, the most universal incentive is the Goods and Services Tax Credit, a refundable credit first paid in late 1990. In addition, filing an income tax return for 1992 made one potentially eligible for the 1993 Child Tax Benefit, also paid to families even if they had no earnings. As noted in the introduction, Frenette, Green and Picot (2004) find that coverage biases are likely empirically important.³

Three types of family income are examined. The first is market income, defined as total income excluding government payments such as Workers' Compensation, the Child Tax Benefit, the Goods and Services Tax Credit, Employment Insurance, Canada Pension Plan/Quebec Pension Plan (CPP/QPP), Old Age Security (OAS), the Spousal Allowance (SpA) and Guaranteed Income Supplement (GIS). It does include Alimony and Support Income, pensions, income from Registered Retirement Savings Plans withdrawals as well as all income derived from employment (including self-employment and business income), net rental income and all capital income.

³ Frenette, Green and Milligan (2007) provide an all-ages income distribution analysis of the 1980s and the 1990s using census data (thereby solving the coverage problem), where they impute taxes using an econometric equation based on the taxfiler data.

The second is before-tax income, defined as income including government payments with no subtraction of any type of tax. It is market income as defined above but with government payments included.

The third is after-tax income, defined as before-tax income less federal and provincial personal income taxes. The Quebec Abatement (a rebate on federal tax) is included in this measure. Quebec provincial income taxes are estimated in the LAD data. Federal or provincial payroll taxes, employment insurance premiums and CPP/QPP premiums are not deducted nor is there any deduction for sales taxes.

Capital gains are included in all the income definitions. In part this is for simplicity: if capital gains were to be excluded then presumably the tax paid on capital gains should also be excluded and this would require an estimate of the marginal tax rate paid on those capital gains, which is not entirely straightforward to calculate. Second, it is not entirely clear that capital gains during retirement should be treated differently than other types of capital income. In any case, the key table is recalculated in the Appendix Table 1 with capital gains excluded. Perhaps more appropriately the analysis should be repeated using three-year or five-year income averages. (The latter step was done by Saez and Veall (2005, 2007) without any important change in the results, a finding consistent with Beach, Finnie and Gray (2001).)

Following Frenette, Green and Picot (2004), zero incomes are included in all samples. Indeed those with negative incomes are also included, although their incomes are set as zero as for example is required for the calculation of Gini coefficients. Individuals who die, immigrate or emigrate in the sample year are excluded.

In all cases, the relevant income measure is at the family level. In the vast majority of cases for those 66+, this is either individual income (for an unattached individual) or the combined income of a couple (most commonly the sum of the incomes from their individual filings but in those cases where the spouse earns little and does not file, any reported spousal income would be included). However there are some cases where there are children in the census family, perhaps children still living with their parents or adopted children, such as adopted grandchildren.

To convert these family incomes to single equivalents, something very close to the Statistics Canada family-size adjustment is used. Family income is divided by the sum of one plus 0.4 if a spouse is present and $0.3 \times (\text{the number of children})$. This would be exactly the Statistics Canada adjustment if adult children were given a weight of 0.4 rather than 0.3. A rival adjustment, used for example in Frenette, Green and Milligan and most calculations involving the Luxembourg Income Study, is to divide by the square root of the number of individuals. Counts and calculations are by individual, where for example if there is a couple both 66 or over, both are included in the calculations but if only one has attained the age of 66, only he/she is included. Note that there are a small number of cases where individuals 66 years of age or over live with their older parents and these children are included in the calculations.

3. Results and Discussion

Table 1 presents some basic measures of inequality, the log of the ratio of the 95th percentile of income to the 5th percentile of income ($\log(95/5)$) as well as $\log(90/10)$ and $\log(50/10)$, similarly defined. The table also presents the Gini coefficients. All results are presented for all three income measures although for market income, none of the log ratio values are defined as the denominator percentile does not exceed zero.

It is easy to summarize the results of the table. The measures of inequality suggest that for market income, total income and after-tax income all became more unequal, with the biggest jump between 1995 and 2000. The timing and size of the changes appear to be broadly comparable across income definitions.

The values for log ratios tend to be smaller than comparable ratios presented in Frenette, Green and Milligan for the all-age population while the Gini coefficients tend to be a bit larger for after-tax income (compare a year 2000 value of about 0.367 here to 0.322 from their study), moderately larger for total income (0.425 versus 0.357) but much, much larger for market income (0.670 versus 0.439). Our after-tax value Gini values⁴ are also somewhat higher than the values presented in Brown and Prus (2004, 2006) which use the Luxembourg Income Study (LIS) data (and hence for Canada the Survey of Labour and Income Dynamics, which is the Canadian dataset used in the LIS).⁵ This is consistent with a very strong polarization of market income and the Frenette, Green and Picot point about coverage. Hence I will turn to a brief discussion of the top end of the income distribution as well as the bottom end.

