Examining the Gender Wage Gap in Canada in 1999 and 2011
Correcting for Self- Selection Bias

By

Wen-Hung Lue

(6513711)

Major paper presented to the Department of Economics, University of Ottawa,

in partial fulfillment of the requirements of the M.A. program

Supervisor: Professor David M. Gray

Ottawa, Ontario

October 5, 2012

In Canada and in most of the OECD countries, men have historically received higher earnings than their female counterparts. This phenomenon is labeled the gender earnings gap. Modern research of the earnings gap frequently separates it in terms of components that can be explained by observable characteristics and components that are unexplainable and are therefore associated with the presence of discrimination. Drolet (2001) used Canadian data to show that the earnings gap is becoming less explainable and shrinking over time. She believes that the earnings gap is shrinking because of progress made by women in the observable component of the gap. This study, along with many before and after it, is based on the wage equation introduced in Mincer (1974). This equation highlights the importance of education and experience as explanatory variables of the wage. While the equation has stood the test of time by continuing to be the foundation for empirical wage equations, it still suffers from the problem of endogeneity. In a wage regression, the two most prevalent causes of endogeneity are unobserved skill and self-selection bias.

The time series that Baker and Drolet (2010) developed shows that the wage gap has been decreasing over time. In light of the fact that this is happening concurrently with the implementation of pro-female policies that encourage female labour force participation, one must consider the interaction between the variables that determine an individual's wage and the variables that determines an individual's participation in the labour force. Self-selection bias in the form of individuals not participating in the labour force means that many individual's observations are omitted from the Mincerian wage equation. The literature has shown that there are demographic groups in addition, such as minority groups and immigrants, that have the tendency to self-select themselves out of the labour force. This suggests that the Mincerian wage equation may be an imperfect specification of the wage variable for measuring the earnings gap in Canada.

Canadian researchers like Baker and Drolet also acknowledge the issue of self-selection in studying the gender wage gap, but they ignore correcting for it. Typically this is done by dropping all non-wage earners from the sample. Researchers do this because it is difficult to robustly estimate the selection coefficient. I believe that self-selection is a very important factor of wage determination, and therefore it affects the calculation of the wage gap.

In this paper I examine the Canadian gender wage gap in two periods, while taking into consideration the complicating issue of sample selection bias. The Canadian Socioeconomic Information Management (CANSIM) database shows that in 2011, the female labour force participation rate is 62.3% while the male participation rate is 71.5%. This poses a problem as far as investigating the wage gap is concerned, because it means simply using an individual's observed wages ignores information about the "would-be wages" of the individual who is not observed to be earning an hourly wage. This also means that if one were to choose to simply rely on the distribution of observed wages, one would risk either over or underestimating the degree of the discrimination of those who are in the lowest tail of the observed wage distribution. In other words, ignoring the counterfactual wage of those who do not earn an hourly wage will result in a misrepresentation of the wage gap.

In hopes of correcting for the issue of self-selection, I will use the technique introduced in Heckman (1977) to infer a wage variable for nonparticipants and compare the regression results with those generated by the standard Ordinary Least Square (OLS) estimates of the wage equation. This step will be combined with an application of the methodologies proposed by Blinder (1973) and Oaxaca (1973) in order to decompose the earnings gap based on the wage equation estimated using OLS. I use an analogous but extended decomposition that is based wage on equations that are estimated using the Heckman procedure. Oaxaca and Neuman (2004) proposed an alternative approach to decomposing wage equations estimated with Heckman's two-step approach.

Using Canadian Labour Force Survey (LFS) data from 1999 and 2011, I discern an apparent reduction in the magnitude ofgender wage gap over this period. However, the results show that the explained portion of the gap becomes dramatically smaller in the second period relative to the earlier period. When selection is taken into account, the gap widens. Moreover, the size of the explained portion becomes smaller. The selection effect is also stronger for men than for women across both periods.

Recent economic policy in Canada has contributed to an imporvement in the socioeconomic characteristics of females. An important milestone in recent history was the Employee Equity Act of 1995. This policy promoted the representation of four key groups in the Canadian labour force, namely females, aboriginals, people with disabilities, and members of visible minorities. Another noteworthy policy was the 1999 court ruling by the Supreme Court of Canada that removed aerobic standards such as a Bona Fide Occupational Requirement, which served as a barrier to advancement for women. The court ruled that it was discriminatory to hold people to an aerobic standard in the work place if it had no relation to how well the individual performed on the job. In the case of firefighting, women had higher chances of failing the aerobic standard, despite performing equally as well as their male counterparts on the job.

The paper will be divided as follows. In Section 2 I survey both the Canadian and the non-Canadian empirical literature dealing with discrimination as well as the socioeconomic characteristics specific to gender-based discrimination. In this section I also survey some of the literature that deals with the issue of self-selection in the context of wage discrimination. In Section 3, I describe the theoretical framework of this study and the literature that is associated with it. Section 4 contains a description of the data. In Section 5, I explain the empirical methodology used in this study. Section 6 contains the results, and Section 7 provides concluding remarks.

2. Literature Review

A lot of the modern literature on wage discrimination has two primary goals: understanding the impact of particular determining characteristics, and improving the methodology used in estimating the wage equation and calculating the decomposition. Some of the earliest studies in this field were motivated by wage discrimination against minorities in America. Ultimately, a finding that is common amongst all empirical studies on discrimination is that a substantial portion of the observed disparity is largely unexplainable with the conventional choices of observable socioeconomic characteristics and variables.

Becker (1957) is among the earliest studies that formally and thoroughly study discrimination. Becker took an entirely theoretical approach to analyzing discrimination. He examined discrimination within a framework of free-market competitive models and sought to understand the disparity in wages between whites and non-white minorities. He believed that employers have a tendency to engage in discriminatory hiring practices that required non-white minorities to have higher values of wage determining variables given the same wage level. This means that when these variables are used to account for observed wage, they their rate of return is lower than it is for their white counter parts. His explanation for this is that employers have a tendency to treat blacks as more expensive to hire than they actually were. He believes that the prejudice of the marginal employer is what determines the size of the wage gap for the black population. The ultimate implication is that minorities with comparable characteristics to their white counterparts have to either receive lower wages or become more productive than their counterparts in the majority group.

Oaxaca (1973) and Blinder (1973) simultaneously developed a decomposition technique that ultimately became a standard approach to conducting empirical studies on wage discrimination. The decomposition is commonly known as the Oaxaca-Blinder (1973) decomposition because of this coincidence of their efforts.

Oaxaca (1973) is motivated by gender disparities in wages. He starts off by defining discrimination as a situation where the relative wages of individuals in one group are different from "the relative wages that would have prevailed if [each group member] were paid according to the same criteria" (Oaxaca, 1973). His main instrument in modeling discrimination is estimating a wage ratio that is completely absent of discrimination that would serve as a benchmark for measuring discrimination. To do this he assumes that either all females begin working under male wage structure, or all males will start working under the wage structure of females, when there is no discrimination. Using this assumption, he is able to decompose the difference in mean log wages in terms of the unexplained component and the effects of different characteristics and variables pertaining to a particular reference group. He applies his methodology using data from the 1967 Survey of Economic Opportunity. His study focuses on discerning discrimination against women, and he compares the results in both a black and white sample. He also runs these tests using the outputs from both a full scale regression and a regression that only includes personal characteristics. He defines a full regression as a regression that includes both personal and work characteristics, since work-related characteristics such as firm and occupation might be considered choice variables. Using the full scale regression, he finds that the unexplained component accounts for 77.7% and 93.6% of the male/female wage gap for whites and blacks, respectively. These types of results became common in subsequent decomposition studies.

Blinder (1973) developed a very similar decomposition technique in his study. The key difference from Oaxaca (1973) is that he looks at the wage discrimination between blacks and whites and compared the results between genders, instead of looking at the wage discrimination of genders and comparing the results between blacks and whites. He uses data from the 1967 Michigan Survey Research Center's "Panel Study of Income Dynamics". This survey contains income variables from a variety of sources including non-labour income. He finds that the effect attributed to difference in endowments—the explained portion—is negative, meaning that the entire wage gap is unexplained.

Heckman (1979) proposes a two-step estimating method to dealing with the self-selection problem. In essence he believes that the problem of self-selection can be treated statistically like a specification error caused by omitted variable bias. His two-step approach involves estimating a selection equation using a probit model in the first stage of the propensity to participate in the labour force. From this probit regression, the researcher can use the estimates to generate a derived variable called the Inverse Mills ratio. The second stage in his two-step approach involves inserting the Inverse Mills ratio back into the original wage equation and treating it as an independent variable. The basic idea behind his approach is that the wage of those who self-select themselves out of the sample can be inferred from the observed wage equation given that one assumes that the unobserved error term in the wage equation is jointly normally distributed with the unobserved error term in the selection equation. The correlation coefficient between these disturbance terms is incorporated into the coefficient of the Inverse Mills ratio. The wage equation should now be completely corrected for selection effects, and therefore the coefficients of the wage -determining variables should be estimated without bias. On a practical level, it is very important to note that the normality assumption is very difficult to verify and that the results generated after the inclusion of the Inverse Mills have proven to be sensitive to the specification of the wage and selection equations.

Using macroeconomic data, Brown (1984) sought to determine the impact of the US civil rights legislation on the black/white wage differential. This is quite a challenge for several reasons, one of which is that the implementation of the legislation coincided with the development of the Great Society programs, which raised the generosity of social assistance payments. This coincident development would be expected to lower the labour force participation rates for low-wages workers of both races, but to the extent that black workers were over-represented in the left tail of the wage distribution, a disproportionate number of blacks (those in the lower tail) would drop out. One would no longer observe their wages. This would cause a censoring of these observations, and the pattern of censorship would not be the same for the wage distributions of whites versus blacks. Unless one can somehow adjust for the differential in the censoring pattern, the estimates for the impact on the wage distribution of the Civil Rights Acts will be biased upward. Some of the improvement that was attributed to the Civil Rights Acts might actually be due to censoring of the data on wages. Brown claims to have developed and applied an adjustment procedure. I note that he does not use micro data, as it was not yet available at the time that he conducted his research. Instead, he uses aggregated, time-series data on median wages by race.

Dolton and Makepeace (1986) is a UK study that examines the issue of self selection in the gender wage gap using 1970-1977 data from the Department of Employment's Unit for Manpower Studies. They used a sample of predominantly male post-secondary graduates. They compared the equations estimated with OLS with the Heckman correction and a modified estimation procedure that utilized the Maximum Likelihood approach. As is true with most research using decomposition methodologies, much of the disparity is unexplained. Using OLS they found out that female wages would be 20% higher in the absence of discrimination. The estimates based on the Maximum Likelihood and Heckman correction procedure gave values of 14% and 12%, respectively after adjusting for selection effect. Their results also show that a choice of a male reference group will always result in a smaller magnitude of the discriminatory component against women than is the case for the female reference group. They also discover that sample selection bias becomes increasingly significant as certain variables like marriage are omitted from the selection equation.

