The gender wage gap among recent post-secondary graduates in Canada: a distributional approach

Brahim Boudarbat *École de relations industrielles, Université de Montréal*

Marie Connolly *Département des sciences économiques, École des sciences de la gestion, Université du Québec à Montréal*

Abstract. Using the National Graduates Survey, we examine the trends in the gender wage gap among recent post-secondary graduates in Canada between 1988 and 2007. Female graduates earn on average 6–14% less than males during the period two to five years after graduation. Decompositions show that observable personal characteristics and job attributes can explain only a small portion of the wage gap. We also find that men earn more than women at every point of the distribution. Interestingly, the wage difference shrank in the lower half of the distribution in recent years, while it increased in the upper half.

L'écart salarial entre les sexes chez les nouveaux diplômés postsecondaires: une approche distributionnelle. À l'aide de l'Enquête nationale auprès des diplômés, nous examinons les tendances dans l'écart salarial entre les sexes chez les nouveaux diplômés postsecondaires au Canada entre 1988 et 2007. Les femmes diplômées gagnent en moyenne de 6 à 14% moins que les hommes durant les 2 à 5 ans suivant l'obtention du diplôme. Des décompositions montrent que les caractéristiques personnelles et les attributs de l'emploi observables ne peuvent expliquer qu'une petite partie de l'écart salarial. Nous trouvons aussi que les hommes touchent plus que les femmes à chaque point de la distribution salariale. De manière intéressante, la différence de salaire a diminué dans la moitié

Boudarbat is also affiliated with CIRANO. Connolly is also affiliated with CIRPÉE and CIRANO. The authors would like to thank Statistics Canada and the Quebec Inter- University Centre for Social Statistics for access to the National Graduates Survey data, as well as Human Resources and Skills Development Canada for funding the research and providing constructive comments. Special thanks to Darren King at HRSDC. Ross Finnie, anonymous referees and seminar participants at the University of Ottawa, the SCSE and the CEA Meetings provided useful comments. The research and analysis are based on data from Statistics Canada and the opinions expressed do not represent the views of Statistics Canada. The technical appendix to this paper is available from the CJE online archive at cje.economics.ca.
Email: connolly.marie@uqam.ca

Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 46, No. 3
August / août 2013. Printed in Canada / Imprimé au Canada

0008-4085 / 13 / 1037–1065 /

inférieure de la distribution lors des années récentes, alors qu'elle a augmenté dans la moitié supérieure.

1. Introduction

According to numerous objective measures, women's conditions have vastly improved over the past decades. In Canada, the women's share of the labour force in 2009 was close to parity at 48%, compared with only 37% in 1976, while the female employment rate reached 58% in 2009, up from 42% in 1976 (Ferrao and Williams 2011). An increasing schooling attainment no doubt played an important role in generating these trends (Chaykowsky and Powell 1999). Data from the 2006 Census shows that 54% of women 25 years and above had completed post-secondary studies, compared with 58% of men. But women now outnumber men in higher education: when we look only at a younger cohort, aged 25 to 34, this percentage increases to 71% for women, compared with 62% of men.[1] Large and increasing returns to education for women have arguably been drivers for this dramatic increase in education (Christofides, Hoy, and Yang 2010; Boudarbat, Lemieux, and Riddell 2010).

Improvements in women's earnings have followed: the general consensus seems to be that male-female wage differentials have shrunk over time, but have not completely disappeared, and that the rate of convergence has slowed down (O'Neill 2003; Fortin 2005; Blau and Kahn 2007; Frenette and Coulombe 2007). Two recent studies, Baker and Drolet (2010) and Drolet (2011), document a narrowing of the gender gap in hourly wages for full-time employed workers of 7.6 percentage points from 1988 to 2008. The authors note a number of factors that can help explain this reduction: a changing composition of the workforce across cohorts, a longer job tenure and occupational changes for older workers, and an increased education and falling unionization for younger workers. But despite this progress in relative wages, a woman on average still earned only 83¢ for every dollar a man received, and it is important to note that the closing of the gap has not been equally felt over the wage distribution. Drolet (2011) reported that substantive gains have been made at the bottom of the distribution, with a narrowing of the gap reaching 11.5 percentage points, while only more modest gains can be reported at the top of the distribution. At the highest wages, the decrease in the gender gap was half of that at the lowest wages, reaching only 6.7 percentage points.

How much of the remaining wage gap can we explain by differences in various observable characteristics of the male and female workforce? And how much is left unexplained or is due to differences in returns to those characteristics?

1 Authors' calculations using Statistics Canada's '2006 Census data products – Topic-based tabulations,' available online at http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/index-eng.cfm.

Economists commonly use decomposition methods to tease out the explained and unexplained parts of the gender gap. While a growing portion of the gap seems to be explainable when the 1970s and 1980s are considered (Baker et al. 1995), a substantial fraction of the gap is left unexplained in Canada and other countries (Chaykowsky and Powell 1999; Blau and Kahn 2007). Decomposing the wage gap is a relevant exercise, as it allows us to determine if the observed gap finds support in the composition of the female workforce and, if so, what factors could explain the persistence of a gap despite large human capital improvements for women. Differences in returns (the unexplained part) could be due to differential treatments in the job market (some form of discrimination), which would be important to know because of the potentially negative effect on the competitiveness of the Canadian economy, but also to assess the effectiveness of gender equity laws that are in place for public-sector workers at the federal and multiple provincial levels, as well as for private-sector employees in Ontario and Quebec. However, other explanations, such as the presence of unobservable factors, unavailable variables, or subjective choices, cannot be ruled out, as the decomposition approaches allow the econometrician only to tell what can be explained by the factors at hand, not to identify a causal effect.

In this paper we take a distributional approach to the gender wage gap and perform decompositions not only at the mean using the standard Oaxaca-Blinder methodology (Oaxaca 1973; Blinder 1973), but also at various percentiles along the wage distribution using the unconditional quantile regression methodology recently developed by Firpo, Fortin, and Lemieux (2009) and Fortin, Lemieux, and Firpo (2010). This allows us to draw a more complete portrait of the recent situation of women in Canada by showing what mean numbers may miss: an approach based on means would inform us on the male and female wage distributions only if they were unimodal, symmetric, and with equal variance, which is likely not the case (Butcher and DiNardo 2002). To the best of our knowledge, this study is the first one taking a distributional approach to wage decompositions in Canada to explain the gender gap.[2] We study recent post-secondary graduates using the nine most recent waves of Statistics Canada's National Graduates Survey (NGS), spanning the years 1988 to 2007. This survey presents many key features that justify its use over other sources that have other advantages that the NGS does not possess, mainly a wider coverage (the whole population, not just post-secondary graduates) or a longer lag after graduation to allow for the full potential of the studies to materialize.[3] By their focus on graduates, the NGS data contain many variables on precise educational attainment (including multiple credentials) and field of study, as well as the usual job market information

2 Arulampalam, Booth, and Bryan (2007) and Booth (2007) provided a similar analysis for Europe, while Kee (2006) took a similar tack studying Australia, Albrecht, Björlund, and Vroman(2003) studied Sweden, and de la Rica, Dolado, and Lorens (2008) shed light on Spain. Cool (2010) looked at OECD countries and her data included Canada, but the focus was limited to the 20th and 80th percentiles and she did not report a detailed analysis.

3 NGS respondents are surveyed two and five years after graduation. A more detailed description of the data set follows in section 3.

such as wages, hours of work, industry, occupation, and type of job (permanent/temporary); the data also contain a more precise measurement of actual work experience, allowing us to control for more than only potential experience, as is common. The field of study information is especially interesting, because it allows an additional dimension of control: we know that men earn on average more than women in every field of study (Finnie and Wannell 2004; Bayard and Greenlee 2009) and that women are less influenced than men by potential future wages when choosing a college major (Boudarbat and Montmarquette 2009).

These data advantages notwithstanding, studying recent graduates is worthwhile for two additional reasons. First, because the targeted population covered is recent graduates of a given cohort, our samples are relatively homogeneous and the potential sample selection biases are accordingly alleviated. Close to 90% or more of individuals in our samples are employed, and most of them work full time, allowing us to shy away from work participation issues, which are especially important in usual gender comparisons due to women's fertility decisions. Second, recent graduates are also younger and closer in age than the general population: more than 70% of our sample were under age 27 at the time of graduation. We generally accept that the gender wage gap should be narrower within a defined and homogeneous group, so any difference we will find is all the more worrying and likely to be a lower bound for the Canadian population at large. Additionally, it is important to understand what could explain a wage gap early in an individual's work life, as differences tend to deepen over a lifetime. As Drolet (2011, 11) summarizes: 'The gender wage gap early in an individual's career is an increasingly good predictor of the wage gap throughout a generation's working life.' Thus, from a public policy perspective, knowing where we stand in terms of wage equality between men and women for our country's young and educated population should be of utmost interest.

As a preview of the results, we find mean wage gaps of over 6% two years after graduation and of more than 8% after five years. But paying attention to the entire distribution reveals different patterns. We do find that men earn more than women at every point of the distribution (except perhaps when only low-earning university graduates are considered). And while gaps remain around the 6% mark two years post graduation between the 10th and the 90th percentiles, they are substantially smaller (around 4%) for the middle percentiles five years after graduation. The tails of the wage distribution is where the most movement has been recorded: the wage gap has narrowed and almost disappeared in the bottom half of the distribution, but the story at the top is the opposite. At the 90th percentile, the wage gap five years after graduation went up from 8.7% in 1991 to 15.2% in 2005. Moreover, we also see a deterioration of women's lot within a given cohort in the three-year span between our data points: the gaps at the top half of the distribution have widened between 2002 and 2005 for the 2000 cohort. When splitting our sample by education level, we find gaps that are more or less constant across the distribution for college and CEGEP graduates, hovering around 10%. For university graduates, we observe a sharp gradient along the distribution, with small gaps at the bottom and rather large ones at the top. Trends for higher-educated graduates thus seem to be driving the results found for graduates overall.

