DISCUSSION PAPER SERIES

IZA

IZA DP No. 8917

The Evolution of Hourly Compensation in Canada between 1980 and 2010

Jean-Yves Duclos
Mathieu Pellerin

March 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Evolution of Hourly Compensation in Canada between 1980 and 2010

Jean-Yves Duclos
Université Laval
and IZA

Mathieu Pellerin
Université Laval

Discussion Paper No. 8917
March 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8917
March 2015

ABSTRACT

The Evolution of Hourly Compensation in Canada between 1980 and 2010*

We consider changes in the distribution of hourly compensation in Canada using confidential census data and the recent National Household Survey over the last three decades. We find that the coefficient of variation of wages among full-time workers has almost doubled between 1980 and 2010. The rapid growth of the 99.9th percentile is the main driver of that increase. Changes in the composition of the workforce explain less than 25% of the rise in wage inequality. However, composition changes explain most of the increase in average hourly compensation over those three decades, while wages stagnate within skill groups.

JEL Classification: J11, J31

Keywords: wage distribution, inequality, Canada, composition effects

Corresponding author:

Jean-Yves Duclos
Department of economics
Pavillon de Sève
Université Laval
Québec, QC, G1V 0A6
Canada
E-mail: jyves@ecn.ulaval.ca

* Financial support from the QICSS, SSHRC and the Industrial Alliance Research Chair on the Economics of Demographic Change is gratefully acknowledged. The analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics, which is part of the Canadian Research Data Centre Network (CRDCN). We thank Thomas Lemieux and Bruce Shearer for useful comments.

1 Introduction

This paper quantifies the contribution of changes in the composition of the workforce to the rise of wage inequality in Canada between 1980 and 2010. Several recent studies have documented a significant rise in inequality over that period. For instance, the Gini coefficient of market income estimated on the basis of survey data has risen by 18% between 1976 and 2009 (Figure 1, Fortin et al., 2012). Using tax data, Veall (2012) finds that the income share of the top 1% has increased from 8% to 12% between 1980 and 2010. Although the observable characteristics of the Canadian labour force have changed significantly in the last 30 years, to the best of our knowledge, no study has attempted to determine if composition changes have had a meaningful role in the rise of inequality.

Composition effects can increase inequality in at least two ways. Firstly, a rise in the dispersion of observable characteristics increases the inequality of wages, unless there is a corresponding fall in skill differentials. In 1980, 58% of full-time workers had a high school degree or less and only 13% had a college degree. By contrast, the majority of workers in 2005 had some education beyond high school, and there were approximately as many workers with a college degree (24.2%) as workers with only a high school degree (24.6%). Since the return to education has remained high (and has even increased between 1980 and 2005 — Boudarbat et al. (2010)), this higher dispersion of educational attainments should explain part of the increase in inequality.

Secondly, composition effects may increase the demographic weight of worker categories with higher within-group inequality. Lemieux (2006) finds that within-group inequality is systematically higher among educated and experienced workers in the US. Furthermore, average years of schooling and experience in the US labour force have increased substantially between 1973 and 2003. The resulting composition effects explain between 28% and 75% of the rise in residual inequality [1] among American women (between 44% and 70% among men) (Tables 1A and 1B, Lemieux, 2006). Since similar demographic changes have occured in Canada, composition shifts towards skill groups with higher within-group inequality should have affected inequality in Canada as well.

In order to quantify the importance of these two mechanisms, we use confidential data from the Census of Canada compulsory long form between 1980 and 2005 as well as confidential data from the new National Household Survey. Confidential census data contain several key demographic indicators and measures of income for 20% of the Canadian population,

[1] Suppose that the population is divided in skill groups A and B. Total variance equals $Var[Y] = E[Var[Y|X]] + Var[E[Y|X]]$, where $X \in \{A, B\}$. We use the term "residual inequality" to refer to $E[Var[Y|X]]$, and the term "within-group inequality" to designate $Var[Y|X = A]$ and $Var[Y|X = B]$.

with a new sample available every 5 years. We use hourly compensation to measure inequality and restrict our sample to full-time workers. Measurement issues arise in the construction of our data since the Census does not measure hourly compensation directly. In Section 2, we argue that census data measure hourly compensation adequately for full-time workers and that our restricted sample yields valuable insights about the evolution of inequality.

The National Household Survey contains the same information as previous censuses, and has a large sample size as well : the main difference is that the NHS is not compulsory, and thus vulnerable to non-response bias. In spite of this caveat, the results we obtain from NHS data are consistent with other data sources, and our conclusions remain the same whether we use the 2006 census or the 2011 NHS as the last year of our analysis.

We define composition as a vector of four characteristics : education, experience, gender and immigrant status. Using a simple variance decomposition, we find that between 73% and 87% of the rise in inequality can be explained by a dramatic expansion of within-group variances. The increase of within-group variances is itself driven by the rapid growth of hourly compensation in the top percentiles, as shown in Figure 1. Composition effects and changes in the returns to various measures of skill play only a small role in the rise of total inequality. Conterfactual scenarios based on the DFL method (DiNardo et al., 1996) confirm that composition effects do not fully account for the rise of top wages.

However, composition effects have had an important impact on the evolution of *average* hourly compensation over the period. Average hourly compensation grew by 15.5 percent in our sample between 1980 and 2010, well behind the 39.8 percent growth of GDP per hour worked. This weak performance masks an even more sluggish (sometimes negative) growth within skill groups. When holding the composition of the workforce constant, we find that average hourly compensation falls by 1% to 8% during the period.

The paper is organized as follows. Section 2 defines the population we study, addresses issues related to the National Household Survey, and explains how we calculate hourly compensation. Section 3 shows the evolution of hourly compensation by education, gender, immigrant status and potential experience. Section 4 presents the results. Section 5 concludes.