Yun (2000) also examines the issue of self-selection with an American dataset. He is interested in knowing if there is significant self-selection involved in computing the male-female wage gap of whites compared to non-whites. He uses data from the March 1995 Current Population Survey (CPS) to compare the wage gap derived from the OLS decomposition, with the wage gap derived from the Heckman Two-Step procedure as well as the wage gap derived from the decomposition method he develops called the Generalized Selection Bias (GSB) approach. The main advantage of the new technique was that it allows a researcher to examine the issue of self-selection without being forced to use a probit model. He used a fully parametric Maximum Likelihood Estimation (MLE) model, a method that produces consistent estimators. He shows that the GSB approach generates a considerably smaller magnitude of selection bias compared to Heckman's two-step procedure. He believes this is largely due to the presence of unobserved characteristics.

Blau and Kahn (2005) use the 1994-1996 International Adult Literacy Survey (IALS) to explain the impact of cognitive test scores on wage disparities in Canada, Switzerland, the Netherlands, Sweden, and the United States. This was one of the first times that this regressor was included in a study of wage discrimination. They use an augmented Oaxaca-Blinder decomposition modified by Juhn, Murphy and Pierce (1993). They ultimately find that in the US, performance on cognitive tests explains a significant portion of wage disparity, but the prices of labour market skills are more important in explaining the disparity.

Finzi (2007) examines gender wages disparities in Switzerland in terms of horizontal and vertical segregation using 2000 data from the Enquete Suisse sur la structure des salaires (ESS) and 1999-2003 data from the Swiss Labour Force Survey (SLFS).[1] She finds that both males and females experience a negative effect on their wages when they work in occupations that have a high concentration of female workers. Her estimates reveal a significant coefficient for the Inverse Mills ratio, which could mean a number of things. For instance, this could be a case of forced selection, such that males and females have a tendency to be forcefully selected into segregated occupations in Switzerland. This is another dimension to the selection problem. The author proposes that it may also occur at the hiring stage rather than only at the stage of labour force participation. For her second research question, she finds out that only 30% of women work in "high level" occupations, compared to 55% of men. Additionally, for women, occupational seniority, marital status, and the number of children play very significant roles in explaining wages. The state of being married and the number of children have a highly negative

[1]Horizontal segregation refers to the concentration of males or females in particular fields or occupations. Vertical segregation refers to the concentration of males or females across levels of responsibility or seniority.

impact on wages. For men the estimated coefficients of marital status, the number of children, and experience are all significant. The key difference for males is that all three of these variables have a positive impact on wages. She also finds that the formation dummy, which is the label she uses for education categories, plays an important role because it will always result in higher levels in the job hierarchy and thus higher wages for women.

Badel and Peña (2010) published a paper dealing with the Columbian labour market. They explore the issue of self-selection in gender wage disparities by using quantile regression techniques to estimate a wage equation. They use 2006 data from the Colombian Household Survey (CHS), a monthly cross-sectional survey, and employ the Machado-Mata (MM) decomposition technique, to determine the wage disparity at every quintile of the wage distribution. They ultimately find that female wages lie well below male wages at the tails of the distribution, but less so in the middle of the distribution. They also find that accounting for self selection results in a decomposition that is 50% smaller than what is observed in all quantiles.

Many of the Canadian studies also have focused on examining the explanatory power of particular wage-determining characteristics on the gender wage gap. Doiron and Riddell (1994) examine the issue with a focus on unions, while Baker and Fortin (2001) look at the issue with a focus on occupational gender characteristics. Drolet (2002) looks at the issue with including ther the variables of work experience, education, and job responsibilities. Baker and Drolet (2010) innovatively sought to create a new time series for the wage ratio using only hourly wage data, in contrast to the earnings data that was used in the past.

Baker and Fortin (2001) explore the impact of occupational gender segregation on wages. Their paper uses cross sectional data from the Canadian Labour Market Activity Survey (LMAS) and examines the period between 1987 and 1988. They use a sample of individuals between the ages of 16-69 years who earn more than $1 an hour and are not full-time students. Their main finding is that the relationship between the wage and the gender composition of a job is generally small and statistically insignificant for women, but uniformly negative for men who worked 'female' jobs.

Drolet (2002) updates the Canadian literature by looking at the wage gap with new data that include the following variables: actual work experience, education levels, major field of studies, and job responsibilities. The goal of this paper was to compare the results with previous American studies and as such used variables that were common in American research. She uses the Survey of Labour and Income Dynamics (SLID) to construct her dataset. Her key finding is that average hourly wage of women is approximately 82-89% of what men earn. Additionally, she discovered that the gender difference in work experience explained 12% of the wage gap, while education reduced of the endowment effect by 5%.

Doiron and Riddell (1994) examined the differences in male and female hourly wages during the 1980s, with an emphasis on the unionization effect on the wage gap. Due to data constraints they were only able to examine 3 years worth of data (namely)1981, 1984, and 1988. For 1981 they use data from the Survey of Work History (SWH), for 1984 they use data from the Survey of Union Membership (SUM), and in 1988 they use data from the Labour Market Activity Survey (LMAS). One of the most obvious weaknesses of this approach is that they draw on three different surveys for each of these years. Nonetheless, all the data shared the same sampling frame given the appropriate weights. Moreover, they also use the same definitions for key variables. The main findings of this paper are that the net effect of unionization on the wage gap is small, and that union status had an offsetting effect on gender earnings. They argue that the effect is small in this study because females are less likely to be unionized then men.

Baker and Drolet (2010) attempted to construct a time series of the gender wage gap from 1981 to 2006. The main contribution is that they are now using wage data instead of earnings data to construct the time series. Their data came from a wide variety of sources, including the SWH, the SUM, the LMAS, the SLID, and the LFS. The strengths of drawing from multiple sources is that the data is nationally representative, and it covers a long period of time. The weakness is that taking data from five different sources is problematic when the definition of wage is different across all the surveys. They draw two important conclusions. First, on average the wage ratio is 10-15% higher than the earnings ratio. This result follows the fact that women have the tendency to work fewer hours than men. Secondly, nearly 100% of the fall in the gender wage gap over time is unexplained, and thus the endowment effect is not at work.

3. Conceptual Framework

This paper appeals to neoclassical labour market theory and in order to situate the empirical results and facilitate their interpretation The determinants of an individual's wage are directly tied to his or her productivity. The phenomena of wage discrimination and self-selection are analyzed within the framework of labour supply and demand, and labour force participation. As such, I start this section by discussing the theory of labour supply. I then move on to the theory of wage determination.

One of the earliest attempts to empirically analyze the slope of the labour supply curve was Paul Douglas's *Theory of Wages* (1934). One key study in this book examines the labour supply of foreign and black men in the 1920s, across 38 US cities, in all age groups. This is a significant study because it is an early analysis of the relationship between the wage and labour force participation of a discriminated-against group. Contrary to his expectations, which were for a positive estimate, he finds that individuals have a negative wage-labour quantity supplied correlation across all age categories, but the results were statistically insignificant for all categories except for the very young and very old. Another key study in this book is the analysis of the relationship between hours of work and hourly wages. He ultimately finds that the elasticity of hours worked with respect to hourly wages lies between -0.1 and -0.2. This unexpected result might be generated by a failure to account for endogeneity, since this publication does not employ two-staged least squares.

Mincer (1962) marks the beginning of labour supply literature that takes into account the measurement of income and substitution effects. Mincer observed the labour force participation of married women. This paper identified very important results regarding the labour force participation patterns of married women. He found that a married woman's participation rate is negatively related to the income of her household, but positively related to their wages assuming that her household income is unchanged. In essence, he decomposed the income effect and substitution effects in the context of labour force participation.

The standard formal, static model for labour supply was introduced in Hicks (1946). This model derives labour supply from the theory of consumer demand in the product market. Each individual is maximizing utility, which is a function of personal preferences between income and leisure, the level of consumption of commodities, and the number hours worked subject to a budget constraint. Mathematically, one can express this maximization problem as :

$$\max_{x \in X} U\,(c, h;\ X, \varepsilon) \tag{1}$$

$$\text{s.t. } pc = wh + y \ \text{ and } T = h + \ell$$

where c is the level of consumption of a commodity, h is the number of hours worked, ℓ is the number of hours of leisure, X is a vector of personal characteristics, ε the individual's "tastes" for income versus leisure, p is the price of commodity c, w is the price of each hour of work, and y is all income not associated with work. It is assumed that the utility function is quasi-concave, and that p and w are exogenous. Optimizing the problem presented in (1) will result in the following first order condition: an individual's real wage being equated to the marginal rate of substitution between income and leisure, which is expressed as:

$$\frac{w}{p} = -MRS\ (c, h; X, \varepsilon) = -\frac{\partial U/\partial h}{\partial U/\partial c}. \tag{2}$$

This equation can be rearranged using the same budget constraint to get information on optimal hours worked and the level of goods consumed. For the working population (also known as an interior solution with a positive number of hours worked), this can be expressed as:

$$\mathrm{c} = \mathrm{c}\ (\mathrm{p}, \mathrm{w}, \mathrm{y};\ \mathrm{X}, \varepsilon) \tag{3}$$

$$\mathrm{h} = \mathrm{h}\ (\mathrm{p}, \mathrm{w}, \mathrm{y};\ \mathrm{X}, \varepsilon)\ \text{ both conditional on h} > 0$$

Researchers who are interested in remedying self-selection, should pay particular attention to the corner solution where h=0. The intuition behind the event of non-participation is that there exists a reservation wage that depends on a particular individual's level of non-labour income, socioeconomic characteristics, and personal tastes towards consumption versus leisure. These individuals are not working because their wage offers are below their reservation wage. By convention, I will use w^* to define the unobserved reservation wage. The reservation wage can be expressed mathematically as a function whose form is similar to that of the wage-earning population. This expression is:

$$\frac{w*}{p} = -MRS\ (c, 0; X, \varepsilon), \text{ where } w \leq w* \Rightarrow h = 0. \tag{4}$$

Suppose we know what the true wage distribution is for all individuals in a population. Using the standard notation for the probability density function (pdf) and the cumulative distribution function (cdf), I will call ϕ (w*) and Φ (w*) the respective pdf and cdf when they are evaluated at the reservation wage. Intuitively, Φ (k) can be interpreted as a function that results in a realization of $\Pr(w * \leq k) = \Phi$ (k). It is the probability that a given individual is willing to accept a job with a wage offer of *k* dollars. This can also be expressed as:

$$\Pr(k; p, y, X) = \Phi(k; p, y, X), \tag{5}$$

where Pr(.) is the probability for the choice of working. Another interpretation of this term is the labour force participation. If we take the derivative of (5) with respect to the reservation wage, w* evaluated at k, we obtain the associated pdf

$$\frac{\partial LFP}{\partial k} = \frac{\partial \Phi(w)}{\partial k} = \phi(k) \geq 0. \tag{6}$$

The term in (6) is the change in the probability of working divided by the change in the wage offer of *k*. Dividing (6) by (5) can be interpreted as the growth rate of an individual's labour force participation probability with respect to the wage across his or her wage distribution. This ratio is known as the Inverse Mill's ratio and it gives information about an individual's probability of working across all possible *would be* wages. In empirical applications it is used to produce unbiased estimates of the wage equation in a way that will be discussed in more detail in Section 5.