Our results seem to indicate that young women have managed to free themselves from the ‘sticky floor,’ but that there is still a way to go to be able to crack the ‘glass ceiling.’ These findings are consistent with those of Wood, Corcoran, and Courant (1993), who studied gender differences in the career dynamics of lawyers from the University of Michigan Law School, and those of Bertrand, Goldin and Katz (2010), who did the same for MBAs from a top US program. Both found growing gaps in earnings and hours worked between men and women as their careers evolved.

A description of the methodology used in the paper is found in the next section. Section 3 presents the NGS data used in the analysis. Section 4 contains the findings of our research, while section 5 concludes.

2. Methodology

2.1. Decompositions at the mean

First, we performed standard Oaxaca-Blinder decompositions (Oaxaca 1973; Blinder 1973) to decompose the mean wage gap into a part due to differences in average characteristics between men and women (the explained part), and a part due to differences in returns to the various characteristics (the unexplained part). As is commonly done, we estimated the following two separate equations using ordinary least squares (OLS):

$$Y_{mi} = \boldsymbol{X}'_{mi}\boldsymbol{\beta}_{\boldsymbol{m}} + u_{mi} \text{ for men} \quad \text{and} \quad Y_{wi} = \boldsymbol{X}'_{wi}\boldsymbol{\beta}_{\boldsymbol{w}} + u_{wi} \text{ for women,} \tag{1}$$

where Y is the logarithm of the hourly wage, $\boldsymbol{X}$ is a vector of observable characteristics, u is a random error, and $i = 1, \ldots, n$ represents the individual observations. After the regression coefficients are estimated, it is straightforward to rewrite the log wage gap as

$$\bar{Y}_m - \bar{Y}_w = (\bar{X}_m - \bar{X}_w)\hat{\beta}_m + \bar{X}_w(\hat{\beta}_m - \hat{\beta}_w). \tag{2}$$

The first component of the right-hand side of equation (2), $(\bar{X}_m - \bar{X}_w)\hat{\beta}_m$, is the explained part of the gap, while the second component, $\bar{X}_w(\hat{\beta}_m - \hat{\beta}_w)$, is the unexplained part of it. This classic decomposition suffers from two commonly known problems (Neumark 1988; Jann 2008; Fortin 2008). The first concerns the choice of counterfactual (base) wage. Equation (2) uses as counterfactual the wage an average woman would have received at men’s returns: $\bar{X}_w\hat{\beta}_m$. It would be possible to use instead the wage an average man would have received at women’s returns, and the results would differ to some extent. To avoid this possible discrepancy, we applied Neumark’s method (Neumark 1988), which uses coefficients from a pooled regression (men and women together) where a dummy variable for women is included. This approach is consistent with that of Fortin (2008).[4] The second issue is the identification of the contribution of categorical variables (e.g., education level, field of study, occupation). In practice, these variables are transformed into a set of dummy variables, where one category is omitted from the regression to avoid perfect multicollinearity. The problem is that the choice of the reference group affects the decomposition results. We used Jann's (2008) solution, which constrains the coefficients on the dummy variables to sum to zero to express them in terms of deviations from the mean of each category. When we do this, the base group can be added back to the regression, thus bypassing the issue of which reference group to pick (Jann 2008; Yun 2005).

A question that arises is which explanatory variables should be included in the $\boldsymbol{X}$ vector. The gender gap literature acknowledges that 'there is no universally accepted set of conditioning variables that should be included' in decompositions (Drolet 2002, 32). We employed the following widely used factors: personal characteristics (marital status, presence of children in the household, age of youngest child, and province of residence),[5] education level and work characteristics (permanent/temporary/seasonal job and a quadratic in experience). Additional and somewhat controversial factors include industry and occupation dummies, and to some extent field of study dummies. As Drolet (2002, 32) added: 'if employers differentiate between men and women through their tendency to hire into certain occupations, then occupational assignment is an outcome of employer practices rather than an outcome of individual choice (Altonji and Blank 1999). Analyses that omit occupation and industry may overlook the importance of background and choice-based characteristics on wage outcomes, while analyses that fully control for these variables may undervalue the significance of labour market constraints on wage outcomes.' For our paper, we estimated four different specifications, ranging from a simple specification, where $\boldsymbol{X}$ includes only education level, experience, and province of residence, to a full specification, where all explanatory variables noted above are included in $\boldsymbol{X}$. In that way, we are able to show the effect on our estimates of adding industry, occupation, and field of study to the model. After presenting findings using the most complete model (including occupation, industry, and field of study dummies) for all cohorts, in section 4.3 we report findings from the four models in detail for the most recent year of available data, 2007. Note that the analysis is performed separately for

4 As a robustness check, we also performed our decompositions using the male wage structure instead of a pooled-gender one. The unexplained portion of the gap tends to decrease slightly when doing so. A difference is to be expected because women comprise more than 50% of our sample and so play a large role in determining the coefficient of the wage structure in the pooled regression. The difference is not dramatic, however, since our main findings across cohorts and along the wage distribution are not changed.

5 Information on immigrant status is not available for the 1995 cohort in 1997 and for earlier cohorts (1986 and 1990), so for comparability purposes it is excluded from the analysis. Computations using the other cohorts showed that the impact of immigrant status on the gender wage gap was minimal and its exclusion does not change our findings.

each data set, that the sampling weights provided by Statistics Canada are used to weight observations in all computations, and that robust standard errors are used in all regressions.

2.2. Decompositions along the distribution

Since the gender gap is not constant over the wage distribution, we think it is important to also consider decompositions along the whole distribution. These types of decomposition, into explained and unexplained parts, are similar in spirit to the Oaxaca-Blinder decompositions at the mean explained in the previous subsection, but adapted to a quantile-based approach. We used the unconditional quantile regression method set forth by Firpo, Fortin, and Lemieux (2009) and Fortin, Lemieux, and Firpo (2010) (hereafter FFL). The general idea of the method is the following (the reader is referred to the FFL references for in-depth presentation).

Suppose we are interested in the men-women wage gap at the *j*th percentile, $\Delta(j)$ $(1 \leq j \leq 100)$:

$$\Delta(j) = q_m(j) - q_w(j), \tag{3}$$

where $q_m(j)$ and $q_w(j)$ are the *j*th percentiles for the male and female distributions, respectively. FFL (2009) showed that $\Delta(j)$ can be decomposed into a part due to composition ($\Delta^X(j)$) and a part due to the wage structure ($\Delta^\beta(j)$), as was done for the mean in equation (2):

$$\Delta(j) = [q_m(j) - q_c(j)] + [q_c(j) - q_w(j)] = \Delta^X(j) + \Delta^\beta(j), \tag{4}$$

where $q_c(j)$ is the counterfactual wage at the *j*th percentile. This decomposition can be performed by first computing a recentred influence function (*RIF*) defined as such:

$$RIF_{ji} = q(j) + [1(Y_i \geq q(j)) - (1 - j)]/f(q(j)), \tag{5}$$

where $q(j)$ is the *j*th percentile of the pooled male-female sample, 1(.) is a binary variable taking value 1 if the log wage (Y) is greater than or equal to $q(j)$ for individual *i* and 0 otherwise, and $f(q(j))$ is the density of the wage at the *j*th percentile (which can be estimated by kernel density, for example). Next, the standard Oaxaca-Blinder approach is applied, but using *RIF* (as computed from (5)) as the dependent variable in OLS regressions akin to those in (1).[6] Empirically, and as in Boudarbat and Lemieux (2010), we will present decomposition results for the 10th and 90th percentiles, but we will also show graphs of the adjusted wage

6 For more details, see also the description of the procedure in Heywood and Parent (2012).

gap (i.e., the wage gap that remains unexplained after removing the explained part) throughout the distribution.[7]

One valid concern when applying the unconditional quantile regression method is that of common support: are men and women comparable at the very bottom and very top of their common distribution, or rather, for each possible combination of characteristics that men (women) have are there enough women (men) with the same characteristics for the approach to yield meaningful results? Overlapping support is one of the assumptions upon which the FFL methodology is based, but as Fortin, Lemieux, and Firpo (2010, 16) themselves noted: 'in the decomposition of gender wage differentials, it is not uncommon to have explanatory variables for which this condition does not hold.' When performing a decomposition at the mean, other authors have suggested alternative approaches based on non-parametric matching (Black et al. 2008; Ñopo 2008). However, these methods are not readily applicable to the decompositions performed here, at various points of the wage distribution. We therefore acknowledge the potential common support problem, but note that (1) our sample sizes are large (more than 15,000 individuals per cohort in the smallest sample) and that (2) common support should be less of an issue in our simplest specification, which includes as explanatory variables only education, experience, and province of residence. Thus, by considering the most parsimonious model and given the sample sizes, we hope that the inherent issue of common support should be alleviated as best as it can be.

3. Data

Our data come from the National Graduates Survey (NGS). Statistics Canada surveyed representative samples of all Canadian post-secondary graduates during given calendar years; for our analysis we used the five most recent cohorts, those of 1986, 1990, 1995, 2000, and 2005. The graduates were surveyed two years after graduation and in some cases also five years after graduation, in a follow-up survey. In all, our data cover five cohorts in nine years: the 1986 cohort, surveyed in 1988 and 1991 (NGS and follow-up), the 1990 cohort, surveyed in 1992 and 1995 (NGS and follow-up), the 1995 cohort, surveyed in 1997 and 2000 (NGS and follow-up), the 2000 cohort, surveyed in 2002 and 2005 (NGS and follow-up), and the 2005 cohort, surveyed in 2007 (NGS).[8,9] The targeted population of each survey is all individuals who graduated from a Canadian public post-secondary

7 Summary statistics, detailed regression results, and coefficients for the mean and the 10th and 90th percentiles are available in the appendix tables in the technical appendix.