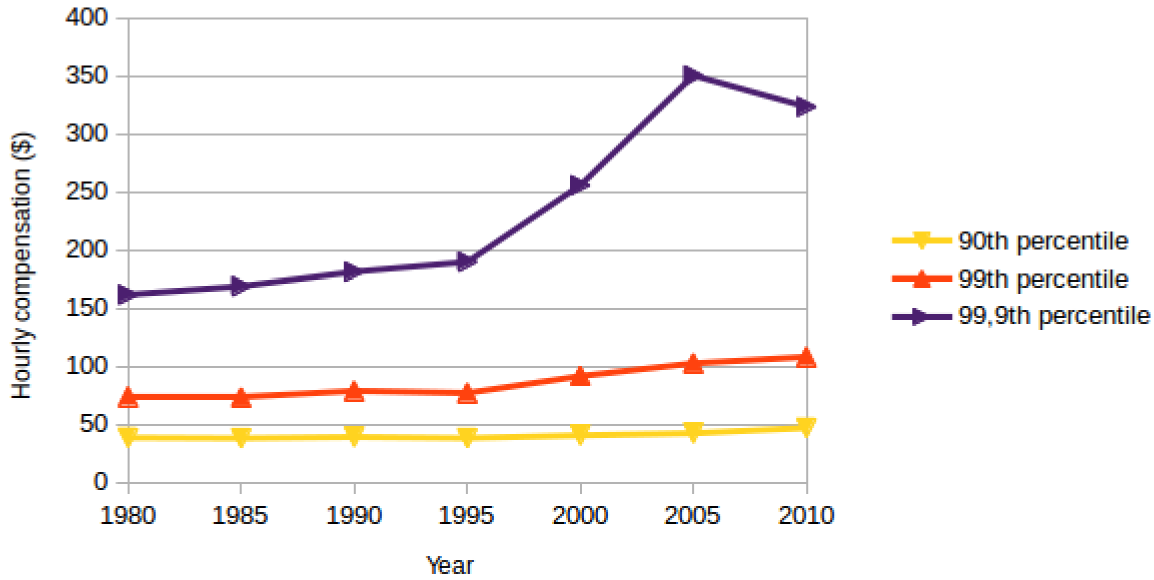

Figure 1: Hourly compensation growth, top percentiles

2 Data sources

2.1 Measurement of income in the Canadian census

The Census of Canada is conducted every five years. Up to the 2006 census, 20% of Canadian households received form 2B, known as the long form. The form contained detailed questions about housing, income, language, ethnicity, schooling and a host of other indicators. Since filling the long form was mandatory, census data suffered little from sample selection bias.

Starting with the 2011 census, the long form was abolished, a move that sparked controversy (Green and Milligan, 2010; Veall, 2010). Instead, 4.5 million households (about 30% of all private dwellings) received the National Household Survey (NHS) questionnaire : responding was optional, and the unweighted response rate was 68.6% (Statistics Canada, 2011b). Previous studies comparing census and voluntary survey data find that low and high incomes are usually underrepresented in surveys, such as the Survey of Consumer Finances and the Survey of Labour and Income Dynamics (Frenette et al., 2007). Since filling the NHS is voluntary, it is susceptible to such biases as well.

However, there are reasons to believe that NHS data may be more reliable than those of the SCF/SLID. Firstly, in order to attenuate non-response bias, Statistics Canada enumerators contacted 400,000 of the 1,200,000 non-respondents in the first wave of the survey to

Variable	**Time of measurement**
Income	Year t - 1
Weeks worked	Year t - 1
Hours worked	Year t reference week
Personnal characteristics	Year t

Table 1: Variables and time of measurement in the Census/NHS of year t

collect their answers (Statistics Canada, 2011b). Secondly, the results we obtain from NHS data, such as a slight decline of top wages between 2005 and 2010, are consistent with results obtained from reliable fiscal data (Veall, 2012).

One last important feature of census data is that in the 2006 census and the NHS, respondents had the option to let Statistics Canada access their tax records instead of self-reporting their income. 82.4% of respondents in 2006 used the new option (Statistics Canada, 2011a) (Brochu et al., 2014). Fortunately, data about income from the 2006 census are still comparable with those from previous censuses. Indeed, Lemieux and Riddell (2015) provide evidence that self-reported incomes are measured as accurately as incomes obtained from tax data. In particular, they find that there is no discrepancy between fiscal and census data for the 95^{th} and 99^{th} percentiles of total income between 1980 and 2005. Data from the 2006 census measure the 99.9^{th} percentile of total income accurately, while data from the 2001 census underestimate it. Still, the total rise of the 99.9^{th} percentile between 1980 and 2005 is the same in both data sources. Since the numerator of hourly wages, employment income, is the main component of total income, we contend that employment income is measured accurately and consistently in census data.

2.2 Hourly compensation in the Canadian census

The Census/NHS does not measure hourly compensation directly. However, we can obtain an approximation from other variables. As Table 1 shows, the questions about income and weeks worked refer to the previous year. By contrast, the question about hours refers to hours worked in the week before filling the questionnaire, the reference week. Therefore, we cannot obtain a measure of hourly compensation for respondents who worked only in year t or t - 1. Similarly, measurement error can happen for workers who are active in both years but change their labour supply over time. Although we don't know the number of hours typically worked by the respondent in the previous year, a question asks if most of the weeks worked were full-time (30 hours or above).

In order to measure hourly compensation as accurately as possible, we focus on workers who meet the following criteria :

- Workers must have worked 30 hours or more during the reference week of year t
- Workers must have worked at least 40 weeks in year t - 1.
- Most of these weeks must have been worked full-time.

We refer to workers who meet these criteria as full-time workers. Our criteria induce truncation bias, most importantly by excluding low earners who move in and out of the labour market. However, our procedure is likely to preserve high income earners such as medical doctors or senior managers, who work full-time and are unlikely to suffer long spells of unemployment.

Work activity	% of sample	Prob. 30 hours or more in t
Worked neither in t - 1 nor in t	29.0%	0.0%
Worked in t only	2.2%	45.0%
1 to 39 weeks part-time in t - 1	7.3%	16.8%
1 to 39 weeks full-time in t - 1	8.8%	51.7%
40 to 52 weeks part-time in t - 1	7.7%	24.2%
40 to 52 weeks full-time in t - 1	**45.0%**	**88.3%**

Table 2: Work activity and conditional probability of working full-time in the reference week, respondents aged 15 and above, 2010-2011 population

Table 2 shows the distribution of work activity of all respondents aged 15 and above in the NHS. Workers who have worked mostly full-time for 40 weeks or more during year t - 1 are more likely to work full-time during the reference week of year t by a wide margin. We censor hours at 84 (7×12).

Table 3 shows similar results for respondents aged between 25 and 54, in order to exclude full-time students and retirees. The sample we use in our study covers $64.6\% \times 89.9\% = 58.0\%$ of this subpopulation : this percentage goes up to 66% if we exclude inactive respondents. Thus, as long as full-time labor supply does not vary too much between years $t - 1$ and t, our findings are probably a good reflection of the wage trends faced by Canadians active in the labour market over the last three decades. Part A of the Online appendix has the same information for each of the 7 years included in this study.

Work activity	% of sample	Prob. 30 hours or more in t
Worked neither in t - 1 nor in t	12.0%	0.0%
Worked in t only	2.2%	57.8%
1 to 39 weeks part-time in t - 1	4.8%	23.9%
1 to 39 weeks full-time in t - 1	9.8%	58.9%
40 to 52 weeks part-time in t - 1	6.7%	29.9%
40 to 52 weeks full-time in t - 1	**64.6%**	**89.9%**

Table 3: Work activity and conditional probability of working full-time in the reference week, respondents aged 25 to 54, 2010-2011 population

We use the sum of wages, self-employment income and income from a non-incorporated farm business as our measure of employment income. The question concerning hours worked in the census includes hours worked in one's own business. Our choice of employment income thus ensures that both the numerator and the denominator of hourly compensation refer to the same concept of work. In our study, we use the terms hourly compensation and hourly wage interchangeably since "wages" include self-employment income. Dollar amounts are in 2010 dollars, unless stated otherwise. We use the national CPI (CANSIM Table 326-0021) to convert nominal wages into 2010 dollars.