From an earlier working paper (Veall, 2006), I present a Figure 1 that illustrates one aspect of the upper end of the senior income distribution. It can be seen that market income is quite concentrated so that in 2002, about 21% of all market income accrued to the top 1% income recipients in Canada. The comparable value for the top 1% of income recipients at all ages is 13%.

Employment income is even more concentrated, with the top 1% of non-self-employment income earners receiving about 68% of all non-self-employment income and only a slightly smaller concentration for employment income. In passing I note that such a concentration of employment income makes it difficult to imagine that the aggregate effect on employment income of incremental changes in GIS or OAS clawback ranges could be very large. The threshold level

⁴ The Appendix Table A1 presents the results with capital gains excluded on the income side. These values predictable estimate a greater degree of equality but the overall trends are similar, with perhaps less evidence of a jump between 1995 and 2000.

⁵ Frenette, Green and Milligan (2007) do not include capital gains in their income tax measure. Brown and Prus (2004, 2006) include capital gains in their income measure and subtract all personal income taxes. Both subtract all personal income taxes in calculating after-tax income.

for the top 1% of senior employment income earners is about \$130,000: such individuals are well beyond the reach of these clawbacks and, as demonstrated, the aggregate labour income earned by others is small. Hence it may be appropriate that analysis of the disincentive effects of these clawbacks has concentrated on the savings disincentive (see e.g. Shillington, 1999, Kesselman and Poschmann, 2001).

Finally Tables 2, 3, 4 and 5 all examine factors that may be associated with either tail of the income distribution. (See Veall, 2007, for a more explicit discussion of poverty in this context involving a larger set of possibly associated variables.) Very briefly, it can be seen that immigration status is very important: immigrants are strongly over represented in the lower income tail because they do not qualify for senior income support. Age does not appear to be obviously important. The final two tables show that for men and particularly for women, unattached individuals (single, separated or divorced as opposed to married or in a common law relationship) are over-represented in the lower tail and under-represented in the upper tail. In addition I note that women are over-represented in the lower tail and under-represented in the upper tail. For example in 2005, about 57% of those 66 years or over were women and this same proportion holds in the middle 90% of the income distribution. In the lower 5% tail of the distribution, however, 63% are women while in the upper end, about 51% were women.

4. Conclusions

Using log ratio or Gini measures, the income distribution of the older population in Canada has become more unequal over the period 1992 to 2005, with the largest jump between 1995 and 2000. Over this same period there has been a continuing surge of income concentration: the share of the top 1% of senior income earners of after-tax income is 20%. Employment income is even more concentrated with the top 1% earning close to 70%. The low end of the income distribution is very disproportionately immigrant and somewhat disproportionately female and unattached (divorced, separated or single as opposed to married or in a common law relationship).

Table 1: Income Inequality Indices, Canada, Population 66 Years of Age or Older, Selected Years				
Market Income				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	n.a.	n.a.	n.a.	0.61631
1995	n.a.	n.a.	n.a.	0.62934
2000	n.a.	n.a.	n.a.	0.66981
2005	n.a.	n.a.	n.a.	0.66957
Total Income (including transfers)				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	1.76724	1.48646	0.52343	0.37584
1995	1.76235	1.49924	0.53757	0.36742
2000	1.90892	1.60293	0.58896	0.42504
2005	1.91944	1.61755	0.60169	0.42206
After-tax Income				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	1.52719	1.27855	0.47047	0.32015
1995	1.50164	1.27139	0.47324	0.30738
2000	1.65986	1.38454	0.51937	0.36712
2005	1.71450	1.43096	0.54186	0.37319
Notes to table: Sample sizes are 573810 for 1992, 628951 for 1995, 710409 for 2000 and 778856 for 2005. There are no log ratio values for market income because the denominator percentiles do not exceed zero.				