8 Unfortunately, Statistics Canada did not conduct a follow-up survey for the 2005 cohort in 2010.

9 Note that we accessed the confidential data files available at one of Statistics Canada's Research Data Centres. Public-use data sets are available to the Canadian research community, but the level of the detail is too aggregated for our purpose (e.g., salary information comes in a categorical variable instead of a continuous one).

institution during the given year, which includes people who earned a trade or vocational degree, a college or CEGEP diploma or certificate,[10] a bachelor's degree, a master's degree, or a doctorate. We note that Statistics Canada decided not to cover trade school graduates for the 2005 follow-up survey of the class of 2000, owing to 'conceptual and sample requirements issues' (Statistics Canada 2007). This makes it difficult to consistently study trade school graduates across time, so we decided to drop them from our analyses.[11] Therefore, our focus will be on college/CEGEP and university graduates.

All the variables used in the analysis are defined in appendix table 1 and their summary statistics are shown in appendix tables 2, 3, and 4, all available in the online technical appendix. Our work experience variable is computed by combining the information contained in various variables to try to obtain the most precise estimate possible, in order to take into account career interruptions and fertility decisions, which have been shown to have a high explanatory power for the gender gap (O'Neill and O'Neill 2006; Phipps, Burton, and Lethbridge 2001). Specifically, we added the pre-diploma full-time experience (as reported by the respondent) to the post-diploma work duration.[12]

The NGS data provide detailed schooling information, giving not only the degree obtained in the cohort year but also any additional diploma held at the time of the reference year. For our analyses, we used as the education level the highest one attained at the time of each survey. So, if for example an individual who graduated in 2000 with a bachelor's degree subsequently completed a master's program in 2004, she would show up in our sample for 2002 (NGS) as having a bachelor's degree, but in our sample for 2005 (follow-up survey) we would classify her as having a master's degree.

For our analysis sample, we kept only graduates who lived in one of the 10 provinces,[13] were 50 years old or less at the time of graduation, and were employed full time (30 hours per week or more) during the survey's reference week; we dropped the unemployed and also the self-employed and those in part-time employment. The age restriction is intended to focus our attention on people who study to invest in their human capital as way to gain later returns on the job market (not those older individuals who study as a leisure pursuit). It does not

10 CEGEP stands for 'Collège d'enseignement général et professionnel,' or College of General and Vocational Education, and is a type of college in the province of Quebec providing either a two-year general degree between high school and university, or a three-year technical degree.

11 Trade school graduates account for a relatively small and declining fraction of the NGS sample, going from 21% to 16% of all graduates between 1986 and 2005 for men, and from 15% to 8% for women over the same period.

12 Note that the latter work duration includes part-time work, but that should be only a minor issue, as most individuals work full time. In addition, our results do not change dramatically when we use the common potential experience measure computed as *Age* − *Years of Education* − 6.

13 Starting with the 2000 follow-up survey of the class of 1995, Statistics Canada also surveys graduates from Canadian institutions who had moved to the United States. We drop these observations because of the labour market differences between the two countries and the small number of individuals concerned.

remove too many observations from our analysis, since over 70% of our sample were under 27 years of age at graduation. The median age at graduation was 23 or 24 for both males and females over the whole period studied. Appendix table 5 in the technical appendix provides detailed age statistics for our analysis samples.

The full-time restriction is meant to allow for a comparison of similar types of workers. Part-time employment is often found as explaining part of the gender wage gap: Drolet (2002) reported that 2% to 8% of the gap calculated from the 1997 Survey of Labour and Income Dynamics could be attributed to part-time employment. Thus, focusing on full-time workers should make men and women in our sample more comparable. One noticeable feature of the data we use is the high rate of full-time employment among women. This is due to the fact that recent graduates are young and tend to devote more time to work, a tendency that may get reversed as women start to have children and shift their focus away from paid employment. Rates of full-time employment by gender and cohort are reported in appendix table 6 in the technical appendix. We observe that rates for men are high, never dipping below 90% and reaching highs of around 96%. For women, rates for all years are in the 86–90% range, with one exception: only 79% of the women from the 1995 cohort in 1997 worked full time. The reasons for this dip in full-time employment are not clear, but nevertheless the overall rates are much higher than those for the Canadian population at large; Ferrao and Williams (2011) reported that 26.1% of women and 11% of men worked part time in 2007.

We decided to analyse the *wage* gap, but not the *earnings* gap or *income* gap. While the latter measures would also be interesting to look at, as a follow-up to Finnie and Wannell (2004), for example, we wanted to abstract away from work intensity considerations: if men worked more than women and that explained most of the earnings gap, the relevant object to study would not be the earnings but rather the hours-of-work decision. Even when multiple waves of the same survey are used, assuring consistency in a variable's definition through time can be challenging. Statistics Canada provides a measure of annual earnings that is comparable for all cohorts. What is interesting is that the measure does not capture annual earnings per se, but rather the annual earnings for the main job *if the respondent were to work his or her usual hours at that job for 12 months*. Thus, we took this measure and divided by the usual weekly hours multiplied by 52 to compute hourly wage.[14] For earlier cohorts (1986 and 1990), annual earnings is the only variable available pertaining to wages. For later cohorts, the NGS also provides a direct measure of wage, which respondents can give at various intervals: hourly, bi-weekly, monthly, annual, and so on. As a robustness check, we compared the hourly wage we computed to that given directly in the NGS. The figures were extremely close to each other, giving us confidence that our

14 Weekly hours are not available for the 1982 cohort, which unfortunately prevents us from adding this cohort to our analysis.

TABLE 1
Sample sizes

	Year of survey	Men	Women	Total
1986 cohort	1988	10,319	9,853	20,172
	1991	9,519	8,922	18,441
1990 cohort	1992	9,189	9,186	18,375
	1995	7,737	7,760	15,497
1995 cohort	1997	10,322	9,861	20,183
	2000	7,505	7,523	15,028
2000 cohort	2002	9,929	11,896	21,825
	2005	6,749	8,578	15,327
2005 cohort	2007	9,380	11,941	21,321

NOTES: The sample consists of full-time salaried graduates with college diplomas or university degrees, aged 50 or less (at graduation). Top and bottom 0.5% of wages are trimmed.

measure based on annual earnings is adequate. Once we calculated a measure of hourly wage for all individuals, we trimmed the top and bottom 0.5% of our sample to delete implausibly low and high values of wages.

Table 1 shows the sample sizes by cohort and year of survey, by gender, and total. The number of observations by gender goes from just under 7,000 to close to 12,000, for total combined gender number of up to around 22,000 by year. The large sample sizes should help alleviate the common support issue discussed in section 2.2, allowing us to compute meaningful decompositions at various points along the wage distribution. Table 1 also shows that some attrition happened between the NGS and its follow-up for all cohorts. While this could potentially bias estimates for the period five years after graduation if attrition is non-random, we always used the weights developed by Statistics Canada to ensure that the follow-up surveys are representative of the target population (Statistics Canada 2007). We therefore believe that the biases related to attrition should be minimized. We observe that the attrition rate increased from one cohort to the next in the time period studied: while we lose only 8.6% of the 1986 cohort in our sample due to attrition, the rate steadily increases to 15.7% for the 1990 cohort, 25.5% for the 1995 one and 29.8% for the 2000 graduates. The reasons for this increase are not clear and further study of the issue could be informative.

4. Findings

4.1. Results at the mean

We start by presenting in figure 1 the gender wage gap at the mean, that is, the difference between the average of the log wages for men and that for women.[15] In this figure, the gap is computed as the value for men minus that for women,

15 Complete results are available in appendix table 7 in the technical appendix.

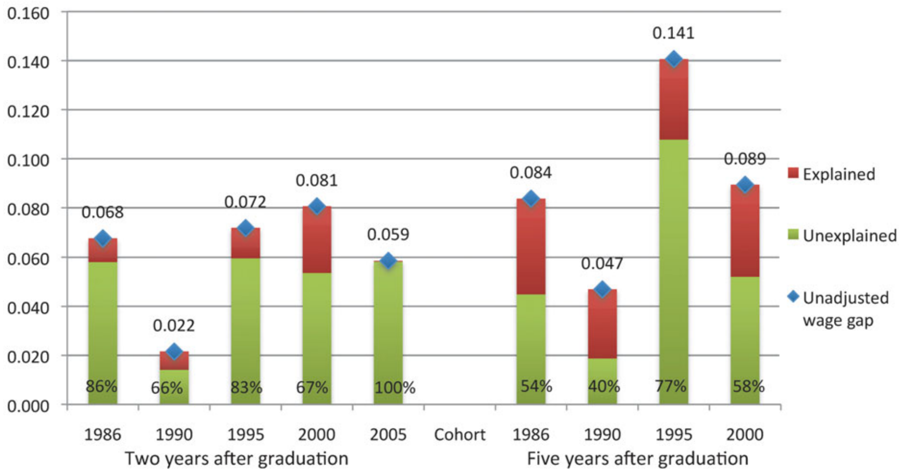

FIGURE 1 Gender log wage gaps at the mean and fraction explained and unexplained
NOTES: The percentage value shown is the fraction of the wage gap that is unexplained, expressed in percentage. The number above the bar is the value of the unadjusted log wage gap. The five bars to the left represent data from two years after graduation and the four on the right data from five years after graduation.