In most empirical applications, the bulk of the earnings equation is drawn from Mincer (1974). The wage regression utilized in this paper is of the form:

$$ln(WAGE) = \beta_0 + \beta_1 SCHOOL + \beta_2 EXP + \beta_3 EXP^2 + e. \tag{7}$$

Mincer defines a proxy for experience that equals age minus years of schooling minus six. This equation assumes that six is the average age of an individual when he or she begins his or her schooling. A squared term for experience is included in this equation to capture the concavity to the earnings equations. It is assumed that the expected marginal returns of experience will be higher for inexperienced individuals. It is expected that the rate of return of education (β_1) is constant, and the return of experience (β_2) and the return of squared experience (β_3) are positive and negative, respectively.

The use of a log-linear specification over the linear one is motivated by Mincer (1958) who justified the use of the log specification over the level specification of the dependent variable by pointing out the relevance of the distributions of each variable. Earnings distributions are not symmetrically or normally distributed, in contrast to the distribution of the log of earnings, which tends to have a very symmetric and "near-normal" distributions. This is demonstrated in the following section. Another reason for using the log-linear specification is that the parameters reflect rate of returns that can be interpreted as percentages.

4. Data

The data used in this paper come from two waves of the LFS, September 1999 and September 2011. The LFS is a monthly cross-sectional survey of households that provides very detailed information on labour market conditions and socioeconomic characteristics. The unit of observation is the individual. This survey was first conducted in 1945, and its prime objective since then has been to produce data on the working-age population. There are three states of labour market activity: employed, unemployed, or outside of the labour force. The employment information that this survey provides pertains to job-related information (industry, occupation, public and private sector, hours worked, etc…) and regional information (at the level of provinces, territories, and sub provincial regions). For industry data the LFS uses the North American Industry Classification System (NAICS) 2002 as its primary classification scheme. The LFS surveys approximately 54,000 households each month. These individuals consist of civilians above the age of 15 years.[2] One of the biggest strengths of the LFS is its timeliness. Furthermore, parliamentary legislation forcing people to participate in this survey renders it representative. The LFS sample survey consists of 102,740 observations in 1999 and 105,592 observations in 2011.

For both years, I restrict my sample to include individuals who are of prime age (between the ages of 20-60), are not full-time students, and are not self-employed. I choose to select a prime-aged sample because those below the age of 20 are likely still engaged in schooling. Those above the age of 60 are dropped because they are of retirement age, or are approaching it. I exclude full-time students because I do not want to include individuals who are not working as their primary activity. I exclude the self-employed because their earnings are not generated in labour markets. I choose the years of 1999 and 2011 because there is a long time span between them, and both periods have very similar national unemployment rates (around 7.5%). The underlying labour market conditions are broadly similar. With this sampling restriction, the total number of observations is reduced to 32,650 women and 29,171 men in 1999, and 32,650 women and 29171 men in 2011. I note that these exclusions cut the size of the sample by almost half.

[2] The LFS excludes those who are committed to institutions, living on a reserve or Aboriginal settlement, or a full time member of the Canadian Armed Forces.

Table 1 shows the 31 variables used in this study.[3] Figure 1 shows that the log wage has the appearance of a normally distributed, symmetric random variable for both years. My occupational characteristics include the seven industry dummies I generate using the NAICS 2002 codes, a dummy for part-time status, a dummy for union coverage, a continuous variable for tenure of current job in years, a labour market experience proxy, and a continuous variable indexing the number of hours worked.[4] My family characteristics include a dummy for marital status, a dummy for divorce status, and a continuous variable indexing the family size. I have four dummies for educational attainment. I also use 10 provincial dummies as controls for geographical fixed effects. I dropped the indicators for *Ontario*, *Food Services*, and *Undergraduate* from the specification; these indicators thus serve as the reference for their respective sets of dummy variables. The reference person is a male, who is neither divorced nor married, working full time, with no union coverage, from Ontario, working in Food Services, and holding an undergraduate degree. I choose not to include weights in the estimating equations, because weighting does not have a noticeable impact on estimates of the equations estimated either by OLS or by the Heckman procedure.

Perhaps the biggest weakness of my sample is the use of the Mincerian proxy for experience. The main problem with this variable is that it directly contains age and as such is collinear with the *Age variable*. Because these variables are important determinants of both wages and labour force participation, the inverse mills ratio generated by the selection equation will be very sensitive to the choice of these variables. However, since the LFS does not survey individuals about their experience, I am forced to either employ a proxy or omit it altogether. I test various specifications of the wage and selection equations to see if the results generated by

---

[3] Not including the derived variables that are *Age Squared* and *Expected Experience Squared*.

[4] I use the Mincerian proxy for expected experience. I also express tenure of the current job in units of years to keep the units consistent.

Heckman's procedure are robust. I ultimately decided to include the Mincerian proxy, for reasons that will be explained in more detail in the Section 6.

I am interested in both the working and non-working population, and as such I generate six different partitions of the main sample for each year. Each year I partition the main sample by gender, then once again into three groups: the whole sample, the working sample, and the non-working sample. According to the usual definition, labour force participants include the unemployed who are actively seeking work. Following the approach used in Yun (2000), I do not include the unemployed as labour force participants. I choose to exclude this group because in this case, labour force participants only consist of wage earners. By doing this I am able to identify the non-wage earning sample. Of the variables I mentioned, age, experience, the education dummies, the marriage and divorce dummies, and family variables are defined the non-working sample. I used these variables as well as squared age and squared experience as the explanatory variables for my selection equation. I choose to omit years of education as a repressor of all equations because it produces inaccurate measurements for its coefficient.

I chose to express income as the hourly wage following the rationale of Baker and Drolet (2010). They state that it is a more suitable variable for measuring income gaps because wage "corresponds more closely to the price of labour". This is a particularly important point, especially in the context gender disparity. Women have a higher tendency to participate in work that is not full-year, full-time. Since yearly earnings are based on a value that encompasses both prices and quantities (ie. number of hours worked), it cannot be determined whether the disparity is due to a difference in the number of hours worked or to discrimination. Looking at the working samples in Table 2a and 2b, one can see that in 1999 and 2011, females have a considerably higher incidence rate for part-time work than males in part-time work. In 1999

21% of working females are part-time workers compared to 4% of men. In 2011 18% of female workers are part-time workers, while the figure for men remains at 4%.

Looking again at Table 2a and 2b, one sees that the total sample size is roughly the same in both periods; it is 58,525 and 61,821 in 1999 and 2011, respectively. Overall the composition of the sample between males and females changes in favour of females over the two periods. Between 1999 and 2011, the proportion of females who are part of the non-working sample decreases by approximately 3 percentage points, whereas the proportion of males in the non-working sample increases by approximately 2 percentage points. For women this reflects a 10% decrease in their share, and for men this is a 12% increase. In both years, women tend to be more educated than men in the whole sample. This gap appears to be widening in favor of women due to an increase in female participation in graduate school programs. In 1999 the share of all women in every category of education except for the group of individuals with less than a high school education and the bracket of those with a graduate education exceed the share of all men. In 2011 the share of women exceeds the share of men in every category except *Less than High School*.

There are more married and divorced women in the whole sample than men. In 1999 the proportion of women who are married is around 4 percentage points larger than the proportion of men who are married, while the proportion of women who are divorced is around 3 percentage points larger. In 2011 the patter is similar; the percent of women who are married or divorce are higher than they are for men. Based on the findings of Finiz (2007), women's wages have a different relationship with these variables than men's wages and are ultimately more sensitive to variations in these variables. These suggest that the estimated coefficients in the women's participation equation could potentially be different than the case for men.

An interesting fact about the family size variable is that women who are working tend to have smaller families; however men who are working tend to have slightly larger families. This fact appears to be consistent across both years. This pattern is considerably stronger in the male sample in 1999. If I compare the proportions across different participation groups, I can see that working women tend to have families that are 2% smaller than non-working women, a relatively small difference. Working men, on the other hand, tend to have families that are 15% larger than their nonworking counterparts. In 2011 the figures do change much. Working women have families that are 3% smaller than their non-working counterparts, while working men have families that are 14% larger than their non-working counterparts. In 2011 the sizes of families become smaller for both men and women, regardless of whether or not they are working. .

If we look further at the education variable, we see that there is an education gap in favor of women in both the working and non-working samples. The variable *Years of Education* is larger for women than men in both years in both the working and the non-working sample. If we consider anybody with at least a high school diploma, it is clear that males and females with undergraduate degrees make up the largest proportions of their respective educated samples. This is followed closely by the proportion of males and females with high school diplomas. We can see that the proportion of females who hold a high school diploma and the proportion of females who hold undergraduate degrees are both larger than the proportion of males with the same level of education. Interestingly enough, there is a greater proportion of men enrolled in graduate programs than women in 1999, but this also changes in favor of women in 2011. There are is a considerably higher proportion of women employed part-time than men in both years.

In terms of industry of employment, there is a tendency for women to work in the service industries, while men appear to have a higher presence in the non-service industries. This fact seems to hold true in both periods. Overall, the share of women employed in the service industry is larger than the share of men. On the contrary, the primary sector appears to employ a larger proportion of men. The variable *Other Services* is a residual label that includes all services under the Sector 81 label of the NAICS, many of which are traditionally "female" occupations. This includes services such as Repair and Maintenance, Personal and Laundry Services, Private Household Services, and Religious services, to name a few.

