Education and the Evolution of the Gender Income Gap in Canada

By Fan Zhang

Student number: 3730332

Major paper presented to the
Department of Economics of the University of Ottawa
in partial fulfillment of the requirements of the M.A. Degree

Supervisor: Professor Kathleen M. Day

ECO 7997

Ottawa, Ontario

April 2009

Education and the Evolution of the Gender Income Gap in Canada

Abstract

In my major paper, I will follow the same structure and techniques as Christie and Shannon (2001) to examine the evolution of the gender income gap between 1990 and 2000 using sample data selected from the Canadian censuses of 1991 and 2001. I will also look at the composition of the unexplained and explained components, which can be related to educational attainment, field of study and other major characteristics, using both the Oaxaca (1973) and Cotton (1988) decompositions.

I. Introduction

As a critical economic standard of society nowadays, the gender income gap between men and women is a regular subject of discussion among scholars. Consequently, large numbers of studies have been conducted on this topic. Nowadays, the gender income gap has become a vital issue for policy makers and politicians, leading to the following questions: Why is there an income gap between genders? What effect will this gap have on society? How can society reduce and eliminate the gap to fight gender discrimination?

Since human capital theory was first developed, it has become a popular theoretical foundation for studies of the gender income gap. However, there is still controversy about which variables best explain the income gap. Education is one of the contributing variables. Generally speaking, people with higher education would earn higher income in most cases in the labour market. Thus if average educational levels and returns to education for women and men are quite different, education may be a powerful tool for explaining the gender income gap. Indeed, it has been used as a key determinant of worker's income in recent research (e.g., Christie and Shannon 2001, Finnie and Wannell 2004). Most of the past research and analysis shows that there is a strong correlation between education and income.

This paper has two primary objectives. The first one is to compare the differentials in earnings between males and females in Canada in 1990 and 2000 using Census data in order to examine the evolution of the income gap. The second is to measure the contribution of education towards changes in the gender income gap. Basically, my paper is inspired by Christie and Shannon (2001). They used Canadian Census data for 1986 and 1991 to measure how the gender income gap had varied between 1985 and 1990. In addition, they used the decomposition techniques of Oaxaca (1973) and Cotton (1988) to estimate the contribution of explained and unexplained factors to the gender income gap, in particular the share of education in the explained component of the gender income gap. In my major paper, I will follow the same structure and techniques as Christie and Shannon (2001) to examine the evolution of the gender income gap between 1990 and 2000 using sample data selected from the Canadian censuses of 1991 and 2001. I will also look at the composition of the unexplained and explained components, which can be related to educational attainment, field of study and other major characteristics, using both the Oaxaca (1973) and Cotton (1988) decompositions.

The structure of my paper is as follows. Section II will review related research on the evolution of the gender earnings gap. Section III will briefly discuss the data and sample selection used in my analysis based on data from the 1991 and 2001 Canadian censuses. In section IV, I will examine at the aggregate level the correlation between educational attainment, field of study and male-female earnings differentials. Section V, I will describe the specification and explanatory variables of the regression model based on human capital theory. Section VI presents the econometric analysis; the majority of this section will focus on the discussion of the OLS estimation results. In this section, I will examine the development of gender income gap between 1991 and 2001 and show how education contributes into this change. Section VII will use two decomposition techniques (Oaxaca 1973 and Cotton 1988) to illuminate the explained and unexplained portion of the gap. From the analysis we can determine how strong the relationship between education and the gender income gap is. In the final section, I will draw some conclusions from my empirical analysis.

II. Previous Literature on the Gender Income Gap

Over the past decades, many researchers have made tremendous contributions to the analysis of the gender income gap. The traditional view is that education plays an important role in the determination of labour income; it seems straightforward that education is a critical human capital investment that enables labour to acquire higher skills in order to earn higher pay. Thus if there are male-female differences in average educational attainment or field of study, these might explain a large part of the gender income gap.