so that a positive number means that men earn more than women. The gap is presented in log points, which for reasonably small values can be approximated as the percentage difference between men and women.[16] We show the wage gap on the left y-axis, as well as fractions explained and unexplained, as computed after the Oaxaca-Blinder decompositions presented in section 2. Note that unless otherwise specified, we used a full set of covariates, including industry, occupation, and field of study dummies.[17] The number shown above the bars is the total unadjusted wage gap, while the percentage shown close to the x-axis is the fraction of the wage gap that is unexplained, expressed in percentage. The height of the unexplained bar gives the adjusted wage gap, after taking into account all covariates. We first note that two years after graduation, Canadian women earned on average 6–8% less than men in all of our cohorts, with the notable exception of the 1990 cohort, which showed a much smaller gap of 2.2%. We return to the 1990 cohort in the discussion in section 4.5. That exception aside, the gap appears to have been relatively constant between 1988 and 2007, with no value of the wage gap statistically different from cohort to cohort. Five years after graduation, the gaps reached 8–14% (again with the exception of the 1990 cohort, with a gap of 4.7% in 1995), meaning that they widened in the three-year span between the NGS and follow-up surveys. However, the rate of increase is

16 Throughout the text we will make no distinction between a difference in log points and percentage points, even though the approximation becomes less precise as the difference grows.
17 An investigation of decomposition results using more parsimonious models is presented in section 4.3.

shrinking between cohorts: in the case of the 1990 and 1995 cohorts the gap more than (or almost) doubled, while it increased by only 0.008 log points for the class of 2000. Comparing our results with those of Finnie and Wannell (2004), we find much smaller gaps than they do for the 1986 and 1990 cohorts, which is likely due to the fact that we are interested in hourly wage differences, whereas they used annual earnings. Greater disparities at the annual level would imply that hours of work plays an important role in explaining the earnings gap, an intuition confirmed by Finnie and Wannell's inclusion of hours of work in their earnings decomposition.

There is considerable variation in the fraction of the gap that is unexplained, the percentages ranging from 40% to 100%, and no clear trend over time and across cohorts.[18] The adjusted wage gaps (the height of the unexplained bars) range from 1.4% to 10.8%. To sum up, at the mean of the wage distributions, we can explain up to 60% of the gender wage gap, but we still observe adjusted gaps of 0.014 to 0.108 log points. For the latest cohort, women's characteristics have the least explanatory power, accounting for virtually nothing of the mean wage gap.

4.2. Results along the distribution

Next we turn to decomposition results along the distribution of wages. To give a more complete look at what happens at all points of the distribution, rather than at the mean, we present in figures 2 and 3 the gender log wage gaps by percentile for each cohort. Figure 2 shows the wage gaps adjusted for covariates two years after graduation. The 1990 cohort is again an outlier, to which we return in section 4.5. Looking at the middle of the distribution for the other cohorts, we can see that all the wage gaps are more or less flat between the 20th to 80th percentiles, and in the 4–7% range. As for the gaps at the mean, no obvious upward or downward time trend emerges, especially when considering that the standard errors of the various cohorts' adjusted gaps are such that their confidence intervals overlap. In the bottom 20% of the distribution, the gaps are generally below those of the middle, except for the 1986 cohort in 1988, for which we observe a sharp decline from over 0.1 log points at the 2nd percentile to below 0.06 by the 20th percentile. This is in marked contrast with the most recent year available, the 2005 cohort in 2007, with the same figures going from 0.004 to 0.058. Past the 80th percentile, the gaps start to increase, especially sharply after the 85th or 90th percentiles are reached. While that was true for each cohort (except maybe the 2000 cohort, for whom the gap peaks at the 90th percentile), we also note a deterioration over the years for women at the very top of the distribution. At the 95th percentile, the 1995 cohort experienced a gender gap of

18 To the best of our knowledge, there were no changes in the definition or coding of the variables in the data that could explain the changes in the explanatory power of the variables across cohorts.

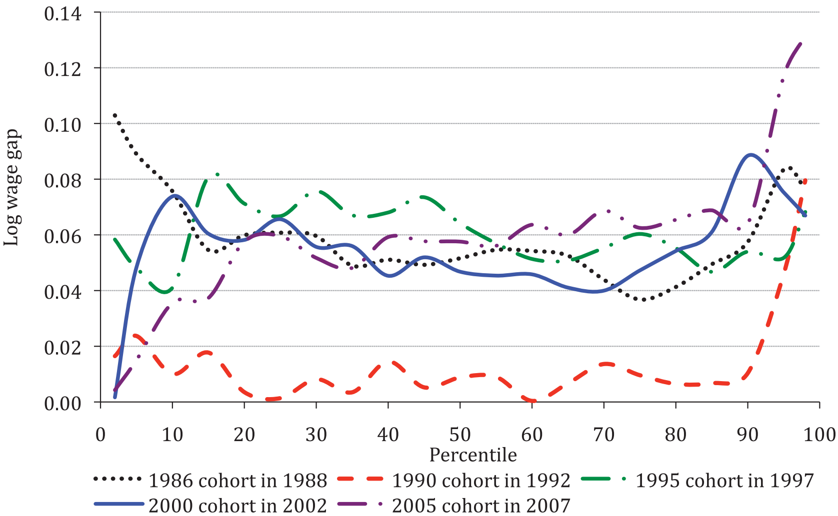

FIGURE 2 Gender log wage gap two years after graduation by percentile (adjusted for covariates)
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

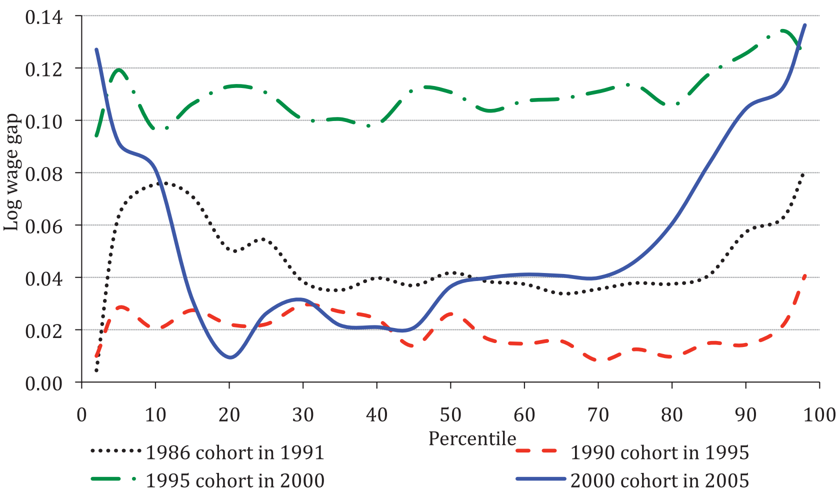

FIGURE 3 Gender log wage gap five years after graduation by percentile (adjusted for covariates)
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

0.0518 log points in 1997. The gap increased to 0.0753 log points in 2002 for the 2000 cohort, and to 0.1170 in 2007 for the class of 2005.

Figure 3 is similar to figure 2 but shows the adjusted gaps using the follow-up surveys – so five years after graduation. We immediately notice that the 1990 cohort has consistently small gaps throughout the wage distribution, reaffirming

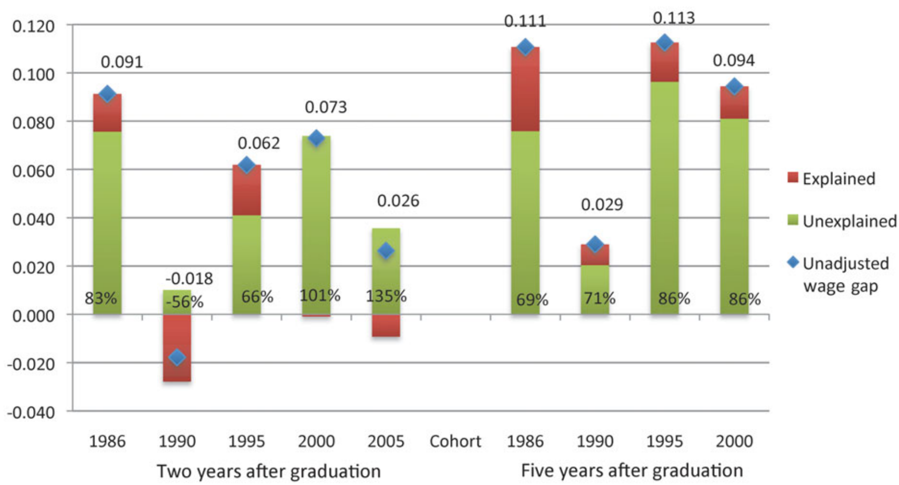

FIGURE 4 Gender log wage gaps at the 10th percentile and fraction explained and unexplained
NOTES: The percentage value shown is the fraction of the wage gap that is unexplained, expressed in percentage. The number above the bar is the value of the unadjusted log wage gap. The five bars to the left represent data from two years after graduation and the four on the right data from five years after graduation.

its outlier status among the NGS cohorts studied here. We also note that the 1995 cohort in 2000 exhibits high and stable gaps all along the distribution, with gaps around 10–11% until the 80th percentile and inching up to 12–13% from the 85th percentile on. This is a very different pattern from that of the other two cohorts: 1986 in 1991 and 2000 in 2005. For the older cohort, the gaps increased at first, to peak at the 10th percentile, then decreased to reach a stable level up to the 85th percentile, past which they increase again. In 2005, the 2000 cohort had its peak gaps at both the very bottom and the very top of the distribution. The smallest gap is at the 20th percentile (0.9%), which is followed by gradual increases until the 75th percentile (4.6%) is reached. The gaps then shoot up to a peak of 13.6% at the very top of the distribution, the largest adjusted gap we computed. We also note that the line for the 2000 cohort almost always lies under that of the 1995 cohort, showing that the lot of female graduates has improved relative to men between 2000 and 2005, or five years after their graduation. The improvement gets smaller as we move towards the top of the distribution and is wiped out at the very top percentiles.