The distributions of the variables *Experience*, *Tenure*, and *Union Coverage* have some interesting characteristics. Between the two survey years, the share of women with union coverage becomes larger than the share of men with union coverage. Similarly, the length of female tenure and experience become larger for women when compared to men. In 1999 36% of women are covered by unions compared to 38% for men, whereas in 2011, 37% of women were covered by unions compared to 33% of men. Based on the findings of Doiron and Riddell (1994), I expect this to have an impact in decreasing the wage gap in 2011. The findings regarding the relative value of the tenure variable are similar. In 1999 women remain in their current job for an average of 7.2 years, compared to men who stay for an average of 8 years. In 2011 women have longer tenures than their male counterparts, at 7.5 years compared to men's 7.3 years. In the literature tenure has been used as a proxy for experience, but that is problematic because it undervalues the experience of an individual who frequently is changing jobs. The *Experience* variable appears to be positively correlated with the *Tenure* variable across samples. Much like the case of the tenure variable, females tend to be less experienced than males in 1999, but more experienced than males in 2011.

Overall it appears that the figures for both men and women are relatively similar across time. The changes that did occur appear to be in favour of women. However, since these changes are quite minor I expect that the impact of these variables on the decomposition will also be relatively minor regardless of whether or not selection is taken into consideration.

5. Econometric model

In the following section, I will talk about the methodologies I use to estimate my equations and decompositions.

The Wage Equation

This study uses two different models of wage determination in order to estimate the explained and unexplained components of the gender wage gap. I use the standard OLS technique and Heckman's two-step procedure to estimate wage equations. While both models are log-linear, the second model also depends on a selection equation for the latent variables of the wage and the selection. In the econometric model shown below, individuals are indexed as *i* and groups are indexed by *g*:

$$LW_i^g = X_i^g \beta^g + \varepsilon_i^g, \ where\ i = 1, \dots, n_g\ and\ g = M\ or\ F. \tag{7}$$

Where LW_i^g is log wage of individual i, X_i^g is a matrix of characteristics that impacts an individual's wage given that they are a part of group *g*, and β^g is a vector of coefficients associated with each characteristic. ε_i^g is a vector of i.i.d. white noise terms associated with each individual for group *g*.

The problem with this wage equation is that it only involves the observed wage, which considers only n_g individuals in group *g* who explicitly earn a wage. Those who are not a part of the wage earning sample will be completely dropped from the estimation. This is inappropriate because it means that the estimating sample is no longer completely random and representative of the labour force. When we seek to estimate the wage determination equation, unless the underlying relationships are the same for those who are included and those who are excluded from the sample, we will obtain a bias for the estimated coefficients.

Using the approach presented in Heckman (1979), I now express log wage as a latent variable that depends on a latent selection equation such that:

$$LW *_j^g = X_j^g \beta^g + \varepsilon_j^g, \tag{8}$$

$where\ LW *_j^g = LW\ if\ S *_j^g > 0, \quad else\ 0$

$$S *_j^g = Z_j^g \gamma^g + \varphi_j^g, \tag{9}$$

$where\ j = 1, \dots, n_g\ and\ g = M\ or\ F.$

$Selection_j^g = 1\ if\ S *_j^g > 0\ \ and\ S_j^g = 0\ if\ S *_j^g \leq 0$

In the above equation, $S *_j^g$ is an unobserved latent variable, Z_j^g is a matrix of socioeconomic characteristics that explains an individual's propensity to participate in the labour force given that they are a part of group g, γ^g is a vector of coefficients associated with Z_j^g, and φ_j^g is an i.i.d. white noise term associated with group g.

The primary feature of this wage equation is that ε_j^g is no longer i.i.d in the presence of selection. This is because $LW *_j^g$ is now a partially latent variable that is only observed if $Selection_j^g = 1$. The significance of this point is that we should adjust the wage determination equation to account for the wages of non-participants. This can only be done if the proper exclusion restrictions are made in the selection equation. That is, there should be at least one variable present in the selection equation but excluded from the wage equation. The non-linearity of the form of the selection equation and the linearity of the wage equation make it possible for identification without exclusion restrictions, however it is highly recommended that one applies a restriction involving the inclusion of variables in the selection equation to achieve identification.

Selection bias can be interpreted as a form of omitted variable bias. This can be shown by expressing the expectation of the outcome variable conditional on the observables variables, Z_j^g, and the noise term, φ_j^g, in the selection equation.

$$E\left[LW *_j^g \left|Z_j^g, \varphi_j^g\right.\right] = X_j^g \beta^g + E[\varepsilon_j^g | Z_j^g, \varphi_j^g] \tag{10}$$

$$= X_j^g \beta^g + \beta^\lambda \varphi_j^g.$$

In the above equation, β^λ is the coefficient measuring selection bias. A standard assumption of the Heckman two-step procedure is that the variables $LogWage *_j^g$ and $Selection *_j^g$ have a bivariate normal distribution in ε_i^g and φ_j^g. This allows us to go from the second to the third line in (10). The main issue with (10) is that φ_j^g is clearly unobservable. Therefore an estimator must be computed for the purpose of empirical analysis, a role that is fulfilled by the inverse mills ratio .

To derive an unbiased estimator for the selection variable, which is the first step of Heckman's two-step procedure, the law of iterative expectations can be used to create an expression for the expected value of $LW *_j^g$ conditional on observable variables, Z_j^g, and $Selection_j^g$. We first substitute the selection variable in for the white noise term φ_j^g to get:

$$E\left[LW *_j^g \mid Z_j^g, S_j^g\right] = E[E\left[LW *_j^g \mid Z_j^g, \varphi_j^g\right] | Z_j^g, S_j^g]. \quad (11)$$

Combining (10) with (11) gives us:

$$E\left[LW *_j^g \mid Z_j^g, S_j^g\right] = E[X_j^g \beta^g + \beta^\lambda \varphi_j^g | Z_j^g, S_j^g] \quad (12)$$

$$= X_j^g \beta^g + \beta^\lambda E[\varphi_j^g | Z_j^g, S_j^g].$$

In our model we assume that all labour force participants are assigned a value such that $Selection_j^g = 1$. As such, we need to find $E[\varphi_j^g | Z_j^g, Selection_j^g = 1]$. Our assumptions about the model imply that :

$$E\left[\varphi_j^g | Z_j^g, S_j^g = 1\right] = E[\varphi_j^g \mid \varphi_j^g > -Z_j^g \gamma^g] \quad (13)$$

$$= \frac{\phi(-Z_j^g \gamma^g)}{1 - \Phi\left(-Z_j^g \gamma^g\right)} = \frac{\phi(Z_j^g \gamma^g)}{\Phi\left(Z_j^g \gamma^g\right)}$$

$$= \lambda(Z_j^g \gamma^g) \quad .$$

The derivation of the second line of (13) can be found on page 785 in Greene (2003). Assumptions about symmetry in a normal distribution allow us to go from line two to line three of (13). Using these results we can now fully parameterize expression (10) such that:

$$E\left[LW *_j^g \mid Z_j^g, S_j^g = 1\right] = X_j^g \beta^g + \beta^\lambda \lambda(Z_j^g \gamma^g) \quad , \quad (14)$$

where $\beta^\lambda = \rho_{LW,S}^g \sigma_{LW}^g$.

In the above equation $\rho_{LW,S}^g$ is the correlation between the disturbance term of the wage equation and the selection equation, and σ_{LW}^g is the standard deviation of the wage equation. This factor serves to scale the β^λ coefficient such that it is consistent with the wage variable in the wage determination equation.

Decomposition Techniques

The decomposition techniques I use separate the explainable component from the unexplainable components. These components are differentiated based on portions attributed to observable characteristics (i.e. the endowment effect) and portions attributed to the returns of the attributes as they generate earnings, which are reflected in the coefficients.

The means of the variables are always easy to calculate—the point is that the equation estimated via OLS might yield biased results. When applying OLS, we assume that the noise term, ε_i^g, is i.i.d., which implies that it is conditionally uncorrelated with the observable characteristics, X_i^g. That is, we are assuming that $E\left[\varepsilon_i^g | X_i^g\right] = 0$. Following this line of reasoning, taking the conditional expectations of (7) one obtains:

$$E\left[\, LW_i^g | X_i^g\right] = E[X_i^g]\, \beta^g. \tag{15}$$

The fitted equation (15) evaluated at the sample mean is used in the decomposition and can be expressed as:

$$\overline{LW^g} = \overline{X^g} b^g, \tag{16}$$

where b^g is the OLS estimator for β^g. As described in Oaxaca (1973), the difference in means can be decomposed if one chooses the reference group. The choice of the reference group is important for determining the decomposition. The two possible decompositions that can be created from (16) are:

$$\overline{LW^m} - \overline{LW^f} = \left(\overline{X^m} - \overline{X^f}\right) b^f - \overline{X^m}(b^m - b^f), \text{ or} \tag{17}$$

$$\overline{LW^m} - \overline{LW^f} = \left(\overline{X^m} - \overline{X^f}\right) b^m - \overline{X^f}(b^m - b^f). \tag{18}$$

The right hand side of (17) and (18) show different decompositions of the wage differential. In both of these decompositions, the term on the left is the explained portion of the disparity, and the term on the right is the unexplained portion of the disparity. These terms can also respectively be referred to as the endowment and coefficient effects. In equation (17) the difference in the means of the characteristics and determining variables are remunerated according to the estimated coefficients of the female population, while the difference in the coefficients are applied to the mean values of the characteristics of the males. Equation (18) is set up according to the reverse. The choice of (17) and (18) is simply a decision of whether one believes that males or females are the non-favored group. For the purpose of this study, I use equation (17), because I believe that females are the ones who are being discriminated against. I want to know how much females would earn given a male wage structure after adjusting for sample selection bias.