Table 2: Immigration and the Top and Bottom Ends Of the Senior After-tax Income Distribution in Canada, 2005				
	Year	% in bottom 5%	% in middle 90%	% in top 5%
Immigrants (landed in previous 10 years)				
	1992	51.7%	46.4%	1.9%
	1995	60.6%	38.7%	0.7%
	2000	61.7%	37.2%	1.1%
	2005	70.7%	28.6%	0.7%

Table 3: Age and the Top and Bottom Ends
Of the Senior After-tax Income Distribution in Canada, 2005

	Year	% in bottom 5%	% in middle 90%	% in top 5%
Individuals 80 years of age and over				
	1992	8.1%	88.3%	3.6%
	1995	6.3%	90.1%	3.7%
	2000	5.2%	90.7%	4.1%
	2005	5.0%	90.9%	4.1%

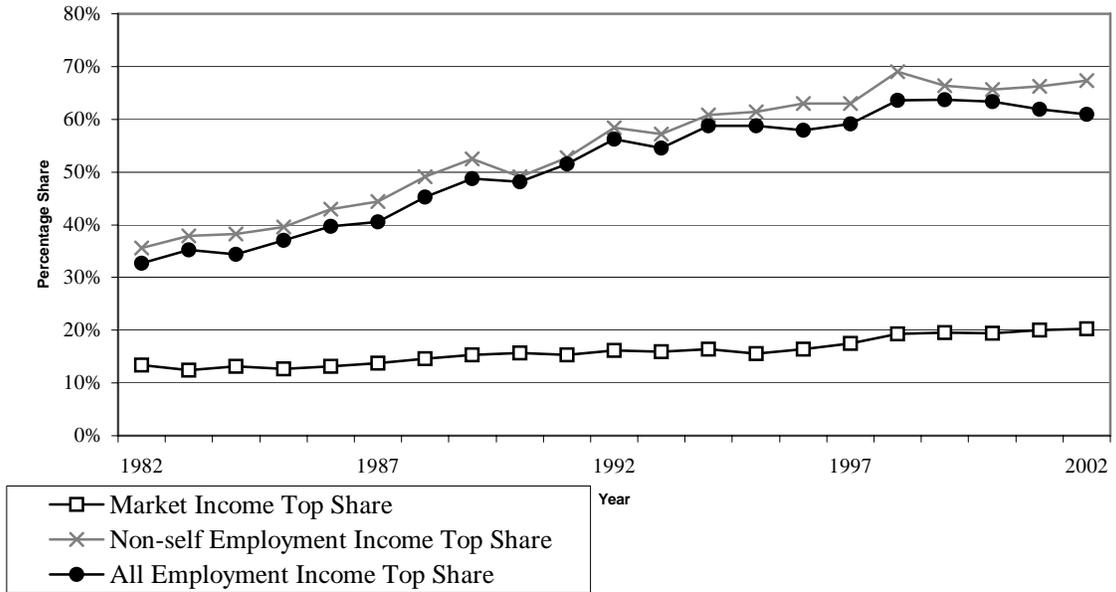
Note: Rows may not add to 100% due to rounding.

Table 4: Marital and Parental Status and the Top and Bottom Ends Of the Senior After-tax Income Distribution in Canada, Men, 2005				
	Year	% in bottom 5%	% in middle 90%	% in top 5%
Attached men with kids				
	1992	4.8%	85.9%	9.3%
	1995	4.5%	86.3%	9.2%
	2000	3.6%	88.1%	8.2%
	2005	3.2%	88.9%	8.0%
Attached men without kids				
	1992	2.6%	91.2%	6.2%
	1995	3.3%	90.5%	6.3%
	2000	4.7%	89.0%	6.3%
	2005	3.9%	89.8%	6.3%
Unattached men with kids				
	1992	6.1%	85.6%	8.3%
	1995	6.4%	84.8%	8.8%
	2000	5.9%	86.9%	7.3%
	2005	5.4%	87.5%	7.1%
Unattached men without kids				
	1992	6.8%	89.6%	3.7%
	1995	5.9%	90.5%	3.6%
	2000	5.4%	90.9%	3.8%
	2005	5.9%	90.5%	3.7%

Note: Rows may not add to 100% due to rounding.

Table 5: Marital and Parental Status and the Top and Bottom Ends Of the Senior After-tax Income Distribution in Canada, Women, 2005				
	Year	% in bottom 5%	% in middle 90%	% in top 5%
Attached women with kids				
	1992	4.3%	86.5%	9.2%
	1995	2.8%	87.7%	9.6%
	2000	2.6%	89.3%	8.0%
	2005	1.9%	90.4%	7.6%
Attached women without kids				
	1992	2.5%	91.2%	6.3%
	1995	2.7%	90.9%	6.4%
	2000	4.4%	89.6%	6.1%
	2005	3.3%	90.5%	6.2%
Unattached women with kids				
	1992	5.5%	87.2%	7.4%
	1995	5.7%	87.7%	6.6%
	2000	5.0%	89.5%	5.5%
	2005	5.1%	89.9%	5.0%
Unattached women without kids				
	1992	8.6%	89.3%	2.1%
	1995	8.2%	89.8%	2.0%
	2000	6.1%	91.5%	2.4%
	2005	7.8%	90.0%	2.3%
Note: Rows may not add to 100% due to rounding.				