We next turn to the detailed decomposition results forming the basis for figures 2 and 3 for two points of the distribution of wages. Figures 4 and 5 show the results of the FFL decompositions at the 10th and 90th percentiles, respectively, in the same manner as figure 1 showed for the mean.[19]

19 Complete results are available in appendix tables 8a and 8b in the technical appendix.

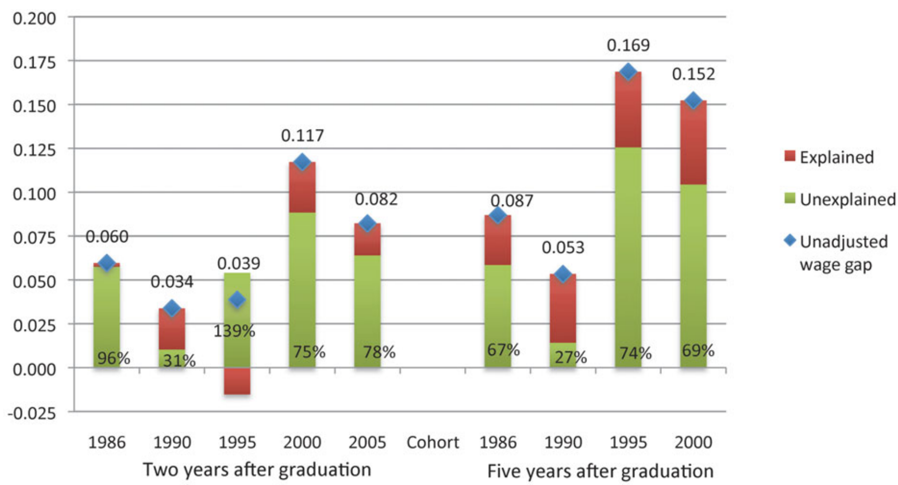

FIGURE 5 Gender log wage gaps at the 90th percentile and fraction explained and unexplained
NOTES: The percentage value shown is the fraction of the wage gap that is unexplained, expressed in percentage. The number above the bar is the value of the unadjusted log wage gap. The five bars to the left represent data from two years after graduation and the four on the right data from five years after graduation.

Looking at figure 4, we again notice how the 1990 cohort stands out from the rest, with a negative unadjusted gap (meaning that women actually earned more than men) in 1992 and a small gap of 2.9% in 1995, although both are not statistically different from zero. The 1986 cohort exhibits high values of both unadjusted (0.091 and 0.111 in 1988 and 1991, respectively) and adjusted gaps (0.076 for both years). These figures are higher than the gaps at the mean for the same year, so women at the bottom of the income distribution fared worse relative to men than those earning a wage around the mean. The unadjusted wage gaps two years after graduation decreased more or less gradually between 1988 and 2007 (1992 aside), to reach a low (and not statistically different from 0) of 2.6% in 2007. A feature to note is that within any given cohort, both the unadjusted and adjusted wage gaps increase during the period of two to five years after graduation. Thus, even for low earners, the gender gap can substantially increase during the early years of their work life. This situation is worrisome, as we usually expect the gap between males and females to increase over the years (albeit not as fast as in the previous decades; see Drolet 2011), and our findings suggest that the rate of increase is high even soon after graduation.

The unexplained fraction of the gap varies from 66% to 135% (with the exception of 1992, in which the fraction is −56%, the adjusted gap being 0.1 log points but the unadjusted one −0.018). As for the mean, a trend both within and across cohorts is not obvious. In 2007, the last year for available data, the variables we used explain a negative 35% of the gap, meaning that women's characteristics actually play in their favour, at least at the 10th percentile of the wage distribution.

Figure 5 shows our results at the 90th percentile. Some of our findings from the 10th percentile persist: the 1990 cohort again appears to be an outlier, with relatively low values of unadjusted and adjusted wage gaps, and the gaps increased in the two-to-five-years interval after graduation. Another trend emerges: the gaps increased between the early cohorts (1986 and 1990) and the later ones (1995, 2000, and 2005), although the 1997 results for 1995 are more in line with the those from the early cohorts. Indeed, we found adjusted wage gaps between 0.01 and 0.059 log points for the years 1988 to 1997 but between 0.064 and 0.126 log points between 2000 and 2007. The corresponding values for the unadjusted gaps are 0.034 to 0.087 and 0.082 to 0.169, respectively. What we found is thus an increase in the male-female wage disparity at the top of the income distribution in the first decade of the 21st century.

The 1990 cohort notwithstanding, the unexplained portions of the gap range from 67% to139%, leaving only up to 33% to be explained by the variables in our model. The part of the gap due to potential discrimination or other unobservable factors is substantial, as was true for at the 10th percentile of the wage distribution as well as at the mean, albeit to a slightly lesser extent.

In considering the conditions of women in the workforce, a traditional glass ceiling – one that prevents women from reaching the top rungs of the career ladders – appears to be only part of the story, as we find evidence of equally invisible impediments from the very beginning of a woman's career. These findings echo those of Bertrand, Goldin, and Katz (2010) on MBAs, who found growing gaps as years from graduation increase, reaching 60 log points 10 to 16 years after graduation. They attributed their results to three main factors: differences in business school courses and grades, in careers interruptions, and in weekly hours worked. We address the fertility decisions and hours worked in section 4.5 as a robustness check on our results.

4.3. The explanatory power of various factors

All the findings presented so far come from a model with a full set of explanatory variables, including industry, occupation, and field of study dummies as well as the other covariates listed in section 2.1. To investigate the effect of the choice of variables on the decomposition results, we tried four different model specifications and – for the sake of conciseness –estimated them using only the latest year available: 2007 data pertaining to the 2005 cohort. We computed the decompositions at the mean and at the 10th and 90th percentiles. Table 2 presents our findings and consists of three panels, each with four columns: panel A contains results for the decomposition at the mean, panel B at the 10th percentile, and panel C at the 90th percentile. The figures in column (4) correspond to the full model, presented in the previous sections. Model 1 is the most parsimonious, containing only education level, experience, and province. Model 2 adds fields

TABLE 2
Factors accounting for decomposition of gender log wage gap at mean, 10th percentile, and 90th percentile, 2005 cohort in 2007

	Model 1 (1)	Model 2 (2)	Model 3 (3)	Model 4 (4)
Panel A: log wage gap at mean (unadjusted gap: –0.059)				
Fraction explained	– 9.0%	10.8%	0.5%	0.4%
Fraction unexplained	109.0%	89.2%	99.5%	99.6%
Fraction explained by				
Education level	– 15.4%	– 16.9%	– 12.9%	– 13.0%
Field of study	–	21.9%	31.5%	31.8%
Experience	5.2%	4.4%	3.2%	2.7%
Permanent job	–	–	5.6%	5.7%
Occupation	–	–	– 37.7%	– 37.1%
Industry	–	–	9.4%	10.2%
Presence of children	–	–	–	0.1%
Age of youngest child	–	–	–	– 0.2%
Marital status	–	–	–	– 1.3%
Province	1.2%	1.4%	1.5%	1.5%
Panel B: log wage gap at 10th percentile (unadjusted gap: –0.027)				
Fraction explained	– 24.4%	44.8%	– 34.0%	– 35.1%
Fraction unexplained	124.4%	55.2%	134.0%	135.1%
Fraction explained by				
Education level	– 41.8%	– 42.5%	– 28.8%	– 28.9%
Field of study	–	71.5%	106.3%	107.4%
Experience	15.1%	13.2%	9.8%	9.7%
Permanent job	–	–	18.7%	19.7%
Occupation	–	–	– 112.6%	– 112.3%
Industry	–	–	– 29.9%	– 32.7%
Presence of children	–	–	–	1.6%
Age of youngest child	–	–	–	– 2.6%
Marital status	–	–	–	0.4%
Province	2.3%	2.7%	2.5%	2.5%
Panel C: log wage gap at 90th percentile (unadjusted gap: –0.082)				
Fraction explained	4.1%	8.5%	21.1%	22.1%
Fraction unexplained	95.9%	91.5%	78.9%	77.9%
Fraction explained by				
Education level	– 0.1%	– 1.5%	– 0.2%	– 0.3%
Field of study	–	5.9%	– 8.0%	– 6.2%
Experience	3.2%	2.9%	2.1%	1.4%
Permanent job	–	–	3.1%	3.0%
Occupation	–	–	– 17.3%	– 17.7%
Industry	–	–	40.1%	41.2%
Presence of children	–	–	–	– 0.7%
Age of youngest child	–	–	–	1.9%
Marital status	–	–	–	– 1.7%
Province	1.0%	1.1%	1.2%	1.3%

NOTES: The sample consists of full-time salaried graduates with college diplomas or university degrees, aged 50 or less (at graduation). Top and bottom 0.5% of wages are trimmed. Weighted using survey's sampling weights. Full decomposition results available in appendix tables 9a, 9b, and 9c in the technical appendix.
SOURCE: Authors' calculations from National Graduates Survey data.

of study dummies and Model 3 job permanency, occupation, and industry dummies. All fractions are presented as percentages of the unadjusted wage gap. As a reminder, a negative percentage as the fraction explained by a given factor would be interpreted as the factor being beneficial to women: evaluated at an average return (because we are using Neumark's approach), the wage difference between males and females should be smaller because of that factor.