Since some non-incorporated businesses incur losses during the year, respondents might have a negative hourly compensation. This poses no problem when the results are presented in levels. However, negative values preclude the use of the logarithm function. Therefore, all of our decompositions based on the log of wages exclude values below 1$ per hour. In part B of the Online appendix, we show that excluding negative wages has little impact on the growth and level of inequality.

We also show in part C of the Online appendix that the ratio of employment income to total income is fairly stable through time, even when the ratio is broken down by income percentile. In particular, this ratio increases little in the top percentiles. Therefore, the rise in top wages does not appear to be a spurious trend caused, *inter alia*, by business owners paying themselves wages instead of capital income.

3 Descriptive statistics

Year	$n =$	Mean	Std. dev.	CV ($\frac{\sigma}{\mu}$)	p_{10}	p_{50}	p_{90}	p_{99}	$p_{99.9}$
1980	1,379,610	23.2	16.6	0.71	8.9	20.9	38.7	74.1	161.9
1985	1,406,217	22.5	17.3	0.77	7.4	20.3	38.2	73.4	169.1
1990	1,623,128	23.1	19.6	0.85	7.4	20.8	39.3	79.0	181.6
1995	1,637,213	22.4	20.4	0.91	6.6	20.0	38.4	77.2	189.9
2000	1,847,926	23.8	27.3	1.15	7.1	20.5	40.7	91.8	255.9
2005	2,049,233	24.6	37.4	1.52	6.3	20.4	42.5	102.8	350.5
2010	2,147,883	26.8	36.4	1.36	6.6	22.4	47.2	108.2	323.6

Table 4: Evolution of the hourly wage distribution among full-time workers

Table 4 reports statistics on the evolution of hourly wages among full-time workers between 1980 and 2010. The coefficient of variation (CV) increased sharply starting from 1995. Mean hourly compensation grew by 15.5% between 1980 and 2010, while the median increased by 7.1%. The 10^{th} percentile fell sharply. The 90^{th} percentile grew slowly over the period, at a rate of 0.66% per year, suggesting that the rise in inequality is caused by the upper tail of the distribution. The rapid rise of inequality starting in 1995 coincides with an 85% rise of the 99.9^{th} percentile until 2010. For the bottom percentiles, most of the growth has taken place between 2005 and 2010, a result corroborated Morissette et al. (2013) and based on the Labour Force Survey.

The 1^{st} percentile is not shown in Table 4 since it remains stable at zero. Respondents who work for a family business without formal arrangements might report zero income but sufficient hours to be included in our sample. We show in sections B and E of the Online appendix that rising inequality in our sample is driven by increasing compensation at the top rather than negative or zero values. In particular, when the observations above the 99.9^{th} percentile are removed, the growth of the coefficient of variation between 1980 and 2010 falls from 90% to 32%.

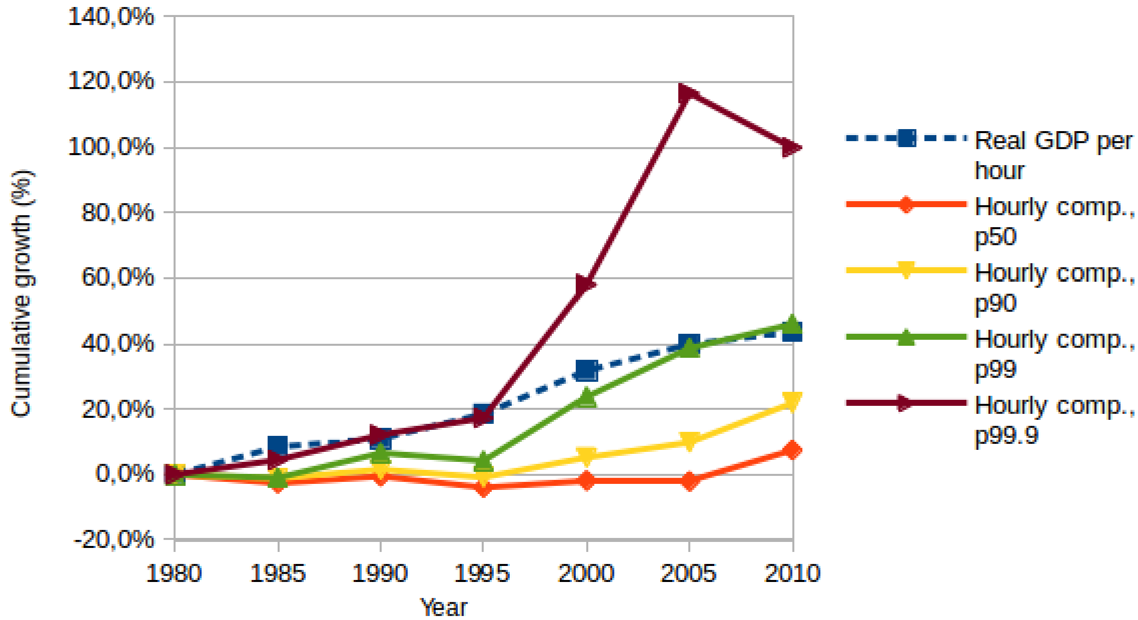

Figure 2: Hourly compensation among full-time workers and real GDP per hour, 1980-2005. Real GDP per hour obtained from FRED, serie CANRGDPH.

In order to put the values of Table 4 into context, Figure 2 compares the evolution of hourly compensation to the evolution of average labour productivity, as measured by real GDP per hour worked. The cumulative percentage growth is normalized to start at 0. Three facts stand out from Figure 2. Firstly, the 50^{th}, 90^{th} and 99^{th} percentiles moved roughly in tandem until 1995, when the 99^{th} percentile began to grow much more rapidly than the 90^{th}. Secondly, the cumulative growth of the 99^{th} percentile over the period is roughly equal to labour productivity growth : even those workers located at the 99^{th} percentile barely kept pace with productivity growth. Finally, the 99.9^{th} percentile grew at the same rate as productivity until 1995, before increasing dramatically and falling slightly after the great recession of 2008.

The next sub-sections detail the evolution of the hourly wage according to different characteristics and quantify the changes that occurred between 1980 and 2010 regarding the composition of the full-time work force.

3.1 Education

Workers are grouped in five categories : no high school degree, high school degree only, some college (this includes vocational diplomas and CEGEP[2]), bachelor's degree only and completed graduate degree. Table 5 shows that full-time workers are about two times more likely to hold a college degree (BA or higher) in 2010 than in 1980. It also shows that the high school dropout status among full-time workers fell from 37.3% to 9.1%, with a more sudden drop between 2000 and 2005.

Year	No degree	HS	Some coll.	BA	>BA
1980	37.3%	20.9%	29.1%	7.8%	5.0%
1985	33.3%	21.8%	29.6%	9.6%	5.7%
1990	25.9%	24.2%	32.6%	10.9%	6.4%
1995	21.3%	23.6%	34.7%	13.0%	7.4%
2000	19.1%	23.2%	36.1%	14.0%	7.7%
2005	11.6%	24.6%	39.6%	15.4%	8.8%
2010	9.1%	22.9%	40.1%	17.9%	10.1%

Table 5: Evolution of education status among full-time workers

Table 6 shows that average hourly compensation fell in 2 categories out of 5 between 1980 and 2010, and in 4 categories out of 5 if we use 2005 as the endpoint. Even for BA holders, the group with the fastest wage growth, average hourly compensation increased by 8.7% over the period, versus 15.5% among all full-time workers. Together, Tables 5 and 6 suggest that rising educational attainments have compensated for sluggish wage growth within skill groups.