The basic structure of most studies of the gender income gap is similar: a human capital earnings equation is used to analyze the correlations between the explanatory variables and labour income for women and men separately, and the total gender income gap is then decomposed into explained and unexplained components using either the Oaxaca (1973) decomposition or the Cotton (1988) decomposition, or both. The estimating equation can be expressed as follows: [1]

$$\ln W_{is} = X'_{is}\beta_s + \varepsilon_{is} \quad , \qquad (1)$$

where W_{is} is the earnings of individual i in group s, where s can be m for men or f for women. X_{is} is a vector of explanatory variables, while β_s is the vector of unknown coefficients. Finally, ε_{is} is a random error term. The earnings equations are typically estimated using OLS. Mincer's proxy for work experience, which is age-years of schooling-6 or age are the two most popular ways to represent work experience, due to the difficulty of obtaining actual work experience in some databases. Experience squared is usually included as well as

[1] This human capital earnings equation should be estimated for men and women separately.

experience to allow the rate of return to experience to decrease as the individual ages. Labour income is measured as annual income or the hourly wage depending on which database the author used.

Traditionally, the natural log of earnings is used as the dependent variable, while the independent variables most commonly included are education, work experience, personal characteristics and region. In some recent articles industry and occupation are also included in the equation, in order to increase the explanatory power of the regression model.[2]

Once the parameter vectors β^m and β^f have been estimated using OLS, they can be used to decompose the difference in mean log earnings. In trying to estimate the contribution of discrimination against women to the gender income gap in the US using the 1967 Survey of Economic Opportunity (SEO) database Oaxaca (1973) divided the difference between female and male hourly wages measured in natural logs into an explained part and an unexplained part. The specification of this decomposition is as follow:

$$\ln W^m - \ln W^f = (X^m - X^f)b^m + X^f(b^m - b^f), \qquad (2)$$

where $\ln W^m$ and $\ln W^f$ are the means of the logs of earnings of men and women respectively. X^m and X^f are row vectors containing the averages of the productivity variables for males and females. b^m and b^f are vectors of OLS estimates of the coefficients for males and females respectively, coefficients which reflect the returns to productivity characteristics. The first term of the right hand side of the equation is the explained component, while the second is the unexplained component. The explained component describes the income gap due to differences in the average related characteristics

[2] See for example, Drolet (2001) and Baker Desaulniers, Grant (1995)

of males and females, weighted using the estimated returns to males, while the unexplained component measures the gap due to differences in returns between males and females, weighted using the characteristics of women.

Cotton (1988) criticized Oaxaca's (1973) technique on the grounds that the unexplained term does not measure discrimination very well, so that the unexplained component may under or overstate the extent of discrimination. The basic logic of Cotton's decomposition is similar to that of Oaxaca's, except that Cotton (1988) reformulated Oaxaca's decomposition as

$$\ln W^m - \ln W^f = (X^m - X^f)\, b^c + X^m(b^m - b^c) + X^f(b^c - b^f) \quad (3)$$

where b^c is the weighted average of the coefficient vectors for men and women, and the definitions of all other terms are unchanged. The weights used to calculate b^c are the shares of men and women in the total population. The unexplained component is divided into men's advantage in returns relative to women (the second term) and women's disadvantage in returns relative to men (the third term). Men have an advantage when their coefficients lie above the weighted average coefficients, while women have a disadvantage when their returns lie below the weighted average return. In his opinion, the veracity of estimation of the components of the gender income gap is improved by this procedure. To test it, he used the 1% public use sample of the 1980 US Census to decompose the racial hourly wage gap (in logs) in the U.S. The results suggested that the Oaxaca decomposition overstated the unexplained component, which in this case contained racial discrimination in the US.

Although the structure of most papers on the gender income gap is similar, the use of slightly different measures of income or work experience or different model specifications has led to controversy in the analysis of the gender income gap. Nonetheless all these studies are beneficial as a background for my paper. In the following paragraphs, I will review a number of papers on the gender pay gap in chronological order to provide an overview of a generation's worth of studies focused on the analysis of the gender income gap. Since my paper focuses on the Canadian gender income gap, most of the studies I review use Canadian data.

Turning now to applications of these methods to Canadian data, Holmes (1976) carried out an empirical analysis of the gender income gap in Canada using data from the 1967 Survey of Consumer Finance (SCF), but without using regression analysis or either the Oaxaca or Cotton decompositions. He found that the ratio of female to male annual income increases with the level of education. The ratio varied from 0.35 to 0.40 for those with less than a high school diploma but rose to 0.45 to 0.50 for higher educational levels in 1967. This means that women with higher educational attainment were more competitive in terms of earning income in the labour market at that time. However, he included women who worked part-time in his initial sample, which can make the differentials between females and males larger, because women are more likely to work part-time. When only full-time workers were included in the sample, the gender annual income ratio increased to 0.45 to 0.55 for the lower education level and 0.60 to 0.65 for the more highly educated.