Starting with Panel A, we can account for the wage gap at the mean. The fraction explained varies by model, from −9% (Model 1) to 11% (Model 2). Models 3 and 4 explain very little of the gap (0.5% and 0.4%, respectively). These figures hide dramatic differences among the explanatory power of specific variables, but specific variables tend to have a similar impact across models. For example, the education level consistently explains between −17% and −13% of the gap, consistent with the higher educational attainment of women. In contrast, fields of study explain 22–32% of the gap, thus contributing to the wage difference between genders, and consistent with previous findings in the literature that women choose less lucrative fields than men. In our 2007 sample, twice as many women studied in education compared with men (11.6% of women compared with 5.9% of men), and the percentage of men who studied in architecture, engineering, and related technologies was more than six times as high as that of women (25.3% of men compared with 4.1% of women).[20]

Both occupations and industry have large but opposing effects: occupations account for −37% or –38% of the gap and industry accounts for 9% or 10% of it. Other factors have smaller effects: experience explains between 2.7% and 5.2% of the wage difference and holding a permanent job 5.6% or 5.7%. Our socio-demographic variables (presence of children, age of youngest child, marital status, and province of residence) explain at most 1.5% of the gap.

At the 10th percentile (panel B), most of the findings from the mean still apply, albeit with amplified effects. Education level now accounts for up to −43% of the gap and field of study up to 107%. Occupational choice appears to be driving −113% of the gap. The effect of industry has switched signs: industry is now responsible for −33% to −30% of the gap. Experience and having a permanent job have a larger effect than at the mean (10–15% and 19–20%, respectively), while the impact of other factors remains small.

At the 90th percentile (panel C), most effects are dampened but remain similar to those at the mean, with three exceptions. First, there is a sign reversal in the effect of fields of study: in models 3 and 4 (those including occupation and industry), field of study is accountable for a negative fraction of the gap: –6 or –8%. Second, the fraction explained by occupations is smaller but still a sizeable −17% or −18%. Finally, industry takes the lion's share of the explained part of the gap, with fractions explained of 40% or 41%.

20 Appendix table 3 in the technical appendix contains statistics on the distribution of fields of study by gender and cohort.

Table 2 presented the combined effect of groups of variables, for example the combined effect of all occupation dummies. To investigate which of the individual categories have the highest explanatory power, we now present in table 3 the disaggregated fractions explained by the education level, field of study, occupation, and industry dummies, again for the 2005 cohort in 2007. Given that the effect of these variables does not change much from model to model, we show results only from the estimation of model 4, the most complete model and the one corresponding to the results from sections 4.1 and 4.2. Looking at the highest level of studies, we see that the total negative effect is driven largely by the negative effect of having obtained at most a college degree. In 2007, 39% of men in our sample had a college degree for highest educational level compared with 35% of women. This appears to play in favour of the women, reducing the wage gap at the mean and the 10th and 90th percentiles. At the mean, the impact of a bachelor's degree is much smaller and still negative, while it remains large at the 10th percentile and becomes positive at the 90th percentile. More men than women attain a postgraduate level of education, and that is reflected in the positive effect on the gap (it explains 2–4% of the gender wage gap).

The overall effect of the field of study is positive for the decompositions at the mean and the 10th percentile. By far the largest contributor to this effect is 'Architecture, Engineering and Related Technologies,' a fact consistent with the difference in field of study distributions between men and women; as noted above, six times more men than women chose that major. Somewhat offsetting this effect, the fields of 'Business, Management and Public Administration' and 'Health, Parks, Recreation and Fitness' show negative effects, albeit not as large as the engineering field. These findings are similar to those of Drolet (2002), who reported that 15% of the average wage gap could be explained by the fact that men were more likely to graduate from engineering and applied sciences programs.

In terms of occupations, the main driver of the overall negative impact on the gap is by far the health occupations. Consistent with the negative effect of studying in a health-related field, it appears that women's predominance in the health professions – three to four times more women than men work in health occupations[21] – should contribute to a narrowing of the wage gap. This could be explained partially by the stereotypical difference between men and women within the health category: although falling under the same banner, men become doctors and women become nurses. The second biggest contributor, this time on the positive side, is the natural and applied sciences occupations, which is clearly a male-dominated (and lucrative) field: while 21–27% of men work in this field, only 6–9% of women do.

The contributions of the industry categories to the explained wage gap are not as clear-cut as those of the occupation and field-of-study groups. At the mean,

21 Appendix table 4 in the technical appendix contains statistics on the distribution of occupations by gender and cohort.

TABLE 3
Detailed factors accounting for decomposition of gender log wage gap at mean, 10th percentile, and 90th percentile, 2005 cohort in 2007

	Mean	10th percentile	90th percentile
Log wage gap (unadjusted)	– 0.059	– 0.026	– 0.082
Fraction explained (%)	0.4	– 35.1	22.1
Fraction unexplained (%)	99.6	135.1	77.9
Fraction explained by (%)			
Highest level of studies	– 13.0	– 28.9	– 0.3
College or CEGEP diploma or certificate	– 14.8	– 18.9	– 11.6
Bachelor's degree	– 1.6	– 12.3	7.0
Postgraduate	3.4	2.3	4.4
Field of study	31.8	107.4	– 6.2
Education	– 1.6	– 0.2	– 1.0
Visual and Performing Arts, and Communications Technologies	1.2	5.3	0.1
Humanities	1.4	1.8	0.6
Social and Behavioral Sciences and Law	4.1	– 0.2	1.6
Business, Management and Public Administration	– 5.9	– 17.1	– 3.7
Physical and Life Sciences and Technologies	0.1	0.3	0.1
Mathematics, Computer and Information Sciences	6.0	12.6	0.1
Architecture, Engineering, and Related Technologies	38.5	120.5	10.7
Agriculture, Natural Resources and Conservation	– 0.2	2.0	– 1.4
Health, Parks, Recreation and Fitness	– 12.1	– 13.7	– 13.4
Personal, Protective and Transportation Services	0.2	– 3.9	0.2
Occupation	– 37.1	– 112.3	– 17.7
Management Occupations	2.3	4.9	2.3
Business, Finance and Administrative Occupations	3.5	– 12.9	5.1
Natural and Applied Sciences and Related Occupations	18.7	72.3	9.3
Health Occupations	– 34.7	– 84.0	– 17.7
Occ. in Social Science, Education, Government Service and Religion	– 10.2	– 44.1	– 6.5
Occ. in Art, Culture, Recreation and Sport	– 0.6	– 5.5	0.1
Sales and Service Occupations	– 4.0	– 18.2	– 0.6
Trades, Transport and Equipment Operators and Related Occ.	– 4.1	11.3	– 4.3
Occupations Unique to Primary Industry	– 4.7	– 19.1	– 1.8
Occupations Unique to Processing, Manufacturing and Utilities	– 3.4	– 17.1	– 3.5
Industry	10.2	– 32.7	41.2
Agriculture, Forestry, Fishing, Mining, Oil and Gas	8.4	13.3	9.2
Utilities, Accommodation and Food Services, Other Services	– 3.6	6.4	– 2.5
Construction, Transportation and Warehousing	4.3	21.2	0.7
Manufacturing – Durables and Non-durables	– 4.1	– 22.5	0.5
Wholesale Trade, Retail Trade	1.3	5.9	– 0.1
Finance, Insurance, Real Estate and Leasing	– 0.1	18.2	– 1.9
Professional, Scientific and Technical Services	– 1.8	– 4.5	– 0.5
Business, Building and Other Support Services	– 9.3	– 49.3	6.1
Educational Services	7.0	– 44.9	28.8
Health Care and Social Assistance	– 0.8	– 2.9	– 0.5
Information, Culture and Recreation	1.6	7.0	0.2
Public Administration	7.3	19.4	1.2
Other factors	8.6	31.4	5.1

NOTES: The sample consists of full-time salaried graduates with college diplomas or university degrees, aged 50 or less (at graduation). Top and bottom 0.5% of wages are trimmed. Weighted using survey's sampling weights.
SOURCE: Authors' calculations from National Graduates Survey data

the overall effect of industry is positive, the largest positive contributors being 'Agriculture, Forestry, Fishing, Mining, Oil and Gas' and 'Educational Services' and the largest negative ones being 'Business, Building and Other Support Services' and 'Manufacturing.' These effects differ when the decompositions at the 10th and 90th percentiles are considered. At the 10th percentile, three categories offer a large negative effect (thus favourable to women): manufacturing, business, and education. At the 90th percentile, the overall effect of industry is positive, the single largest effect coming from the educational services industry.

4.4. Separate analyses by highest level of education

In this section we present figures similar to figures 2 and 3, showing the adjusted wage gaps two and five years after graduation, but this time separately for college/CEGEP graduates and university graduates. We lumped together individuals with bachelor's, master's, and doctoral degrees, given the low number of postgraduate diploma holders. Figures 6a and 6b show the adjusted wage gaps for college and CEGEP graduates, while figures 7a and 7b do the same for university graduates.

For college graduates, the overall shape of the curves for the various cohorts is similar to that found in figures 2 and 3. What differs is the level of the curves: whereas the gaps oscillated around 6% two years after graduation and were mostly under 5% (except for the 1995 cohort) five years after graduation in figures 2 and 3, the gaps are generally in the 6–14% range two years after graduation and more around 10% five years after graduation when only college graduates are considered. This can be explained by the previous finding that having a college education as the highest education level is a negative contributor to the gender wage gap, which would imply that when only college graduates are considered, the wage difference between men and women is larger.

The reverse is true when we look at university graduates in figures 7a and 7b. There the magnitudes of the wage differences are smaller. Figures 7a and 7b also show clearer trends along the wage distribution and across cohorts. The most striking feature of the two figures is how clearly we see an increase of the wage gap along the wage distribution. At the lower percentiles, the wage gap for the period two years after graduation is −1% to 2–3%, steadily increasing to reach 6–10% at the upper percentiles. Five years after graduation, most of the lower half of the distribution actually shows negative gaps (women earning more than men) for all cohorts except 1995. The wage differences increase dramatically to reach close to 15% at the top of the distribution, even after accounting for the explanatory variables in our model. Over time, the wage gaps at the bottom of the distribution appear to have become smaller and those at the top larger, at least for the period five years after graduation. For the period two years after graduation, the curve representing the gaps for the latest cohort (2005 in 2007) in figure 7a lies above those for all the other cohorts between the 10th and 80th percentiles, showing a possible deterioration of the lot of women in recent years.