Table 7 shows the evolution of the coefficient of variation of hourly compensation. Inequality growth was substantial within each group but was highest among college educated workers. The coefficient of variation does not increase monotonically with education, but rather seems to have a U-shaped relation, a finding that differs from studies based on US data (Lemieux, 2006)[3]. This might be caused by higher measurement error among full-time workers with less education, especially if their labour supply varies more over time.

[2] The CEGEP is an institution specific to Quebec. Students can earn either a terminal 3-year diploma that is the equivalent of an associate degree, or a 2-year degree which typically leads to undergraduate studies.

[3] Strictly speaking, Lemieux (2006) uses the log of wages. Calculating the variance of the log with Canadian data yields the same U-shaped relation as the coefficient of variation.

Year	No degree	HS	Some coll.	BA	>BA
1980	19.5	21.2	24.0	32.0	40.7
1985	18.5	20.5	22.6	30.8	39.4
1990	18.3	20.6	23.2	30.8	39.3
1995	17.3	19.7	22.0	28.8	37.0
2000	17.7	20.2	22.9	31.7	39.3
2005	16.7	19.9	23.1	32.9	40.5
2010	17.7	21.0	24.8	34.8	41.6
Total growth, 1980-2005	-14.5%	-6.1%	-3.8%	+2.7%	-0.4%
Total growth, 1980-2010	-9.3%	-0.9%	+3.5%	+8.7%	+2.4%

Table 6: Mean hourly compensation among full-time workers, by year and education

Year	No degree	HS	Some coll.	BA	>BA
1980	0.71	0.64	0.60	0.67	0.71
1985	0.80	0.69	0.60	0.74	0.76
1990	0.95	0.81	0.69	0.79	0.78
1995	1.18	0.78	0.73	0.86	0.86
2000	0.96	0.93	0.81	1.38	1.22
2005	1.72	1.24	1.05	1.60	1.75
2010	1.24	1.19	0.96	1.53	1.42
Total growth, 1980-2005	143.6%	92.0%	75.6%	139.8 %	145.9 %
Total growth, 1980-2010	75.8%	83.9%	59.3%	129.7 %	99.3 %

Table 7: Coefficient of variation of hourly compensation, by year and education

Using either 1980 or 2005 as a reference year for within-group variances, we find that weighting the variances with the composition of 2005 rather than 1980 results in a higher variance in both cases. Repeating this exercise using 1980 and 2010 shows that the composition of 2010 also results in the highest residual inequality. It follows that rising educational attainments should explain part of the increase in residual inequality.

3.2 Potential experience

Year	Potential experience (Exp) 0-9	10-19	20-29	30-39	40+
1980	29.8%	26.7%	19.4%	16.4%	7.7%
1985	25.2%	30.6%	21.8%	15.4%	7.0%
1990	22.4%	31.6%	25.3%	14.7%	6.0%
1995	18.6%	30.3%	29.7%	16.2%	5.1%
2000	19.1%	26.4%	31.2%	18.3%	5.0%
2005	19.9%	23.6%	29.7%	20.5%	6.2%
2010	19.6%	23.1%	26.9%	22.8%	7.8%

Table 8: Evolution of potential experience among full-time workers

We define potential experience as $Exp = Age - Years\ of\ schooling - 6$. Since the 2006 census and the NHS do not measure years of schooling, we allocate them based on the highest degree completed and demographic characteristics. Part D of the Online appendix details the procedure. Table 8 shows the evolution of potential experience in our sample. Almost 27% of full-time workers are part of the category with 20-29 years of experience in 2010, up from 19.4% in 1980. On the other hand, the 0-9 category underwent a 10% decline over the same period.

Table 9 shows that average compensation grew in every category between 1980 and 2010. Interesingly, hourly compensation growth was fastest in the categories that also experienced the biggest increase in demographic weight. In particular, growth in the 20-29 category was faster than average compensation growth among all full-time workers. The interaction of this trend and population aging explains a significant proportion of compensation growth between 1980 and 2010. This pattern differs from the trends in Tables 5 and 6, where rising educational attainments offset stagnating wages within each category. Note that the unconditional wage gap between younger and older workers has expanded over the period.

Year	0-9	10-19	20-29	30-39	40+
1980	19.6	25.1	25.7	24.9	20.6
1985	18.2	24.04	25.0	24.2	19.9
1990	18.7	24.0	25.8	24.7	20.1
1995	17.6	22.8	24.8	24.0	19.5
2000	18.6	24.7	25.8	25.3	20.6
2005	18.4	25.5	27.3	26.6	21.7
2010	20.4	27.8	30.0	28.5	23.6
Total growth, 1980-2005	-6.1%	1.5%	6.1%	7.0%	5.4%
Total growth, 1980-2010	4.5%	10.7%	16.7%	14.7%	14.6%

Table 9: Mean hourly compensation among full-time workers, by year and potential experience

Year	0-9	10-19	20-29	30-39	40+
1980	0.59	0.67	0.74	0.76	0.87
1985	0.64	0.68	0.77	0.85	1.04
1990	0.65	0.73	0.81	0.99	1.42
1995	0.74	0.75	0.78	1.11	1.81
2000	0.84	0.98	1.00	1.39	2.17
2005	0.82	1.24	1.41	1.80	2.58
2010	0.74	0.95	1.27	1.43	2.79
Total growth, 1980-2005	38.2%	84.9%	90.9%	137.6%	197.9%
Total growth, 1980-2010	25.8%	42.8%	71.0%	89.6%	222.4%

Table 10: Coefficient of variation of hourly compensation, by year and potential experience

Finally, Table 10 shows that within-group inequality increases monotically with experience, in agreement with a model of human capital accumulumation in which workers invest in on-the-job training at differing rates (Mincer, 1974). The relationship between experience and inequality is much more convex in 2005 and 2010 than in 1980. Regardless of the reference year for within-group variances, the aging of the workforce increases residual inequality.

3.3 Gender

Year	Women	Men
1980	31.4%	68.6%
1985	34.2%	65.8%
1990	37.8%	62.2%
1995	38.7%	61.4%
2000	40.2%	59.8%
2005	41.1%	58.9%
2010	42.6%	57.4%

Table 11: Evolution of gender among full-time workers

Table 11 shows that the proportion of women among full-time workers rose from 31.4% in 1980 to 42.6% in 2010. Since women have lower wages than men, as shown in Table 12, their entry in the labour market can increase between-group inequality. On the other hand, Table 12 also shows that inequality is lower among women and grew more slowly for them during the period, reducing the growth of residual inequality. Since residual inequality is the biggest component of total inequality, the increased labour force participation of women is likely to have curbed the growth of the variance of wages. Lemieux and Riddell (2015) find that men represented 81.2% of the top 1% in 2005, an explanation for why within-group inequality is much lower among women.