Using a wage equation derived by Heckman's two-step procedure, I must use a slightly different decomposition. To do this first one must derive the sample analog of (14), which can be expressed as:

$$\overline{LW^g} = \overline{X^g} b^g + b^\lambda \overline{\lambda(Z^g \gamma^g)}. \tag{19}$$

A common approach, involves taking the difference of (19) across the two genders such that we similarly get:

$$\overline{LW^m} - \overline{LW^f} = \left(\overline{X^m} - \overline{X^f}\right) b^f - \overline{X^m}(b^m - b^f) + (\theta^m \overline{\lambda^m} - \theta^f \overline{\lambda^f}), \text{ or} \tag{20}$$

$$\overline{LW^m} - \overline{LW^f} = \left(\overline{X^m} - \overline{X^f}\right) b^m - \overline{X^f}(b^m - b^f) + (\theta^m \overline{\lambda^m} - \theta^f \overline{\lambda^f}) \tag{21}$$

Much like equations (18) and (19), in both of these decompositions, the term on the left is the explained portion of the disparity, the term next to it is the unexplained portion of the disparity, and the term on the far right measures the size of the selection bias. Oaxaca and Neuman (2004) suggest that (20) and (21) are inappropriate because the term $(\theta^m \overline{\lambda^m} - \theta^f \overline{\lambda^f})$ can be decomposed even further in terms of coefficient and endowment effects. They believe that the selection term can be expressed as:

$$(\theta^m \overline{\lambda^m} - \theta^f \overline{\lambda^f}) = \theta^m \left(\overline{\lambda_0^f} - \overline{\lambda^f}\right) + \theta^m \left(\overline{\lambda^m} - \overline{\lambda_0^f}\right) + (\theta^m - \theta^f)\overline{\lambda^f}. \tag{22}$$

The new term, $\overline{\lambda_0^f}$, is the mean value of the Inverse Mills ratio if females faced the same selection equation as men. Inserting (22) into (21) produces an alternative two-fold decomposition:

$$\overline{LW^m} - \overline{LW^f} = [\overline{X^f}(b^m - b^f) + \theta^m \left(\overline{\lambda_0^f} - \overline{\lambda^f}\right) + (\theta^m - \theta^f)\overline{\lambda^f}] + \tag{23}$$

$$[\left(\overline{X^m} - \overline{X^f}\right) b^m + \theta^m \left(\overline{\lambda^m} - \overline{\lambda_0^f}\right)]$$

$$= [\overline{X^f}(b^m - b^f) + \theta^m \overline{\lambda_0^f} - \theta^f \overline{\lambda^f}] + \left(\overline{X^m} - \overline{X^f}\right) b^m + \theta^m \left(\overline{\lambda^m} - \overline{\lambda_0^f}\right)$$

6. Results

Table 4 shows the outputs for both the equation estimated by OLS and the equation estimated by the Heckman procedure. I will commence by discussing the OLS estimates. As far as statistical significance is concerned, for the 1999 sample, the coefficients of all of the variables except for number of hours worked and the agricultural sector dummy are significant at a 99% confidence level for women, as shown in the top panel of Table 4a. For men the coefficients of every variable except the Alberta dummy are statistically significant at a 95% confidence level, as shown in the bottom panel of Table 4a. In 2011 the coefficients for females are statistically significant at a 99% confidence level except for the experience variable, the numbers of hours worked, and the British Columbia dummy, as shown in the top panel of Table 4b. For males, only the experience term is statistically insignificant, as shown in the bottom panel of Table 4b.

In 1999 the estimated coefficient of the industry variables, the education dummies, and the union coverage variable were mostly numerically significant in the female wages equation. Looking at the industry dummies, one can see that women who work in the wholesale, retail, or transportation sector tend to earn considerably less than those in the other sectors, but around 12% more than those in the reference category. On the other hand, those who are in the service, primary, or public administration sectors tended to earn a significant premium: 30%, 44%, and 37%, respectively. There also appears to be a strong positive relationship between wages and holding a graduate degree, but a weak negative relationship for anyone holding any diploma below an undergraduate degree or college certificate. Graduate degree holders earn 19% more, while those with high school diplomas and individuals below the high school diploma earned 6% and 7% less than the reference group. Having union coverage will increase earnings by 17%, while an additional year of tenure is expected to increase earnings by only 2%. Being a part-time worker decreases the wages of women by 8%. The regression results also show that the wage grows as a positive yet concave function of age, but as a negative yet convex wage function of experience. The former of the two findings is consistent with the model of Mincer (1958), but the negative estimate for experience is nonsensical and most likely a result of multicollinearity between the variables of age and experience. It is important to recognize that the magnitude of the estimated coefficients of both pairs of variables are numerically insignificant; as such the relationship is likely very weak in either direction.

In the case of men, the industry dummies, the education dummies, and the part-time worker status were statistically significant and had the largest magnitudes. The estimated coefficient of the industry dummies appear to be as important in the case of men as they were for women. The main difference with the coefficients of the male industry variables is an overall higher increase in magnitude relative to the reference group when compared to the results for females. Men in primary sector tend to earn the most, at 48% more than the reference group, while men in the agricultural sector tended to earn the least—only 11% more than the reference group. This means that men in the reference group earn the least. The magnitude of the estimated coefficients of the remaining industry groups were also relatively high, lying between 25-40%. The education dummies appear to show that among men, graduate degree holders have the same wage premium as females, at 19% more than the reference group. However, the negative effect of holding a high school diploma is much larger amongst the male population at -9.2% for high school diploma holders, and -14% for any lower educational category. Being a male, part-time worker, is associated with earning 27% less than those who are full-time workers. The fact that this figure is so much larger than the figure for females suggests that it is probably also an institutional factor in play that makes the "penalty" for working part time lower for women. For men the estimated coefficient value of 7% for the union coverage variable is relatively small compared to the female number of 17%. This smaller estimate is likely an outcome of a higher presence of men in highly paid, non-union jobs. This is also consistent with the findings of Card (2001) and Lemieux (1993), who find that women benefit more than men from unionization. The results for the variables of age and experience are near identical to those for females, and the same reasoning as I presented before holds true for men. The estimated coefficient for male tenure is very small, at around 2% per year of tenure.

Overall it seems that men and women appear to be similar in their returns from graduate education, tenure duration, and industry characteristics. The largest differences appear to be with respect to the union coverage, part-time work status, and the two lowest education categories.

For the data set in 2011, the term the British Columbia dummy, and the variable indexing the number of hours worked are not significant at a 90% or greater level of confidence for women.[5] For both women and men, the estimated effects of tenure, union coverage, and part-time work status have remained relatively unchanged between 1999 and 2011, despite the radical change in the mean values of these worker characteristics. In the case of women, the estimated industry effects do not change a lot. Women in the wholesale, retail, and transport sectors do, however, experience a relatively steep decrease from a 12% to an 8% earnings premium compared to the reference group. The estimated coefficients of 2011 education dummies show that women who have less than a high school degree also earn 5% more than the reference group, whereas males earn 7% less. For males the estimated coefficients of the education dummies remains relatively fixed after a 12-year period.

While theoretically elegant, one of the biggest problems with Heckman's procedure is that it is tends not to produce empirically robust results. As mentioned earlier, the estimated coefficient of the Inverse Mills ratio tends to be sensitive to the specification of the wage and the selection equation. Given my data sample and specifications problems could potentially arise

[5] Most are significant at a 99% confidence interval.

because of strong correlations between the explanatory variables of age and experience. As shown in Table 1, I am using an experience proxy that contains age, and so the two variables have a large risk of being multicollinear. To address this problem, I estimate three specifications of the Heckman procedure with various modifications regarding the age and experience variables. In Table 3 I show the estimated coefficients of the Inverse Mills ratio in all three specifications. The results are unusual and unintuitive, since we expect for the selection effect of men to be a positive. However, the results do appear to be robust across the three specifications for men in both years. In the case of women, the results are very non-robust. The coefficients for all three specifications are negative for women in 1999. Additionally, they are only significant for the specification including either *Age* or *Experience*. In 2011, only the specification including both *Age* and *Experience* is significant and with a positive coefficient for the estimated coefficient fo the Inverse Mills ratio.

Contrary to what I expected, in all three cases, the selection issue is actually stronger for men than for women, and the sign of the selection coefficient is negative for men. Because the results for males are relatively robust in all three specifications, my choice of specification was motivated by the coefficients of the lambda parameter for females. In the specification including both *Age* and *Experience* as regressors, the positive and significant estimate for females in 2011 is very natural and believable. It is intuitive to expect the presence of characteristics specific to women, such as take-up of maternity leave or the tendency to work part-time jobs, to cause a disproportionate number of them to be non-participants in the labour force.

It is worth mentioning that Choudhury (2001) found that "many economists do not pay to the precise interpretation of the lambda term"; it is a more standard practice to simply look at the significance of the term to determine the presence of unobserved factor pushing towards participation but pushing down the accepted age, or vice versa. Intuitively, a negative coefficient means that there exists some unobserved factor that is positively related to the selection equation but negatively related to the wage equation, or vice versa. Because the results for males are relatively robust in all three specifications, my choice of specification was motivated by the fact that having both *Age* and *Experience* as regressors provided an intuitive result for 2011.

The estimates from the regressors that apply the Heckman procedure are almost entirely unchanged relative to the results generated from the OLS procedure.[6] Based on the 1999 sample, most coefficients become smaller in magnitude after applying the Heckman procedure when compared to the simple OLS results for both genders. Based on the 2011 sample, all coefficients are smaller except for the ones for the male sample. In the case of women, the coefficients are slightly larger. This is probably the most important finding. While the effect of selection does not have a particularly noticeable presence on the estimated coefficients for any particular variable, it does appear to affect the estimated coefficients of every variable, which can ultimately cause large differences in the decomposition of the overall wage gap.

I present the outputs of the probit equations in Tables 5a and 5b. These regressions include *Divorce Status*, *Marital Status*, and *Family Size* as the exclusion restrictions. I chose these variables because they have historically been associated with female self-selection and ought to affect wage determination less. I expect that these variables should perform well in resolving the identification issue.

[6] This outcome is consistent with the results of Dolton and Makepeace(1986)

Based on the 1999 sample, women's probability of participating in the labour force is positively related to marriage, but negatively related to family size. For men the coefficient values for both these variables are positive. Additionally, men who are divorced are more likely to participate. In 2011 there is very little significant change for either men or women compared to the estimates for 1999. A notable feature of the selection equation is that the effect is negative for very high and very low levels of education for both genders, but the magnitude of the coefficients is noticeably higher for women.

Table 6 contains the outputs for the Oaxaca-Blinder decomposition using the equation that is estimated by OLS. One can see that the mean unadjusted differential is 0.228 in 1999 and 0.142 in 2011. Additionally, the total explained portion of the wage decomposition appears to decrease from 20% to 12%. Baker and Drolet (2009) find that the gap becomes less explainable the closer males and females become in terms of their characteristics. For the 1999 wave, it appears that the role of observable characteristics such as union status is relatively small. The variables *Age, Expected Experience, Less than High school, Other Services* and *Hours of Work* have a negative effect, and therefore explain nothing in regards to the gap. The geographical controls also had a very small positive effect. Overall, the industry choice, tenure duration, and part-time status variables had the largest explanatory power for the wage gap. In 2011 there are more variables that have a negative effect. These variables include: *Age, Union Coverage, Tenure, Hours Worked, Other Services, Public Administration,* and all the education levels.

Table 7 contains the outputs for the decomposition using a Heckman wage equation. The most notable findings regarding these decompositions are that the gaps as calculated by the fitted values of the wage equation are comparatively larger than the ones generated by the standard OLS decompositions. The gap is 0.336 in 1999 and 0.225 in 2011. Additionally, the explained portion of the decomposition is smaller in both years. 15% of the gap was explained in 1999, and 9% was explained in 2011. The decomposition using the Heckman procedure also indicates that size of the selection bias is becoming smaller between the years 1999 and 2011, dropping from 0.108 to 0.082.