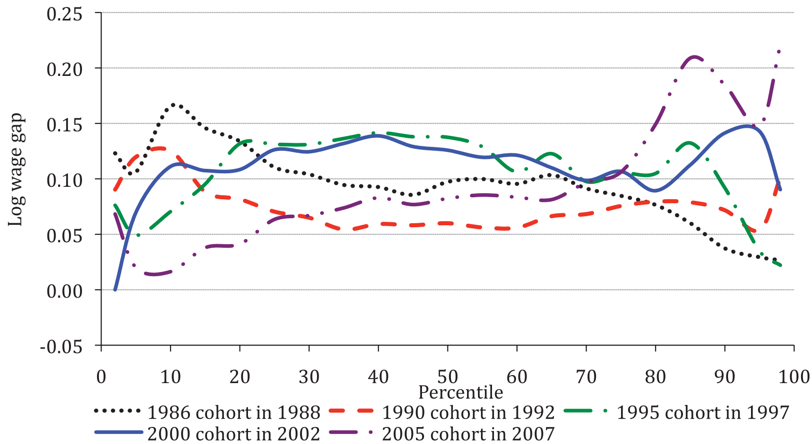

FIGURE 6a Gender log wage gap two years after graduation by percentile (adjusted for covariates), college/CEGEP graduates
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

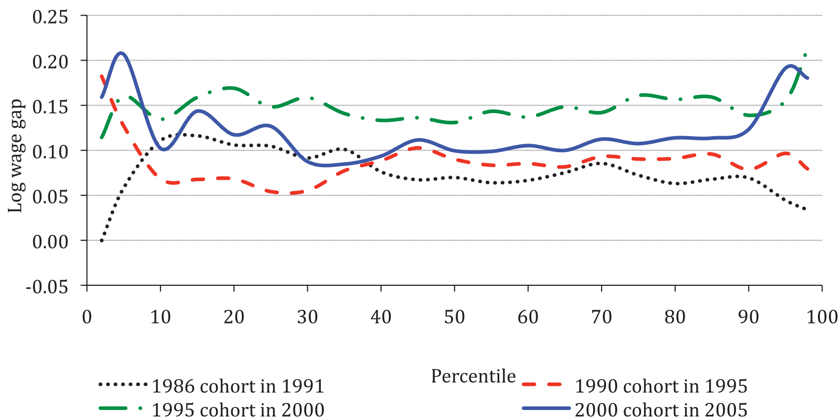

FIGURE 6b Gender log wage gap five years after graduation by percentile (adjusted for covariates), college/CEGEP graduates
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

4.5. Robustness checks and discussion

Next, we performed two robustness checks related to the way we selected our samples. Even if labour force participation is high among recent graduates, we wondered if fertility decisions and parenthood could bias our samples. We restrained our samples to individuals with no children, and found essentially the same results. Note that around 20% of graduates have at least one child two years after graduation, and that mothers have an employment rate of at least 80%.

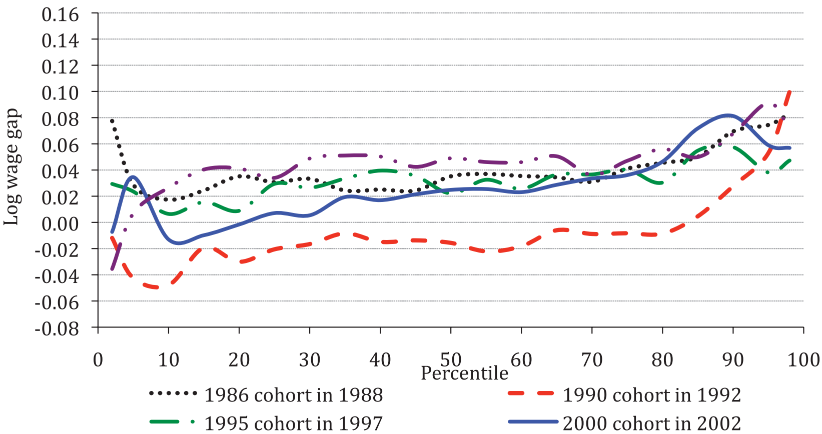

FIGURE 7a Gender log wage gap two years after graduation by percentile (adjusted for covariates), university graduates
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

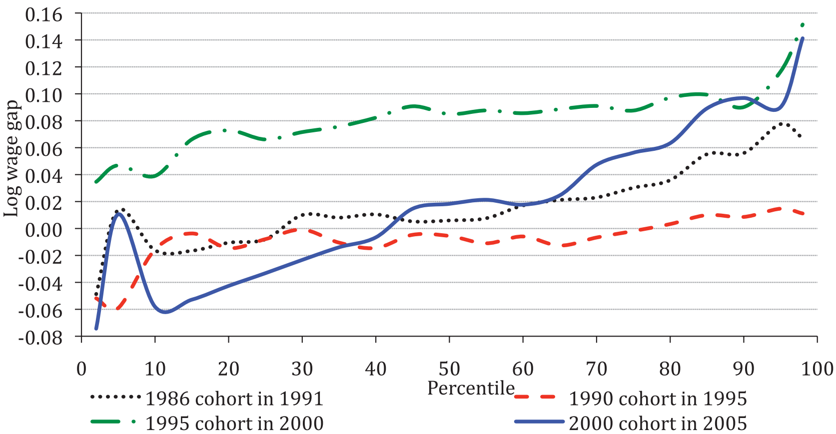

FIGURE 7b Gender log wage gap five years after graduation by percentile (adjusted for covariates), university graduates
NOTE: Each point is the result of a decomposition based on unconditional quantile regressions and shows the value of the adjusted log wage gap.

The second check we did was regarding the types of job that are found at the top of the wage distribution, and more specifically the hours of work performed in those jobs. One explanation for the growing wage gap at the highest percentiles could be that there exists a wage premium (even on an hourly wage basis) for working in high-stress, long-hours jobs, and that men are more willing than women to endure these conditions and reap the benefits. If this explanation were true, we would expect to see work hours rise sharply at the top of the wage distribution. Appendix table 10 in the technical appendix presents the mean of weekly work hours by log hourly wage percentile, separately for men and women and by cohort. Unsurprisingly, we observe that men work more than women all along the wage distribution and in every year of available data. But we also do not observe an increase in hours worked at the top of distribution: if anything, work hours seem to go *down* for the highest wage percentiles. This finding does not support the argument that the wage gap for the best-paid jobs is due to women's preferences regarding long work hours. One possible alternative explanation could be that salaried respondents are more likely to report working a standard 40 hours a week, regardless of their actual hours of work, whereas workers paid by the hour might be more inclined to report actual (perhaps longer) hours. Given that the latter (hourly workers) are on average paid less than salaried workers, this alternative explanation could be consistent with the figures in appendix table 10, while not necessarily invalidating the idea that a growing wage gap at the top of the distribution could be due to jobs requiring longer hours.

Throughout our analysis, we have found consistently small gender wage gaps for the 1990 cohort in both 1992 and 1995. The reasons for this are unclear. To make sure the source of these small gaps is not our methodology, we tried the following two robustness checks. First, we recomputed the gaps using a non-restricted sample (i.e., not restricting based on age, full-time status, self-employment, and no trimming of wages). We found very similar gaps for all the cohorts, including 1990. Next, we also computed the annual earnings gaps instead of the hourly wage gaps. Here we find overall larger gaps, as we would expect, since fewer hours of work for women play a large part in their lower annual earnings, but less so when it comes to wages. However, the discrepancy with the 1990 cohort still remained. It is thus unclear to us exactly why the 1990 cohort displays such smaller wage gaps throughout the wage distribution. We note, however, that there were more graduates with bachelor's degrees in 1990 compared with the surrounding cohorts, and fewer with college degrees (see appendix table 2 in the technical appendix). Additionally, the 1990 cohort does not stand out as much when we conducted our analysis separately for college graduates and university graduates. The higher proportion of university graduates in this cohort may underlie the anomalous result. Another possibility could be that business conditions had a different impact on males and females during those years, helping reduce the gap.

We have documented that recent trends play in favour of women at the bottom of the income distribution, but that wage gaps at the top remain and have widened. What could explain this widening for the best-paid jobs? While it is not the purpose of this paper to test various theories underlining the trends observed, one possible explanation could lie in the composition of the female workforce. We know that women constitute an ever-larger share of the post-secondary graduates. We could imagine that a few decades ago, women with post-secondary education were coming from the upper tail of the skills distribution. Now with more women getting a degree, what we observe on the labour market is additional women selected from a lower part of the skills distribution, thus widening the gap at the top. In short, a woman at the 90th percentile of her wage distribution 20 years ago was probably of higher ability than a woman at the same percentile today. Another explanation could be that highly paid jobs involve more commitment, presence, and investment in human capital while on the job, which could disadvantage women, who on average anticipate shorter and more discontinuous work lives.