Gunderson (1979) estimated separate human capital earnings equations for males and females, and then decomposed the gender income gap into two parts using Oaxaca's (1973) decomposition. The dependent variable in the human capital equation is the natural log of annual labour income. The data were obtained from the public use tapes of the 1971

Canadian Census. The conclusions of his research were that the female to male income ratio was around 60%, and that over 60 percent of this gender income gap could be attributed to wage discrimination, which is the unexplained component.

Using a similar database, Miller (1987) used data from the 1981 Canadian census to provide a newly updated empirical analysis of the gender income gap in hourly wages. The innovation in his work as compared to the previous papers is the addition of family characteristics such as the numbers and ages of children and years of marriage, besides the classic productivity related characteristics. Miller arrived at two conclusions based on his results: firstly, compared with the outcome of Gunderson (1979), the gender income gap had shrunk between 1970 and 1980. Secondly, the explained component of the gap declined over this period. Lastly, the empirical results derived from the human capital equations with family characteristics included were similar to those derived from the human capital equations without family characteristics.

During the same year, Shapiro and Stelcner (1987) also published an analysis of the gender income gap, and in order to illustrate the evolution of differentials in income between females and males from 1970 to 1980, they compared their results with those of Gunderson (1979). In order to estimate the gender income gap in 1980, data from the 1981 Canadian Census were used, while the measure of income was annual income for women and men. They found that the annual gender income gap in natural logs shrank by 0.089 over this decade; using the Oaxaca decomposition, they concluded that around 47% of this reduction was attributable to the explained component, while the rest was due to a decline in wage discrimination. They attributed the small change in the gap to the strong advantage enjoyed by men in returns to productive characteristics. For example, women are usually concentrated in lower-wage occupations as compared to the higher-wage, male-dominated occupations.

In their analysis of the evolution of the gender income gap, Baker, Desaulniers, and Grant (1995) used data from the Canadian Censuses of 1971, 1981 and 1986, as well as the individual files of the Survey of Consumer Finance (SCF), to estimate the income differentials between men and women during this period. They focused on individuals aged 16 to 64 who worked full time and were not self-employed. Their measure of the gender income gap is the annual earnings ratio of females to males. In both data sets, the trend in the evolution of this gap is similar, with the mean of the income ratio rising from 0.60 to 0.66 from 1970 to 1985 in the Census data, and from 0.64 to 0.65 in the SCF data. They concluded that the change in the gap was driven by two factors. First, there was a change in the gender means of productivity related characteristics. Second, the differential in returns of workers who have same levels of characteristics varied over time.

Another primary objective of Baker, Desaulniers, and Grant (1995) was to explore the results of a decomposition of the mean earnings differential as measured by the difference in mean raw log earnings using the Oaxaca (1973) method. They tried to use three specifications that include different explanatory variables to examine the changes of explained component. The specification 2 has all explanatory variables plus controls for marriage status, children and region, while the specification 3 plus controls for occupation based on the explanatory variables of specification 2. Their results show that as more explanatory variables are included in the specification, the contribution of differences in productivity related characteristics increases, while the explanatory power of the specification is improved. In the last step of their paper, Baker, Desaulniers, and Grant divided the sample of Census data into four age groups and included the different educational attainment levels in the equation for each group. These results showed the positive relationship between actual work experience or age and differentials in income, as well as the negative correlation between education and the earnings gap.

In order to overcome the disadvantage of a lack of accurate information on work experience in the Census database, Coish and Hale (1995) attempted to explain the gender income gap using a human capital earnings equation, demographic characteristics, and actual years of work, using data from the 1993 the Survey of Labour and Income Dynamics (SLID). Their human capital equation is pretty standard, including independent variables such as age, region, educational attainment and field of study, as well as other personal characteristics such as marital status, mother tongue, visible minority status and years of work experience. The dependent variable is the natural log of the hourly wage. Based on the Cotton (1988) decomposition, the gender income ratio is 0.78. Only 12% of this gap can be explained by differences in characteristics; the rest remains unexplained.