The increase in the gaps between the OLS decomposition and the decomposition derived from equations estimated by the Heckman procedure are directly related to wage equations. Recall that the Inverse Mills ratio is large, negative, and significant for men in both the years 1999 and 2011, but shrinks slightly in magnitude in the year 2011. For women it is small, insignificant, and negative in 1999; and slightly larger, significant and positive in 2011. The important point is that the magnitude of the male selection coefficient is much higher than the female selection coefficient for both years. In 1999 this implies that both females and males are both more likely to accept a wage offer that is below their reservation wage. This also means that the estimated wage equations of females and males are both going to increase between the OLS estimate and the equation estimated by the Heckman procedure, but the increase will be greater for males than for females because the male selection coefficient is of a larger magnitude. The reason for this is because the wage gap calculated by the Heckman procedure take the pure wage gap and adds on the selection term. This explains why the wage gap increases in 1999. In 2011 the wage gap increases for a similar reason. The selection effect will change the wage equation of males in the same way in 2011 as it did in 1999, however for females a positive selection coefficient suggests that they are more likely to recieve a wage offer that is above their reservation wage. This means that the estimated wage equation for females will decrease between the OLS estimate and the equation estimated by the Heckman procedure.

In Table 7a and 7b it is clear that the size of the selection bias decreases between 1999 and 2011, following the decrease in the wage gap over time. However, its share of the total wage differential becomes larger over time. In 1999 the selection bias makes up 33% of the true wage gap, and in 2011 it makes up 37% of the wage gap.

7. Conclusion

In this paper I use the 1999 and 2011 waves of LFS to compare the gender wage gap in different periods while correcting for sample selection bias. Overall, the wage gap decreases over these two periods, however the addition of the selection variable ultimately increases the size of the adjusted wage gap and decreases the size of the explained portion of the differential. This suggests that the true wage gap is actually larger than what is typically observed. My initial goal in conducting this study was to investigate the impact of self-selection on wage decompositions. I did so with the expectation that the effect would be considerably more significant in the female sample. To my surprise the effect was not only more significant in the male sample, but also hugely negative. In all three specifications that were estimated by Heckman's procedure, the effect of the selection variable was strictly negative and significant for both periods, while the effect for females was non-robust. Testing with different wage and selection specifications with a different data set might be an interesting and worthwhile exercise for future research. It is possible that much of the noise that complicates the analysis is due to the limitations in the availability of variables such as measurable experience and non-wage income in the LFS. I am curious to see how this will impact the size of the selection bias on the wage gap. I am also interested in knowing if adding tenure spent in previous jobs will change anything.

Bibliography

Badel, A. and X. Peña, "Decomposing the Gender Wage Gap with Sample Selection Adjustment: Evidence from Columbia", Federal Reserve Bank of St. Louis series of Working Papers, 2010.

Baker, M. and M. Drolet. "A New View of the Male/Female Pay Gap.", Canadian Public Policy, 36, 2010, 429-464.

Baker, M. and K, Milligan, "How Does Job-Protected maternity Leave Affect Mother's Employment?", Journal of Labor Economics, 26, October 2008, 655-692.

Baker, M. and N. Fortin, "Occupational Gender Composition and Wages in Canada: 1987-1988", Canadian Journal of Economics, 34, May 2001, 345-376.

Becker, G., "Economics of Discrimination", University of Chicago Press, 1957.

Blau, F. D. and L. M. Kahn, "Do Cognitive Test Scores Explain Higher US Wage Inequality?," The Review of Economics and Statistics, 87, 2005, 184-193.

Brown, C., "Black-White Earnings Ratios Since the Civil Rights Act of 1964: The Importance of Labor Market Dropouts." The Quarterly Journal of Economics, 99, 1984, 31-44.

Blinder, A., "Wage Discrimination: Reduced Form and Structural Estimates.", The Journal of Human Resources, 8, 1973, 436-455.

Card, D., "The Effect of Unions on Wage Inequality in the U.S. Labor Market", Industrial and Labor Relations Review, 54, 2001, 296-315.

Choudhury, S., "Reassessing the Male-Female Wage Differential: A Fixed Effects Approach", Southern Economic Journal, 60, 1993, 327-341.

Doiron, D. and W.C. Riddell, "The Impact of Unionization on Male-Female Earnings Differences in Canada", Journal of Human Resources, 1994, 29, 505-534.

Dolton, P. J. and G. H. Makepeace, " Sample Selection and Male-Female Earnings Differentials in the Graduate Labour Market", Oxford Economic Papers, 38, 1986, 317-341.

Douglas, P. H., "The Theory of Wages", 1934, New York: Macmillan.

Drolet, M., "New Evidence on Gender Pay Differentials: Does Measurement Matter?", Canadian Public Policy, 28, 2002, 1-16

Finzi, I., "Occupational Gender Segregation and Gender Wage Gap in Switzerland", PhD. dissertation, University of Lugano, 2007

Greene, William H., "Econometric Analysis", Upper Saddle River, NJ: Prentice Hall, 2003.

Heckman, J., "Sample Selection Bias as a Specification Error", Econometrica, 47, 1979,153-161.

Hicks, J. R., "Value and Capita: An Inquiry into Some Fundamental Principles of Economic Theory ", 2nd Ed. Oxford University Press, 1946

Jann, B., "The Blinder-Oaxaca Decomposition for Linear Regression Models", The Stata Journal, 8, 2008, 453-479.

Jovanovic, B., "Job Matching and the Theory of Turnover", Journal of Political Economy, 87, 1979, 972 990.

Lemieux, T., "Unions and Wage Inequality in Canada and the United States", Chicago: University of Chicago Press, 1993.

Oaxaca, R., "Male-female wage differentials in Urban Labor Markets", International Economic Review, 14, 1973, 693-709.

Mincer, J. A., "Investment in Human Capital and Personal Income Distribution," Journal of Political Economy, 66, 1958, 281.

Mincer, J. A., " Labor Force Participation of Married Women", In Conference of the Universities National Bureau Committee for Economic Research, Aspects of Labor Economics, 1962, 63-97.

Mincer, J. A., "The Human Capital Earnings Function", National Bureau of Economic Research Chapters, 1974, 83-96.

Neuman, S. and R., Oaxaca, "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note", Journal of Economic Inequality, 2, 2004, 3-10.

Pencavel, J., "Labor supply of men: A Survey", Handbook of Labor Economics,1987, 3-102. edition 1, volume 1, number 1.

Willis, R. J., "Wage determinants: A survey and reinterpretation of human capital earnings functions", Handbook of Labor Economics, 1987, 525-602. edition 1, volume 1, number 1.

Yun, M.,"Selection Bias and the Decomposition of Wage Differentials", Rutgers University Departmental Working Papers, 2000.

Appendix

Table 1: Variable Definitions

Log Wage	A continuous variable that takes the log of all wage earning individuals Omitted for all non-wage earners.
Age	A continuous variable that represents an individual's age.
Years of Education	A continuous variable that represents years of schooling of an individual.
Divorced	1 if divorced, 0 if not.
Married	1 if married, 0 if not.
Family Size	A continuous variable that represents the size of an individual's family
Expected Experience	Age-Years of Education-6
Union Coverage	1 if covered by a union, 0 if not. Omitted for all non-wage earners.
Tenure	A continuous variable that represents an individual's tenure in his or her current job in unit of years. (This starts at 1/12 of a year and topcodes those who worked for over 20 years in their current job)
Part Time	1 if working part time, 0 if not. Omitted for all non-wage earners. Omitted for all non-wage earners.
Hours Worked	The number of hours an individual works in a week. (lies between 1-99) Omitted for all non-wage earners.
Provincial Dummies	1 if (insert province), 0 if not.
Less Than High school	1 if an individual has less than 8 years of lifetime education, 0 otherwise.
High school	1 if an individual has at least some high school education or a high school diploma/certificate, 0 otherwise.
Undergraduate	1 if an individual has a college certificate or a undergraduate degree, 0 otherwise.
Graduate School	1 if an individual has a graduate degree, 0 otherwise.
Agricultural Sector	1 if NAICS 2002 code is 1100-1129, 1151-1152 , 0 otherwise. Omitted for all non-wage earners.
Primary Sector	1 if NAICS 2002 code is 1131-1142, 1153, 2100-2131, and 2211-2213 , 0 otherwise. Omitted for all non-wage earners.
Utility, Construction and Manufacturing Sector	1 if NAICS 2002 code is 2361-2389, 3211-3219, 3271-3279, 3311-3379, 3111-3169, 3221-3262, and 3391-3399, 0 otherwise. Omitted for all non-wage earners.
Wholesale, Retail, and Transportation Sector	1 if NAICS 2002 code is 4111- 4191, 4411- 4543, and 4811-4931, 0 otherwise. Omitted for all non-wage earners.
Food Services	1 if NAICS 2002 code is 7211 - 7224, 0 otherwise.
Other Services	1 if NAICS 2002 code is 5211-5331, 5411-5419, 5511- 5629, 6111-6117, 6211-6244, 5111-5191, 7111-7139, and 8111-8141, 0 otherwise. Omitted for all non-wage earners.
Public Administration	1 if NAICS 2002 code is 9110 - 9191, 0 otherwise. Omitted for all non-wage earners.

Table 2a: Summary Statistics for the 1999 wave

	Women		Men	
	Mean	S.D.	Mean	S.D.
Whole Sample				
Age	40.027	10.437	39.572	10.551
Less than High school	0.199	0.399	0.227	0.419
High school	0.300	0.458	0.279	0.449
Undergraduate	0.461	0.498	0.443	0.497
Graduate School	0.041	0.197	0.051	0.220
Divorced	0.095	0.293	0.069	0.253
Married	0.722	0.448	0.684	0.465
Years of Education[7]	14.113	3.401	13.991	3.590
Family Size	2.982	1.236	2.919	1.298
Sample Size	31043		27482	
Non-Working Sample				
Age	42.178	11.114	42.661	11.662
Less than High school	0.358	0.479	0.415	0.493
High school	0.299	0.458	0.243	0.429
Undergraduate	0.323	0.468	0.313	0.464
Graduate School	0.020	0.141	0.030	0.171
Divorced	0.092	0.289	0.106	0.308
Married	0.732	0.443	0.543	0.498
Years of Education	12.660	4.055	12.280	4.462
Family Size	3.026	1.242	2.592	1.287
Sample Size	9776		4835	
Working Sample				
Log Wage	2.568	0.461	2.796	0.462
Age	39.038	9.956	38.913	10.177
Years of Education	14.781	2.811	14.356	3.260
Divorced	0.096	0.295	0.061	0.239
Married	0.717	0.450	0.713	0.452
Family Size	2.962	1.232	2.989	1.289
Expected Experience	18.269	10.635	18.563	10.809
Union Coverage	0.363	0.481	0.381	0.486
Tenure	7.174	6.687	7.990	7.296
Part Time	0.214	0.410	0.036	0.187
Hours Worked	33.994	9.292	40.728	8.327
Newfoundland	0.034	0.181	0.034	0.181
Prince Edward Island	0.026	0.158	0.021	0.143
Nova Scotia	0.063	0.244	0.058	0.233

[7] This variable is omitted as the repressor of all subsequent equations because it produces incredibly inaccurate measurements.