	Average		CV	
Year	Women	Men	Women	Men
1980	17.8	25.7	0.55	0.71
1985	17.9	24.9	0.57	0.79
1990	18.9	25.7	0.62	0.89
1995	19.2	24.5	0.64	0.98
2000	20.3	26.1	0.77	1.25
2005	21.2	27.0	0.90	1.70
2010	23.6	29.1	0.81	1.54
Total growth, 1980-2005	19.1%	5.1%	63.6%	139.4%
Total growth, 1980-2010	32.6%	13.2%	47.3%	116.9%

Table 12: Mean and coefficient of variation of hourly compensation, by year and gender

3.4 Immigrant status

Year	Average		CV	
	Natives	Immigrants	Natives	Immigrants
1980	23.1	23.6	0.71	0.72
1985	22.4	23.0	0.75	0.85
1990	23.1	23.5	0.80	1.02
1995	22.5	22.1	0.86	1.09
2000	23.8	23.6	1.11	1.27
2005	24.9	23.5	1.45	1.77
2010	27.1	25.6	1.34	1.42
Total growth, 1980-2005	7.8%	-0.4%	104.2%	145.8%
Total growth, 1980-2010	17.3%	8.5%	88.7%	97.2%

Table 13: Mean and coefficient of variation of hourly compensation, by year and immigrant status

The share of immigrants in the full-time workforce rose slightly between 1980 and 2010, from 21.1% to 23.3%. Table 13 shows a reversal in the relative position of natives and immigrants : starting in 1995, hourly compensation becomes higher for natives than for immigrants. This reversal can be explained by other observable characteristics : for instance, native Canadians are aging faster than immigrants and earn higher wages as a result of their higher experience. Boudarbat and Lemieux (2014) also find that immigrants now get lower returns on their education and are increasingly penalized if their language skills are lacking.

Table 13 shows that inequality grew faster among immigrants. This trend is consistent with the sharp decline that occured at the bottom of the wage distribution of immigrants, another fact documented in Boudarbat and Lemieux (2014).

4 Results

4.1 Variance decomposition

The following is true for any pair of random variables :

$$Var[Y] = E_X\left[Var[Y|X]\right] + Var_X\left[E[Y|X]\right] \tag{1}$$

If Y in equation (1) measures income and X are observable characteristics, the first term on the right-hand side corresponds to residual inequality and the second term, to between-group inequality. We use this formula to compute the respective contributions of residual and between-group inequality to the evolution of $Var[Y]$ between two periods :

$$\begin{aligned} Var[Y_t] - Var[Y_s] = &\Big(E_{X_t}\big[Var[Y_t|X]\big] - E_{X_s}\big[Var[Y_s|X]\big]\Big) \\ &+ \Big(Var_{X_t}\big[E[Y_t|X]\big] - Var_{X_s}\big[E[Y_s|X]\big]\Big). \end{aligned} \tag{2}$$

X_t denotes the distribution of observable characteristics at time t, and $Y_t|X$ the conditional distribution of wages at time t for a given vector of characteristics. For instance, $E_{X_t}\big[Var[Y_s|X]\big]$ is the mean of the within-group variances at time s, weighted by the shares of the groups at time t. Finally, we divide each of the right-hand-side components into a composition and a structural effect:

$$\begin{aligned} Var[Y_t] - Var[Y_s] = &\Big(E_{X_t}\big[Var[Y_t|X]\big] - E_{X_t}\big[Var[Y_s|X]\big]\Big) && \text{(I)} \\ &+ \Big(E_{X_t}\big[Var[Y_s|X]\big] - E_{X_s}\big[Var[Y_s|X]\big]\Big) && \text{(II)} \\ &+ \Big(Var_{X_t}\big[E[Y_t|X]\big] - Var_{X_t}\big[E[Y_s|X]\big]\Big) && \text{(III)} \\ &+ \Big(Var_{X_t}\big[E[Y_s|X]\big] - Var_{X_s}\big[E[Y_s|X]\big]\Big). && \text{(IV)} \end{aligned}$$

(I) + (II) gives the contribution of residual inequality, while (III) + (IV) represents the contribution of between-group inequality. (I) captures changes in residual inequality between s and t that are caused solely by the evolution of within-group variances : the composition of the workforce is held fixed. One important point from Section 3 is that wage inequality increased within every education level, every experience category, and so forth. This pervasive rise in within-group variances is quantified by term (I). By contrast, (II) captures the interaction of demographic changes and heteroscedasticity. Within-group variances are set to a baseline level, and the composition of the workforce varies over time. For instance, if educated workers' wages are more dispersed, the impact of increased schooling on residual inequality will be captured by this term. (III) quantifies the impact of changing skill differentials, such as the return to schooling or experience. Finally, (IV) represents the contribution of a change in the dispersion of observable characteristics when skill returns are fixed.

Conceptually, (I) and (III) quantify the impact of changes in the wage structure on residual and between-group inequality, respectively. The skill distribution of year t is used as a counterfactual and the wage structure is allowed to vary. Similarly, (II) and (IV) capture the influence of composition effects on residual and between-group inequality. The composition of the workforce varies while the wage structure of year s serves as a counterfactual.

To compute the decomposition, we drop every observation with hourly compensation below 1$ and use the log of wages in order to remove the effect of a changing mean on the variance. Respondents are allocated to a cell that corresponds to their gender, immigrant status, potential experience and education categories (the categories used are the same as in Section 3). Table 14 presents the results when the skill distribution of 1981 ($t = 1981$) and the wage structure of 2006 ($s = 2006$) are used.

Total variation = 100%			
Residual inequality = 84.5%		Between-group inequality = 15.5%	
Structure (I)	Composition (II)	Structure (III)	Composition (IV)
76.9%	7.7%	12.9%	2.6%

Table 14: Variance decomposition, skill distribution of 1981, wage structure of 2006

As foreshadowed by Section 3, (I), the component linked to within-group variances is the dominant factor. Composition effects account for only 7.7%/84.5% = 9.1% of the increase in residual inequality. Since wage inequality is much lower among women in the wage structure of 2006, the entry of women in the labour force over the period offsets the impact of rising educational attainments and experience levels on residual inequality. (III) is quantitatively important because the wages of dropouts and younger workers fell substantially between 1981 and 2006. The wage gap between natives and immigrants, which was mostly absent in 1981, expanded substantially over the period, also contributing to the between-group, structural component. Table 15 shows that using the wage structure of 2011 generates roughly the same results.