5. Conclusion

In this paper, we performed decompositions of the gender wage gap among recent graduates in Canada. Using data for three different cohorts of post-secondary graduates surveyed two and five years after graduation, we first documented mean wage gaps around 6% after two years and 8% and more after five years. We then went beyond the traditional decompositions at the mean and applied FFL's unconditional quantile regression approach to study the wage gap along the whole distribution of wages. At the middle percentiles, wage gaps were similar to the mean wage gap two years after graduation, but were much smaller after five years. We found that while a gap remains even after controlling for a rich set of covariates, women's situation has somewhat improved in the lower half of the distribution between 1986 and 2007. The picture in the upper half of the distribution, especially for the very best-paid jobs, is more worrisome. The unexplained part of the wage gap remained high at the top of the distribution, and the adjusted gap increased across successive cohorts past the 90th percentile. Moreover, the situation now seems to be getting worse over time for a given cohort: the gaps widened in the top half of the distribution when comparing the 2000 cohort in 2002 with the same cohort in 2005. These trends are especially clear and pronounced when university-level graduates only are considered: data for the 2000 cohort in 2005 show gender wage gaps going from a negative 5% (women earning more than men) at the bottom of the distribution to close to 15% at the very top. Since the wage gaps are relatively stable over time for college and CEGEP graduates, it appears that university graduates are driving the widening of the gaps at the top of the wage distribution.

Our study is limited by the fact that we observe graduates only two and five years after graduation. While we believe that considering only graduates is not a major limitation, especially given the growing educational attainment of the Canadian population, the limited time horizon does not allow the full potential of the studies to materialize, nor does it allow us to shed light on the 'glass ceiling' phenomenon that would usually arise later in a woman's career as she climbs the corporate ladder.

In our paper we focused on general trends across the distribution of wages. We also discussed the role of the various explanatory variables in detail as well as the contribution of each level of education, field of study, occupational, and industry categories. Future research could try to understand *why* the gap has increased at high percentiles, a question for which we offered only brief insights in the discussion section. We believe that it is important to pay attention to where Canada stands in terms of wage equality, especially given the laws to that effect that are in force at the various provincial and the federal levels.

References

Albrecht, J., A. Björlund, and S. Vroman (2003) 'Is there a glass ceiling in Sweden?' *Journal of Labor Economics* 21, 145–77

Altonji, J.G., and R.M. Blank (1999) 'Race and gender in the labor market.' In *Handbook of Labor Economics*, Vol. 3C, ed. O. Ashenfelter and D. Card. Amsterdam, New York and Oxford: Elsevier Science, North-Holland

Arulampalam, W., A. Booth, and M. L. Bryan (2007) 'Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution.' *Industrial and Labor Relations Review* 60, 163–86

Baker, M., and M. Drolet (2010) 'A new view of the male/female pay gap.' *Canadian Public Policy* 36, 429–64

Baker, M., D. Benjamin, A. Desaulniers, and M. Grant (1995) 'The distribution of the male/female earnings differential, 1970–1990.' *Canadian Journal of Economics* 28, 479–500

Bayard, J., and E. Greenlee (2009) *Graduating in Canada: Profile, Labour Market Outcomes and Student Debt of the Class of 2005*. Statistics Canada, Culture, Tourism and the Centre for Education Statistics, Catalogue No. 81–595-M – No. 074

Bertrand, M., C. Goldin, and L.F. Katz (2010) 'Dynamics of the gender gap for young professionals in the financial and corporate sectors,' *American Economic Journal: Applied Economics* 2, 228–55

Black, D.A., A. Haviland, S.G. Sanders, and L.J. Taylor (2008) 'Gender wage disparities among the highly educated.' *Journal of Human Resources* 43, 630–59

Blau, F.D., and L.M. Kahn (2007) 'The gender pay gap: have women gone as far as they can?' *Academy of Management Perspectives* 21, 7–23

Blinder, A.S. (1973) 'Wage discrimination: reduced form and structural estimates.' *Journal of Human Resources* 8, 436–55

Booth, A. (2007) 'The glass ceiling in Europe: why are women doing badly in the labour market?' *Swedish Economic Policy Review* 14, 121–44

Boudarbat, B., and T. Lemieux (2010) 'Why are the relative wages of immigrants declining? A distributional approach.' Canadian Labour Market and Skills Researcher Network Working Paper No. 65

Boudarbat, B., T. Lemieux, and C. Riddell (2010) 'The evolution of the returns to human capital in Canada, 1980–2005.' *Canadian Public Policy/Analyse de politiques* 36, 63–89

Boudarbat, B., and C. Montmarquette (2009) 'Choice of fields of study of university Canadian graduates: the role of gender and their parents' education.' *Education Economics* 17, 185–213

Butcher, K.F., and J. DiNardo (2002) 'The immigrant and native-born wage distributions: evidence from United States censuses.' *Industrial and Labor Relations Review* 56, 97–121

Chaykowsky, R., and L. Powell (1999) 'Women and the labour market: recent trends and policy issues.' *Canadian Public Policy* 35, 2–25

Chiswick, B.R., and P.W. Miller (2009) 'The international transferability of immigrants' human capital,' *Economics of Education Review* 28, 162–9

Christofides, L.N., M. Hoy, and L. Yang (2010) 'Participation in Canadian universities: the gender imbalance (1977–2005).' *Economics of Education Review* 29, 400–10

Cool, J. (2010) *Wage Gap between Women and Men*. Social Affairs Division, Parliamentary Information and Research Service, Library of Parliament, Ottawa, Canada, Publication No. 2010–30-E

de la Rica, S., J.J. Dolado, and V. Lorens (2008) 'Ceilings or floors? Gender wage gaps by education in Spain.' *Journal of Population Economics* 21, 751–76

Drolet, M. (2002) 'The male-female wage gap.' *Perspectives on Labour and Income*. Statistics Canada, Catalogue No. 75–001-XPE, Spring 2002, 29–37

Drolet, M. (2011) 'Why Has the Gender Wage Gap Narrowed?' *Perspectives on Labour and Income*. Statistics Canada, Spring 2011, 3–13

Eide, E. (1994) 'College major choice and changes in the gender wage gap.' *Contemporary Economic Policy* 12, 55–64

Ferrao, V., and C. Williams (2011) *Women in Canada: A Gender-based Statistical Report*. Statistics Canada, Catalogue No. 89–503-XWE

Ferrer, A., and A. Menendez (2009) 'The returns to flexible post-secondary education: the effect of delaying school.' Canadian Labour Market and Skills Researcher Network Working Paper No. 20

Finnie, R., and T. Wannell (2004) *The Evolution of the Gender Earnings Gap amongst Canadian University Graduates*. Statistics Canada, Analytical Studies Branch Research Paper Series, Catalogue No. 11F0019MIE – No. 235

Firpo, S., N. Fortin, and T. Lemieux (2009) 'Unconditional quantile regressions.' *Econometrica* 77, 953–73

Fortin, N. (2005) 'Gender role attitudes and the labour market outcomes of women across OECD countries.' *Oxford Review of Economic Policy* 21, 416–38

— (2008) 'The gender wage gap among young adults in the United States: the importance of money vs. people.' *Journal of Human Resources* 43, 886–920

Fortin, N., T. Lemieux, and S. Firpo (2010) 'Decomposition methods in economics.' National Bureau of Economic Research Working Paper No. 16–045

Frenette, M., and S. Coulombe (2007) *Has Higher Education Among Young Women Substantially Reduced the Gender Gap in Employment and Earnings?* Statistics Canada, Analytical Studies Branch Research Paper Series, Catalogue No. 11F0019MIE – No. 301

Hartog, J. (2000) 'Over-education and earnings: where are we, where should we go?' *Economics of Education Review* 19, 131–47

Heywood, J.S., and D. Parent (2012) 'Performance pay and the white-black wage gap.' *Journal of Labor Economics* 30, 249–90

Jann, B. (2008) 'The Blinder–Oaxaca decomposition for linear regression models,' *Stata Journal* 8, 453–79

Kee, H.J. (2006) 'Glass ceiling or sticky floor? Exploring the Australian gender pay gap.' *Economic Record* 82, 408–27

Kiker, B.F., M.C. Santos, and M.M. de Oliveira (1997) 'Overeducation and undereducation: evidence for Portugal.' *Economics of Education Review* 16, 111–25

Neumark, D. (1988) 'Employers' discriminatory behavior and the estimation of wage discrimination.' *Journal of Human Resources* 23, 279–95

Ñopo, H. (2008) 'Matching as a tool to decompose wage gaps.' *Review of Economics and Statistics* 90, 290–99

Oaxaca, R.L. (1973) 'Male-female wage differentials in urban labor markets.' *International Economic Review* 14, 693–709

O'Neill, J. (2003) 'The gender gap in wages, circa 2000.' *American Economic Review. Papers and Proceedings* 93, 309–14

O'Neill, J., and D. O'Neill (2006) 'What do wage differentials tell us about labor market discrimination?' In *The Economics of Immigration and Social Policy*, ed. S. Polachek,

C. Chiswick, and H. Rapoport. Research in Labor Economics, Vol. 24. Bingley, UK: Emerald Group

Phipps, S., P. Burton, and L. Lethbridge (2001) 'In and out of the labour market: long-term income consequences of child-related interruptions to women's paid work.' *Canadian Journal of Economics* 34, 411–25

Statistics Canada (2007) 'National Graduates Survey (NGS) – Detailed Information for 2005 (class of 2000: follow-up).' Last modified 2 May 2007 http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SurvId=20168&SurvVer=3&InstaId=15771&InstaVer=11&SDDS=5012&lang=en&db=imdb&adm=8&dis=2

— (2008) 'National Graduates Survey (NGS) – Detailed Information for 2007 (class of 2005).' Last modified 18 December 2008. http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5012&lang=en&db=imdb&adm=8&dis=2

Wood, R.G., M.E. Corcoran, and P.N. Courant (1993) 'Pay differences among the highly paid: the male-female earnings gap in lawyers' salaries.' *Journal of Labor Economics* 11, 417–41

Yun, M.-S. (2005) 'A simple solution to the identification problem in detailed wage decompositions.' *Economic Inquiry* 43, 766–72