Figure 1
 Top 1% Senior Incomes as Share of All Senior Income, 1982 - 2002



Appendix Table A1: Income Inequality Indices, Canada, Population 66 Years of Age or Older, Selected Years, Capital Gains Excluded from Income Distribution				
Market Income				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	n.a.	n.a.	n.a.	0.6018
1995	n.a.	n.a.	n.a.	0.6189
2000	n.a.	n.a.	n.a.	0.6366
2005	n.a.	n.a.	n.a.	0.6443
Total Income (including transfers)				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	1.72204	1.46042	0.51931	0.3594
1995	1.73489	1.48128	0.53396	0.3552
2000	1.79085	1.53204	0.57411	0.3814
2005	1.82139	1.55701	0.58747	0.3877
After-tax Income				
Year	Log(95/5)	Log(90/10)	Log(50/10)	Gini
1992	1.47521	1.25039	0.46706	0.2994
1995	1.46451	1.25045	0.47034	0.2924
2000	1.52480	1.30613	0.50650	0.3167
2005	1.59287	1.36089	0.52803	0.3310
Notes to table: Sample sizes are 573810 for 1992, 628951 for 1995, 710409 for 2000 and 778856 for 2005. There are no log ratio values for market income because the denominator percentiles do not exceed zero.				

References

- Atkinson, Anthony B. (2007), "The Distribution of Top Incomes in the United Kingdom, 1908-2000" in Anthony B. Atkinson and Thomas Piketty (eds.) *Top Incomes Over the Twentieth Century*, Oxford University Press, Oxford.
- Beach, Charles, Ross Finnie and David Gray (2001), "Earnings Variability and Earnings Instability of Women and Men in Canada", Queens University School of Policy Studies Working Paper No. 25.
- Brown, Robert L. and Stephen G. Prus (2004), "Social Transfers and Income Inequality in Old Age: A Multi-national Perspective", *North American Actuarial Journal* 8 (4): 30-36.
- Brown, Robert L. and Stephen G. Prus (2006), "Income Inequality over the Later-Life Course: A Comparative Analysis of Seven OECD Countries", SEDAP Research Paper No. 154.
- Frenette, Marc, David A. Green and Garnett Picot (2004), "Rising income inequality in the 1990s: An exploration of three data sources," Statistics Canada, Analytical Studies Research Paper Series, No. 219.
- Frenette, Marc, David A. Green and Kevin Milligan (2007) "The tale of the tails: Canadian Income Inequality in the 1980s and 1990s" *Canadian Journal of Economics*, pp. 743-764.
- Kesselman, Jonathan and Finn Poschmann (2001), "A New Option for Retirement Savings: Tax Prepaid Savings Plans," *C. D. Howe Institute Commentary*.
- Murphy, Brian, Paul Roberts and Michael Wolfson, "High-income Canadians", *Perspectives on Labour and Income*, vol. 8, no. 9, Statistics Canada, Ottawa.
- Piketty, Thomas and Emmanuel Saez (2003), "Income Inequality in the United States, 1913-1998", *Quarterly Journal of Economics*, pp. 1-39.
- Saez, Emmanuel and Michael R. Veall (2005), "The Evolution of High Incomes in North America: Lessons from Canadian Evidence", *American Economic Review*, pp. 831-849.
- Saez, Emmanuel and Michael R. Veall (2007), "The Evolution of High Incomes in Canada, 1920-2000" in Anthony B. Atkinson and Thomas Piketty (eds.) *Top Incomes Over the Twentieth Century*, Oxford University Press, Oxford
- Shillington, Richard (1999), "The Dark Side of Targeting: Retirement Saving for Low-Income Canadians," *C. D. Howe Institute Commentary*.

Veall, Michael R. (2006), “The Top Shares of Older Earners in Canada”, SEDAP Working Paper No. 156.

Veall, Michael R. (2007), “Which Canadian Seniors Are Below the Low-Income Measure, SEDAP Working Paper No. 186.