Total variation = 100%			
Residual inequality = 85.5%		Between-group inequality = 14.5%	
Structure (I)	Composition (II)	Structure (III)	Composition (IV)
73.6%	11.9%	12.4%	2.2%

Table 15: Variance decomposition, skill distribution of 1981, wage structure of 2011

Table 16 shows the same decomposition using the wage structure of 1981 ($s = 1981$, $t = 2006$). The wage structure in 1981 showed much less heteroscedasticity, which explains why composition effects play no role in the rise of residual inequality. Again, the biggest contribution comes from a dramatic rise in within-group variances, and using either 2011 or 2006 as the endpoint does not affect the results. Since more women were part of the labour force in 2006/2011 than in 1981, paying 2006/2011 workers according to the wage

Total variation = 100%			
Residual inequality = 84.5%		Between-group inequality = 15.5%	
Structure (I)	Composition (II)	Structure (III)	Composition (IV)
87.1%	-2.5%	24.8%	-9.4%

Table 16: Variance decomposition, skill distribution of 2006, wage structure of 1981

structure of 1981, with its larger gender wage gap, results in a higher contribution of (III). Also, educational attainments and potential experience are more dispersed in 2006/2011, which magnifies the impact of expanding wage gaps and results in a higher contribution of (III). Finally, (IV) shows that if the wage structure of 1981 would have prevailed, demographic changes such as increased educational attainments among younger cohorts would have lowered between-group inequality.

Total variation = 100%			
Residual inequality = 85.5%		Between-group inequality = 14.5%	
Structure(I)	Composition (II)	Structure (III)	Composition (IV)
85.8%	-0.3%	26.5%	-12.0%

Table 17: Variance decomposition, skill distribution of 2011, wage structure of 1981

Although our variance decomposition incorporates the effect of top wages, they are only based on the first two moments of the wage distribution. The next section addresses this limitation by focusing on several counterfactual percentiles.

4.2 Conterfactual percentiles

Suppose we would like to know the wage distribution that would have prevailed in 2006 if 2006 workers had been paid with the 1981 wage structure. The simplest way to obtain such a distribution is to re-weight the distribution of 1981 characteristics in order to make it identical to the 2006 distribution. Formally, if X denotes workers' characteristics, we want to find Ψ_i such that :

$$P(X = x_i|Year = 1981) \times \Psi_i = P(X = x_i|Year = 2006) \tag{3}$$

for each skill group i. Since we have grouped workers into a number of mutually exclusive cells, the computation of Ψ_i is trivial :

$$\Psi_i = \frac{P(X = x_i|Year = 2006)}{P(X = x_i|Year = 1981)} \tag{4}$$

Equation 4 is a particular case of the DFL (DiNardo, Fortin, and Lemieux, 1996) method. Using this formula, we compute selected counterfactual percentiles of the wage distribution, using both 1981 and 2011 as reference years for the composition of the workforce. Now, suppose that the evolution of the distribution of wages between 1981 and 2011 is caused only by composition effects. The graphs of the counterfactual percentiles will form two horizontal lines that will bound the graph of the percentile under consideration, since the composition of the workforce is held constant in counterfactual scenarios. On the contrary, if changes in the distribution are caused solely by changes in the wage structure, the graphs of the counterfactual percentiles are going to be superimposed on the graph of the observed percentile. Thus, a large gap between the counterfactual percentiles and the true percentile indicates that composition effects drive the evolution of wages at this percentile.

Figure 3 shows that average compensation stagnates or falls by one dollar per hour over the period if the skill distribution is held constant (at 1981 or 2011). The fall is steeper for the median, as can be seen from Figure 4. Median compensation drops between 1.7 and 2.8 dollars per hour when we fix the workforce's composition. Perhaps more surprisingly, the situation is similar for the 90^{th} percentile : a 22% percent gain over the period essentially vanishes when composition effects are removed (Figure 5).

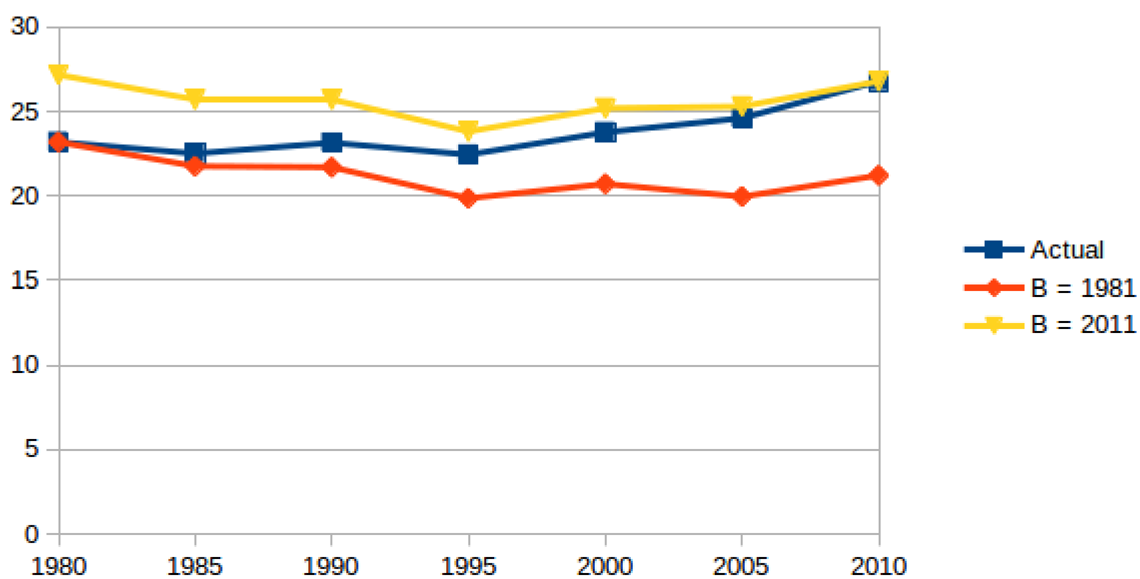

Figure 3: Average hourly compensation among full-time workers, actual and conterfactual with composition of year B

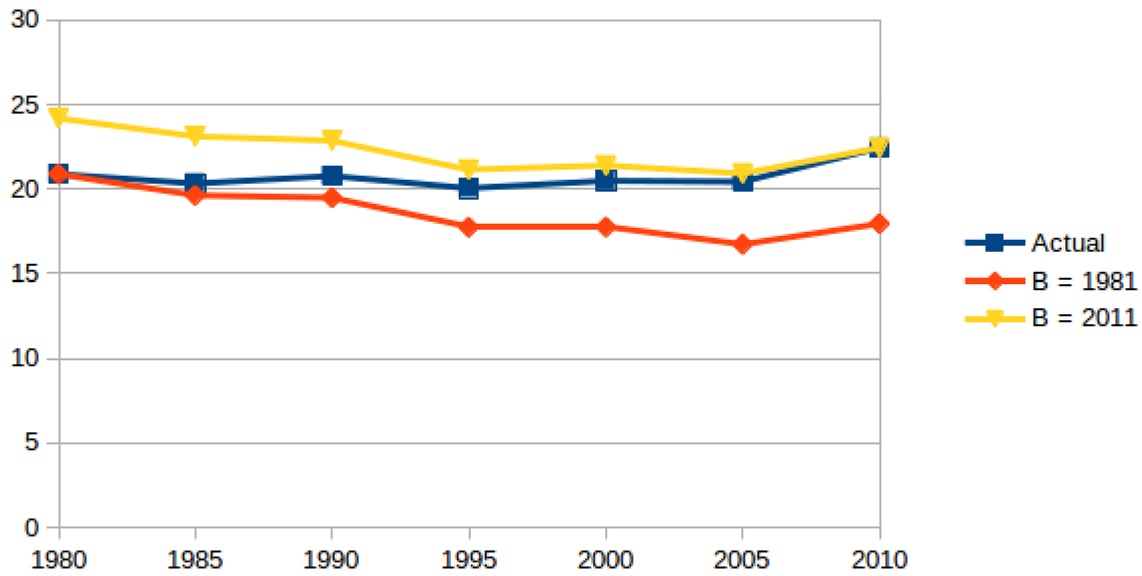

Figure 4: Median hourly compensation among full-time workers, actual and conterfactual with composition of year B