New Brunswick	0.056	0.230	0.055	0.228
Ontario	0.312	0.464	0.319	0.466
Québec	0.182	0.386	0.198	0.399
Manitoba	0.077	0.267	0.073	0.260
Saskatchewan	0.072	0.259	0.065	0.247
Alberta	0.085	0.279	0.086	0.281
British Columbia	0.092	0.289	0.091	0.288
Less than High school	0.125	0.331	0.187	0.390
High school	0.301	0.459	0.287	0.452
Undergraduate	0.524	0.499	0.470	0.499
Graduate School	0.050	0.218	0.055	0.229
Agricultural Sector	0.010	0.100	0.016	0.127
Primary Sector	0.006	0.080	0.049	0.216
Utility, Construction and Manufacturing Sector	0.118	0.323	0.358	0.480
Wholesale, Retail, and Transportation Sector	0.175	0.380	0.213	0.410
Food Services	0.076	0.265	0.033	0.179
Other Services	0.548	0.498	0.261	0.439
Public Administration	0.066	0.248	0.069	0.253
Sample Size	21267		22647	

Table 2b: Summary Statistics for the 2011 wave

	Women		Men	
	Mean	S.D.	Mean	S.D.
Whole Sample				
Age	41.365	11.103	40.798	11.303
Less Than High school	0.108	0.311	0.143	0.350
High school	0.279	0.449	0.294	0.456
Undergraduate	0.554	0.497	0.508	0.500
Graduate School	0.059	0.235	0.055	0.229
Divorced	0.101	0.302	0.074	0.261
Married	0.520	0.500	0.475	0.499
Years of Education	14.974	2.824	14.656	2.971
Family Size	2.880	1.241	2.782	1.300
Sample Size	32650		29171	
Non-Working Sample				
Age	42.880	11.676	43.889	12.077
Less Than High school	0.227	0.419	0.276	0.447
High school	0.330	0.470	0.320	0.467
Undergraduate	0.407	0.491	0.374	0.484
Graduate School	0.036	0.187	0.029	0.169
Divorced	0.099	0.299	0.111	0.314
Married	0.535	0.499	0.342	0.475
Years of Education	13.803	3.493	13.401	3.652
Family Size	2.950	1.282	2.488	1.288
Sample Size	9196		5765	
Working Sample				
Log Wage	2.980	0.461	3.122	0.462
Age	40.831	10.845	40.163	11.031
Years of Education	15.386	2.415	14.914	2.741
Divorced	0.102	0.303	0.066	0.248
Married	0.515	0.500	0.502	0.500
Family Size	2.856	1.225	2.842	1.294
Expected Experience	19.464	11.303	19.256	11.356
Union Coverage	0.371	0.483	0.330	0.470
Tenure	7.519	6.783	7.342	6.936
Part Time	0.178	0.383	0.044	0.204
Hours Worked	34.723	8.662	40.255	8.484
Newfoundland	0.036	0.186	0.034	0.182
Prince Edward Island	0.027	0.162	0.023	0.151
Nova Scotia	0.052	0.223	0.047	0.212
New Brunswick	0.051	0.220	0.051	0.220
Ontario	0.291	0.454	0.282	0.450

Québec	0.167	0.373	0.172	0.378
Manitoba	0.090	0.286	0.091	0.288
Saskatchewan	0.069	0.254	0.070	0.255
Alberta	0.105	0.307	0.119	0.323
British Columbia	0.112	0.315	0.110	0.312
Less Than High school	0.067	0.249	0.115	0.319
High school	0.261	0.439	0.289	0.453
Undergraduate	0.605	0.489	0.535	0.499
Graduate School	0.067	0.250	0.061	0.239
Agricultural Sector	0.008	0.086	0.013	0.115
Primary Sector	0.010	0.097	0.056	0.230
Utility, Construction and Manufacturing Sector	0.085	0.280	0.315	0.465
Wholesale, Retail, and Transportation Sector	0.170	0.375	0.215	0.411
Food Services	0.071	0.257	0.037	0.189
Other Services	0.581	0.493	0.291	0.454
Public Administration	0.076	0.264	0.072	0.259
Sample Size	23454		23406	

Table 3: Coefficients of the Inverse Mills With Different Specifications

	1999			
	Female		Male	
Specifications	Coefficient	S.E.	Coefficient	S.E.
Age and Experience	-0.045	0.045	-0.386***[8]	0.030
Age	-0.231***	0.044	-0.427***	0.046
Experience	-0.192***	0.046	-0.433***	0.029
	2011			
	Female		Male	
	Coefficient	S.E.	Coefficient	S.E.
Age and Experience	0.151**	0.06	-0.296***	0.034
Age	-0.048	0.056	-0.320***	0.034
Experience	-0.009	0.061	-0.294***	0.033

[8] *=significant on a 90% confidence level, **=significant on a 95% confidence level, and ***=significant on a 99% confidence level

Table 4a: Log Wage Equations for the 1999 Wave

	Women			
	OLS		Heckman	
	Coefficient	S.E.	Coefficient	S.E.
Constant	0.674***	0.094	0.789***	0.151
Age	0.082***	0.005	0.077***	0.008
Age Squared	-0.001***	0.000	-0.001***	0.000
Expected Experience	-0.042***	0.003	-0.039***	0.004
Expected Experience Squared	0***	0.000	0***	0.000
Union coverage	0.17***	0.005	0.17***	0.005
Tenure	0.017***	0.000	0.017***	0.000
Part Time	-0.077***	0.010	-0.077***	0.010
Hours Worked	0.000	0.000	0	0.000
Newfoundland	-0.305***	0.013	-0.305***	0.013
Prince Edward Island	-0.217***	0.015	-0.217***	0.015
Nova Scotia	-0.242***	0.010	-0.242***	0.010
New Brunswick	-0.225***	0.011	-0.225***	0.011
Québec	-0.095***	0.007	-0.095***	0.007
Manitoba	-0.161***	0.009	-0.161***	0.009
Saskatchewan	-0.165***	0.010	-0.166***	0.010
Alberta	-0.069***	0.009	-0.069***	0.009
British Columbia	0.049***	0.009	0.049***	0.009
Less than High school	-0.074***	0.016	-0.066***	0.017
High school	-0.062***	0.008	-0.062***	0.008
Graduate School	0.195***	0.014	0.198***	0.014
Agricultural Sector	0.032	0.025	0.034	0.025
Primary Sector	0.443***	0.030	0.443***	0.030
Utility, Construction and Manufacturing Sector	0.237***	0.011	0.237***	0.011
Wholesale, Retail, and Transportation Sector	0.12***	0.010	0.12***	0.010
Other Services	0.302***	0.010	0.302***	0.010
Public Administration	0.374***	0.013	0.375***	0.013
λ	-	-	-0.044	0.045
R^2	.4649			
F	709.7			
Sample Size		21267		

	Men			
	OLS		Heckman	
	Coefficient	S.E.	Coefficient	S.E.
Constant	1.09***	0.080	1.833***	0.104
Age	0.073***	0.004	0.037***	0.005
Age Squared	-0.001***	0.000	0***	0.000
Expected Experience	-0.021***	0.003	-0.007**	0.003
Expected Experience Squared	0**	0.000	0.000	0.000
Union coverage	0.068***	0.005	0.067***	0.005
Tenure	0.014***	0.000	0.014***	0.000
Part Time	-0.274***	0.015	-0.268***	0.015
Hours Worked	-0.004***	0.000	-0.004***	0.000
Newfoundland	-0.29***	0.013	-0.298***	0.013
Prince Edward Island	-0.301***	0.017	-0.303***	0.017
Nova Scotia	-0.265***	0.011	-0.266***	0.011
New Brunswick	-0.228***	0.011	-0.232***	0.011
Québec	-0.114***	0.007	-0.115***	0.007
Manitoba	-0.116***	0.010	-0.114***	0.010
Saskatchewan	-0.124***	0.010	-0.124***	0.010
Alberta	0.001	0.009	0.001	0.009
British Columbia	0.04***	0.009	0.042***	0.009
Less than High school	-0.139***	0.014	-0.119***	0.015
High school	-0.092***	0.008	-0.105***	0.008
Graduate School	0.192***	0.013	0.21***	0.014
Agricultural Sector	0.109***	0.023	0.103***	0.022
Primary Sector	0.476***	0.017	0.463***	0.017
Utility, Construction and Manufacturing Sector	0.367***	0.014	0.355***	0.013
Wholesale, Retail, and Transportation Sector	0.253***	0.014	0.242***	0.014
Other Services	0.327***	0.014	0.319***	0.014
Public Administration	0.404***	0.016	0.394***	0.016
λ	-	-	-0.386***	0.030
R^2	.4197			
F	629.13			
Sample Size		22647		

Table 4b: Log Wage Equations for the 2011 Wave

	Women			
	OLS		Heckman	
	Coefficient	S.E.	Coefficient	S.E.
Constant	0.916***	0.106	0.6***	0.165
Age	0.081***	0.006	0.096***	0.008
Age Squared	0***	0.000	-0.001***	0.000
Expected Experience	-0.05***	0.004	-0.058***	0.005
Expected Experience Squared	0.000	0.000	0.000	0.000
Union coverage	0.161***	0.006	0.16***	0.006
Tenure	0.015***	0.000	0.015***	0.000
Part Time	-0.093***	0.011	-0.093***	0.011
Hours Worked	0.001	0.000	0.001	0.000
Newfoundland	-0.113***	0.014	-0.112***	0.014
Prince Edward Island	-0.121***	0.015	-0.121***	0.015
Nova Scotia	-0.141***	0.011	-0.14***	0.012
New Brunswick	-0.159***	0.012	-0.158***	0.012
Québec	-0.071***	0.007	-0.07***	0.007
Manitoba	-0.076***	0.009	-0.075***	0.009
Saskatchewan	0.027***	0.010	0.028***	0.010
Alberta	0.097***	0.009	0.098***	0.009
British Columbia	0.003	0.009	0.004	0.009
Less Than High school	0.054***	0.020	0.011	0.027
High school	-0.019*	0.010	-0.027***	0.011
Graduate School	0.077***	0.014	0.063***	0.016
Agricultural Sector	0.083***	0.029	0.079***	0.029
Primary Sector	0.471***	0.027	0.47***	0.027
Utility, Construction and Manufacturing Sector	0.232***	0.012	0.231***	0.012
Wholesale, Retail, and Transportation Sector	0.084***	0.011	0.083***	0.011
Other Services	0.289***	0.010	0.289***	0.010
Public Administration	0.39***	0.013	0.39***	0.013
λ	-	-	0.151**	0.060
R^2	.3813			
F	532.24			
Sample Size		22483		