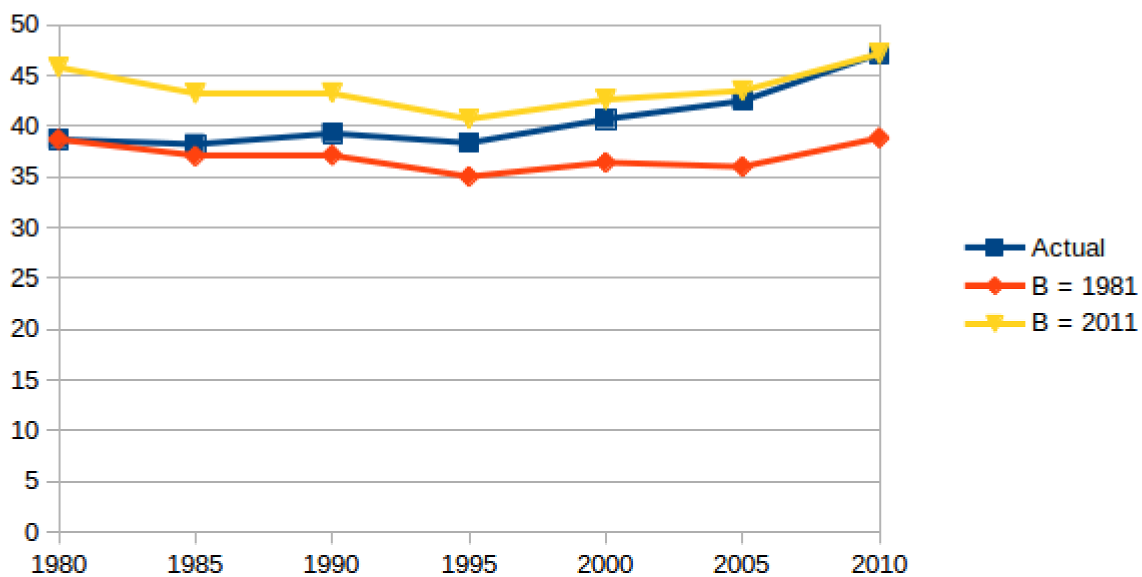

Figure 5: Hourly compensation among full-time workers, 90^{th} percentile, actual and conterfactual with composition of year B

Education	99^{th} percentile 1980	2005
No degree	53	52
HS degree	56	63
Trade certificate	56	60
CEGEP, non-university diploma	57	69
University certificate $<$ BA	71	88
B.A.	94	140
University certificate $>$ BA	109	151
Medecine	185	250
Master's degree	98	178
Ph.D.	101	158

Table 18: Evolution of the 99^{th} percentile of hourly compensation by education group, 2005 constant dollars

Figures 6 and 7 show the evolution of the 99^{th} and 99.9^{th} percentiles, respectively. Interestingly, the 99^{th} percentile and the 99.9^{th} percentile both grow more slowly in counterfactual scenarios, especially when the skill distribution is held constant at its 1981 level. Table 18 provides the intuition behind this result. If we look at the 99^{th} percentile by education group, we see that the rapid increase of top wages essentially happened among full-time workers with a bachelor's degree or above. When the skill distribution of 1981 is used, the categories that generate much of the rise in top incomes are under-represented in the data. Since workers in 1981 were half as likely as 2006 workers to hold a college degree, the effect of the resulting difference is quite important. However, when the composition of 2011 is used, the counterfactual and actual 99.9^{th} percentiles move in tandem, indicating that composition effects begin to lose their explanatory power at this level in the wage distribution.

In summary, composition effects account for a diminishing but still substantial part of hourly compensation growth as we look towards higher percentiles. In particular, they explain a large proportion of the growth of the 90^{th} and the 99^{th} percentile, a fact not visible from the variance decompositions. Section 4.1 showed that composition effects explain a relatively small portion of the rise in inequality; this section shows that this is because they do not explain the rise of very high wages. The substantial growth of within-group variances appears indeed to be caused by the growth in compensation among the top wages (mostly the top 0.1%), the change in the rest of the distribution being largely explained by variations in observable characteristics.

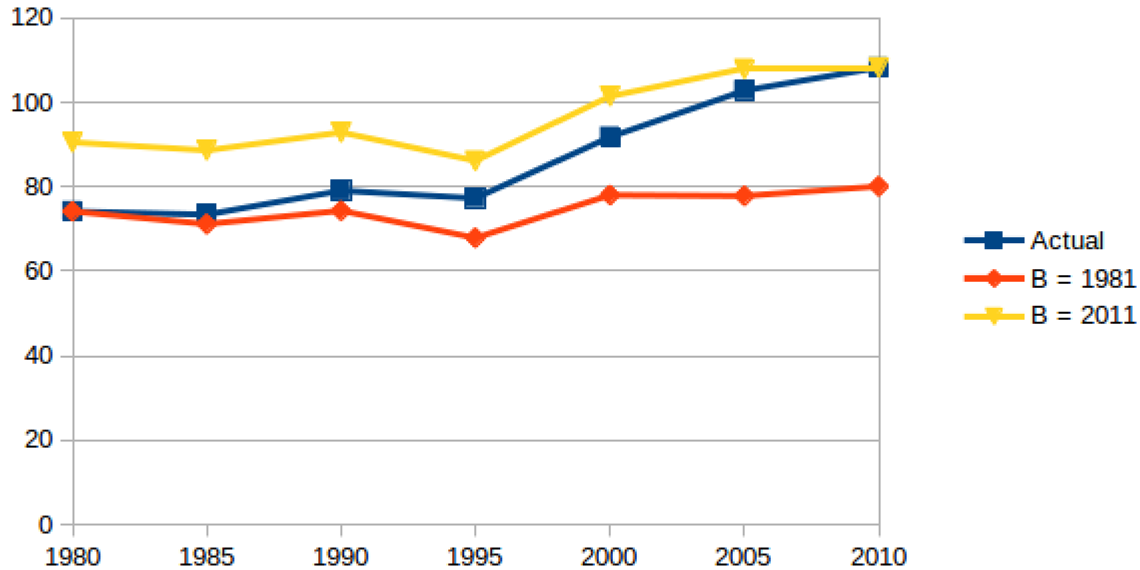

Figure 6: Hourly compensation among full-time workers, 99^{th} percentile, actual and conterfactual with composition of year B

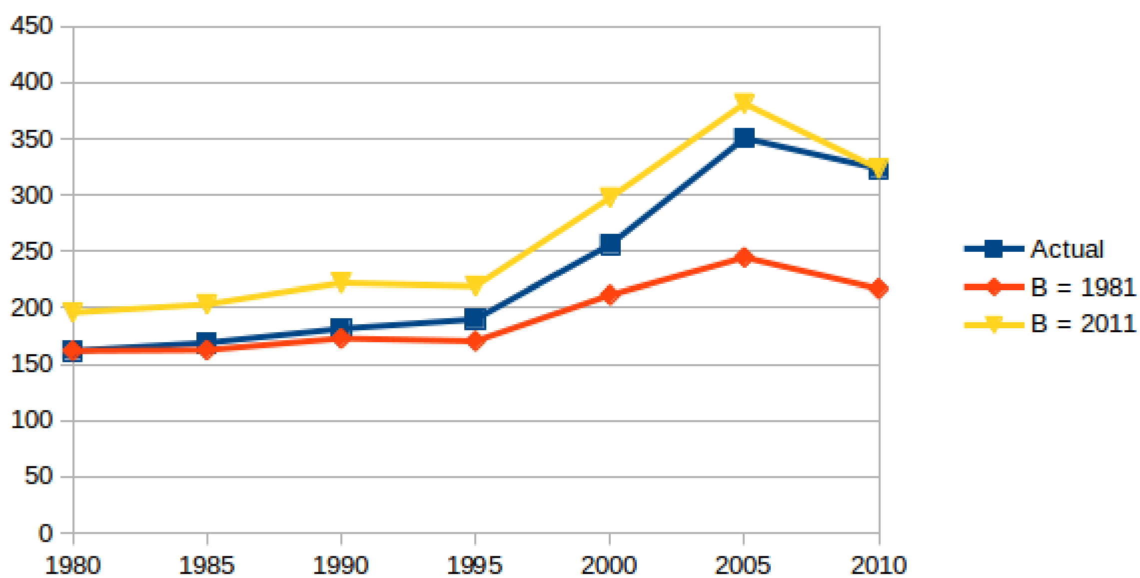

Figure 7: Hourly compensation among full-time workers, 99.9^{th} percentile, actual and conterfactual with composition of year B

5 Conclusion

This paper uses confidential census data and the recent NHS to identify and explain features of the rise of wage inequality in Canada between 1980 and 2010.

1. Wages within educational and potential experience groups have stagnated between 1980 and 2010.

2. Hourly compensation growth among full-time workers is driven largely by the aging of the workforce and by rising educational attainments. Once we remove the wage effects of changes in the composition of the labor force (the so-called "composition effects"), average hourly compensation falls by 1% to 8% over the period.

3. The aging of the workforce and the concomitant growth of potential experience has offset to a certain extent slow wage growth within skill groups, but this aging effect has arguably run its course.

4. Between 75% and 85% of the increase in wage inequality between 1980 and 2010 is explained by the increase of within-skill-groups inequality, that is, by other factors than composition effects or rising skill differentials.

5. The growth of percentiles up to the 99^{th} is mostly driven by changes in the composition of the workforce.

6. An immediate corollary of (4) and (5) is that the growth of within-group variances, the most important factor in the growth of inequality, must be caused by the growth of percentiles higher than the 99^{th}. Section 4.2 shows that composition effects do not account for the rise of wages in the top 0.1% of the distribution.

7. Section E of the Online appendix shows that the growth of inequality between 1980 and 2010 falls from 90% to 32% when the observations in the top 0.1% are removed, confirming our intuition that these observations drive the growth of inequality.

More generally, the slow growth of wages within skill groups appears to be caused by deep macroeconomic trends. Karabarbounis and Neiman (2014) find that the labour share of income is decreasing in most countries and industries since 1980, and that half of that decline is caused by a fall in the relative price of investment goods. The fact that the decline of the labour share is a worldwide phenomenon suggests that slow wage growth (relative to output) is not caused by circumstances specific to Canada, and that further research is needed to pin down the exact forces behind the decline of the labour share of income.

References

Boudarbat, Brahim and Thomas Lemieux (2014), "Why are the Relative Wages of Immigrants Declining? A Distributional Approach." *Industrial and Labor Relations Review*.

Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell (2010), "The Evolution of the Returns to Human Capital in canada, 1980–2005." *Canadian Public Policy*, 36, 63–89.

Brochu, Pierre, Louis-Philippe Morin, and Jean-Michel Billette (2014), "Opting or Not Opting to Share Income Tax Information with the Census: Does It Affect Research Findings?" *Canadian Public Policy*, 40, 67–83.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, 64, 1001–1044.

Fortin, Nicole, David A. Green, Thomas Lemieux, Kevin Milligan, and W. Craig Riddell (2012), "Canadian Inequality: Recent Developments and Policy Options." *Canadian Public Policy*, 38, 121–145.

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/Revue canadienne d'économique*, 40, 734–764.

Green, David A. and Kevin Milligan (2010), "The Importance of the Long Form Census to Canada." *Canadian Public Policy*, 36, 383–388.

Karabarbounis, Loukas and Brent Neiman (2014), "The Global Decline of the Labor Share." *The Quarterly Journal of Economics*, 61, 103.

Lemieux, Thomas (2006), "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 461–498.

Lemieux, Thomas and W. Craig Riddell (2015), "Top Incomes in Canada: Evidence from the Census." Technical report, University of British Columbia.

Mincer, Jacob A. (1974), *Schooling, Experience and Earnings*. Columbia University Press.

Morissette, René, Garnett Picot, and Yuqian Lu (2013), "The Evolution of Canadian Wages over the Last Three Decades." Technical report, Statistics Canada.

Statistics Canada, Official Document (2011a), "National Household Survey Income Reference Guide." Technical Report 99-014-X2011006, Statistics Canada, URL http://www12.statcan.gc.ca/nhs-enm/2011/ref/guides/99-014-x/99-014-x2011006-eng.pdf.

Statistics Canada, Official Document (2011b), "National Household Survey User Guide." Technical Report 99-001-X, Statistics Canada, URL http://www12.statcan.gc.ca/nhs-enm/2011/ref/nhs-enm_guide/99-001-x2011001-eng.pdf.

Veall, Michael R (2010), "2b or Not 2b? What Should Have Happened with the Canadian Long Form Census? What Should Happen Now?" *Canadian Public Policy*, 36, 395–399.

Veall, Michael R. (2012), "Top Income Shares in Canada: Recent Trends and Policy Implications." *Canadian Journal of Economics/Revue canadienne d'économique*, 45, 1247–1272.

Online appendix

The online appendic is hosted at the following address : http://www.cedia.ca/sites/cedia.ca/files/onlineappendix_xl.xlsx

Part A contains tables detailing the labour supply of all respondents in the Census, for every year.

Part B compares the evolution of inequality when negatives wages are excluded from our sample to the results obtained from the full sample.

Part C shows the proportion of employment income in total income for each percentile of total income.

Part D has the details concerning the allocation of years of schooling in the 2006 census and the NHS.

Part E compares the evolution of inequality when wages in the top 0.1% are excluded from our sample to the results obtained from the full sample.