	Men			
	OLS		Heckman	
	Coefficient	S.E.	Coefficient	S.E.
Constant	1.597***	0.094	1.994***	0.108
Age	0.055***	0.005	0.037***	0.006
Age Squared	0***	0.000	0***	0.000
Expected Experience	-0.016***	0.003	-0.01***	0.004
Expected Experience Squared	0*	0.000	0*	0.000
Union coverage	0.076***	0.006	0.077***	0.006
Tenure	0.012***	0.000	0.012***	0.000
Part Time	-0.238***	0.015	-0.236***	0.015
Hours Worked	-0.001***	0.000	-0.001***	0.000
Newfoundland	-0.079***	0.015	-0.079***	0.015
Prince Edward Island	-0.229***	0.017	-0.228***	0.017
Nova Scotia	-0.153***	0.013	-0.151***	0.013
New Brunswick	-0.181***	0.012	-0.18***	0.012
Québec	-0.089***	0.008	-0.077***	0.008
Manitoba	-0.091***	0.010	-0.09***	0.010
Saskatchewan	0.038***	0.011	0.039***	0.011
Alberta	0.168***	0.009	0.17***	0.009
British Columbia	0.048***	0.009	0.051***	0.009
Less Than High school	-0.11***	0.018	-0.071***	0.019
High school	-0.093***	0.009	-0.084***	0.010
Graduate School	0.118***	0.014	0.123***	0.015
Agricultural Sector	0.099***	0.026	0.093***	0.025
Primary Sector	0.515***	0.017	0.51***	0.017
Utility, Construction and Manufacturing Sector	0.359***	0.014	0.353***	0.014
Wholesale, Retail, and Transportation Sector	0.219***	0.014	0.215***	0.014
Other Services	0.343***	0.014	0.339***	0.014
Public Administration	0.476***	0.017	0.471***	0.017
λ	-	-	-0.296***	0.034
R^2	.3186			
F	408.86			
Sample Size		23406		

Table 5a: Selection Equations for the 1999 wave

	Women		Men	
	Coefficient	S.E.	Coefficient	S.E.
Constant	-3.562***	0.257	-2.814***	0.260
Age	0.262***	0.014	0.223***	0.015
Age Squared	-0.002***	0.000	-0.002***	0.000
Expected Experience	-0.117***	0.009	-0.108***	0.009
Expected Experience Squared	0.001***	0.000	0.001***	0.000
Less than High school	-0.356***	0.042	-0.163***	0.046
High school	-0.053**	0.024	0.069**	0.028
Graduate School	-0.050	0.050	-0.032	0.054
Divorced	0.051	0.033	0.213***	0.039
Married	0.046**	0.023	0.572***	0.027
Family Size	-0.092***	0.007	0.02**	0.008

Table 5b: Selection Equations for the 2011 wave

	Women		Men	
	Coefficient	S.E.	Coefficient	S.E.
Constant	-2.83***	0.285	-1.884***	0.288
Age	0.221***	0.016	0.175***	0.016
Age Squared	-0.002***	0.000	-0.002***	0.000
Expected Experience	-0.104***	0.010	-0.074***	0.010
Expected Experience Squared	0.001***	0.000	0.001***	0.000
Less Than High school	-0.543***	0.052	-0.281***	0.053
High school	-0.16***	0.027	-0.077***	0.029
Graduate School	-0.138***	0.046	-0.011	0.055
Divorced	-0.007	0.030	0.085**	0.036
Married	-0.022	0.020	0.472***	0.024
Family Size	-0.073***	0.007	0.02**	0.008

Table 6a: Log Wage Differential for the 1999 Wave Based on OLS Estimation

	Explained	Unexplained	% contributed to wage-gap
Constant	-	0.416	0.00%
Age	-0.009	-0.339	-4.04%
Age Squared	0.003	-0.011	1.38%
Expected Experience	-0.006	0.372	-2.75%
Expected Experience Squared	0.001	-0.051	0.44%
Union coverage	0.001	-0.037	0.55%
Tenure	0.012	-0.021	5.08%
Part Time	0.049	-0.042	21.41%
Hours Worked	-0.027	-0.148	-11.83%
Less than High school	-0.009	-0.008	-3.78%
High school	0.001	-0.009	0.55%
Graduate School	0.001	0.000	0.46%
Agricultural Sector	0.001	0.001	0.30%
Primary Sector	0.020	0.000	8.84%
Utility, Construction and Manufacturing Sector	0.088	0.015	38.68%
Wholesale, Retail, and Transportation Sector	0.010	0.023	4.28%
Other Services	-0.094	0.013	-41.15%
Public Administration	0.001	0.002	0.50%
Province Fixed Effects	0.003	0.005	1.19%
Total	0.046	0.182	20.12%
Wage-Gap	0.228		100.00%

Table 6b: Log Wage Differential for the 2011 Wave Based on OLS Estimation

	Explained	Unexplained	% contributed to wage-gap
Constant	-	0.681	0.00%
Age	-0.037	-1.074	-25.90%
Age Squared	0.019	0.004	13.59%
Expected Experience	0.003	0.659	2.35%
Expected Experience Squared	0.000	-0.050	0.35%
Union coverage	-0.003	-0.031	-2.23%
Tenure	-0.002	-0.025	-1.49%
Part Time	0.032	-0.026	22.48%
Hours Worked	-0.006	-0.064	-4.31%
Less Than High school	-0.005	-0.011	-3.75%
High school	-0.003	-0.019	-1.83%
Graduate School	-0.001	0.003	-0.52%
Agricultural Sector	0.001	0.000	0.41%
Primary Sector	0.024	0.000	16.81%
Utility, Construction and Manufacturing Sector	0.082	0.011	58.00%
Wholesale, Retail, and Transportation Sector	0.010	0.023	7.06%
Other Services	-0.100	0.031	-70.05%
Public Administration	-0.002	0.007	-1.21%
Province Fixed Effects	0.003	0.006	2.35%
Total	0.017	0.125	12.11%
Wage-Gap	0.142		100.00%

Table 7a: Log Wage Differential for the 1999 Wave Based on the Heckman Procedure

	Explained	Unexplained	% contributed to wage-gap
Constant	-	1.159	0.00%
Age	-0.005	-1.756	-1.39%
Age Squared	0.001	0.555	0.38%
Expected Experience	-0.002	0.639	-0.59%
Expected Experience Squared	0.000	-0.089	-0.07%
Union coverage	0.001	-0.037	0.37%
Tenure	0.011	-0.023	3.38%
Part Time	0.048	-0.041	14.20%
Hours Worked	-0.028	-0.154	-8.40%
Less than High school	-0.007	-0.006	-2.19%
High school	0.001	-0.013	0.43%
Graduate School	0.001	0.001	0.34%
Agricultural Sector	0.001	0.001	0.19%
Primary Sector	0.020	0.000	5.84%
Utility, Construction and Manufacturing Sector	0.085	0.014	25.40%
Wholesale, Retail, and Transportation Sector	0.009	0.021	2.79%
Other Services	-0.092	0.009	-27.31%
Public Administration	0.001	0.001	0.33%
Province Fixed Effects	0.003	0.004	0.81%
Total	0.049	0.287	14.53%
Wage-Gap	0.336		100.00%
Size of Selection Bias	0.108		

Table 7b: Log Wage Differential for the 2011 Wave Based on the Heckman Procedure

	Explained	Unexplained	% contributed to wage-gap
Constant	-	1.078	0.00%
Age	-0.024	-1.828	
Age Squared	0.010	0.319	
Expected Experience	0.002	0.780	0.91%
Expected Experience Squared	0.001	-0.051	0.23%
Union coverage	-0.003	-0.031	-1.43%
Tenure	-0.002	-0.027	-0.93%
Part Time	0.032	-0.025	14.12%
Hours Worked	-0.006	-0.066	-2.88%
Less Than High school	-0.003	-0.008	-1.53%
High school	-0.002	-0.017	-1.05%
Graduate School	-0.001	0.003	-0.34%
Agricultural Sector	0.001	0.000	0.24%
Primary Sector	0.024	0.000	10.55%
Utility, Construction and Manufacturing Sector	0.081	0.010	36.23%
Wholesale, Retail, and Transportation Sector	0.010	0.022	4.40%
Other Services	-0.099	0.029	-43.93%
Public Administration	-0.002	0.006	-0.76%
Province Fixed Effects	0.003	0.008	1.53%
Total	0.020	0.204	9.12%
Wage-Gap	0.224		100.00%
Size of Selection Bias	0.082		

Distributions of Log Wage

Figure 1a

Log Wage 1999

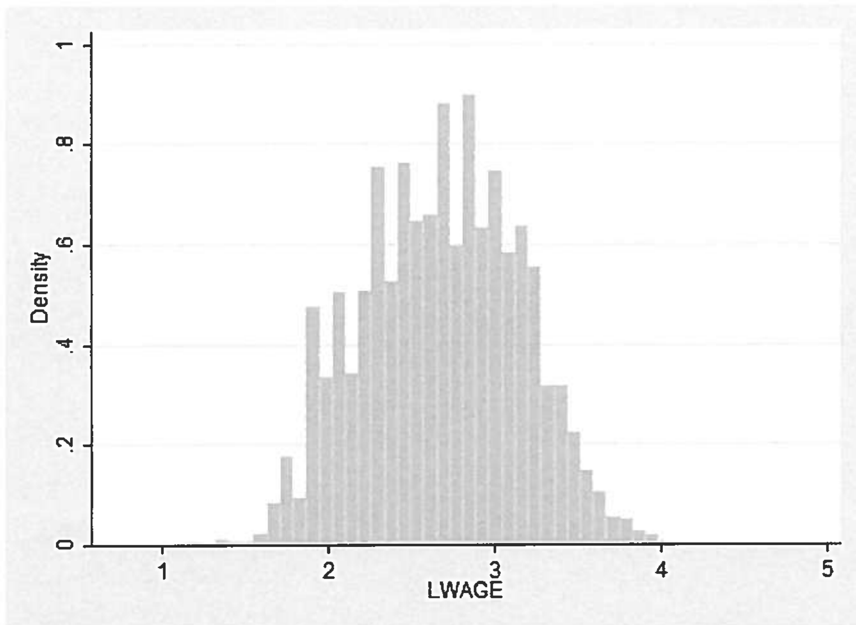

Figure 1b

Log Wage 2011