In addition to Baker, Desaulniers, and Grant's paper, several other important papers made outstanding contributions to the analysis of the gender income gap. Drolet (2001) provided some new evidence on the Canadian gender wage gap through her empirical analysis. In her opinion, the persistence of the gender income gap is a serious problem in society, although there has been a huge improvement in female productivity related characteristics as compared to those of males. For the purposes of her analysis, she collected a sample of individuals from the 1997 SLID. She estimated the human capital equation separately with different measure of work experience such as the Mincer proxy and actual experience and used the Oaxaca decomposition to decompose the gap. Based on her results from estimation of a human capital equation with the hourly wage as the measure of income, she drew several conclusions. Firstly, the constant term in the equation is significantly larger for males than for females, which means there is a large and significant gender income gap that is unrelated to the choice of explanatory variables. Secondly, the signs of the relationships between the log of earnings and the independent variables are consistent cross genders. Thirdly, the choice of measure of work experience can affect the weight of the explained component and the unexplained component. Lastly, the magnitude of returns to characteristics, including work experience and educational attainment, is very different for men and women.

There are also some interesting findings in her analysis of the decomposition of the gender income gap, the most obvious being that the share of the unexplained component of the gender earnings gap is around 50% to 75%, which means the explained component of the gap is relatively small, when industry and occupation are excluded from the equation. On the other hand, differences in returns to field of study contribute about 6% to the explanation of the gender income gap, but the contribution of particular fields of study vary widely.[3] One more thing to mention here is that the explanatory variables for industry and occupation can explain around 9% and 11% of the wage gap respectively.

In another paper, Drolet (2002) used the same data from the SLID and also individual micro data files from the SCF. The basic idea of this paper is to illustrate the importance of

[3] For example, the fact that male students preferred to take engineering and applied science, can explain 17% of the gender income gap. On the other hand, the prevalence of female students taking health science and education program can reduce the gender income gap by 8%.

the measure of income used and the choice of decomposition methods; in other words, the size of the gender income gap is very sensitive to the measures and methodology chosen. In her empirical results, the female to male income ratio varies greatly, from 61.6% to 80.9% depending on the data source, for the same year (1997). In her opinion, a sample consisting solely of full time and full year workers excludes a large proportion of female workers from the analysis, and therefore the validity of estimates of the gender income gap may be affected. She argues that using hourly wage income instead of annual earnings, as well as using actual work experience instead of potential work experience to calculate the gap, will overcome this problem. The results of the Oaxaca decomposition show that 0.109 of the gender hourly wage gap can be explained by differences in explanatory variables.

Christie and Shannon (2001) used micro data collected from the 1986 and 1991 Canadian Censuses to examine the relationship between educational attainment, field of study and the gender wage gap. The sample only included full-year, full-time workers who were aged 25 to 64. A human capital model including traditional characteristics such as educational attainment, field of study, region, age, age squared and other personal characteristics, as well as industry and occupation, is estimated, while the dependent variable is the natural log of annual income. Although an analysis based on sample averages showed that educational attainment is positively related to the income of both females and males, their decomposition analysis indicated that the contribution of educational attainment to the explained component of the gender income gap is very low. The contribution of educational attainment to the explained component is only around -0.002 (0.52%) to -0.003 (0.78%) based on both Oaxaca's and Cotton's decomposition techniques in 1990, but the contribution of field of study ranged from 0.018 (4.65%) to 0.025 (6.46%) depending on the decomposition technique.

Finally, in a recent study Finnie and Wannell (2004) used the National Graduates Survey (NGS) data for 1984/87, 1988/91 and 1992/95 to examine the evolution of the gender income gap among university students who graduated in 1982, 1986 and 1990 from two to five years after graduation. The specification of the regression model is straightforward, as it is based on the traditional human capital model, while the method of decomposition is that of Oaxaca (1973). The conclusions are first, that the overall gender income gap of all cohorts is narrowing, but the reduction is larger for students who graduated two years ago than for students who graduated 5 years ago. In other words, wages seem to increase more slowly for women as work experience accumulates. Second, the explained component of the gender income gap becomes more important relative to the unexplained component for a given cohort over time.

Although the methodology, independent variables or measures of income and work experience used by the studies discussed above differs, Table 1 shows that they have a number of findings in common. First of all, the value of the gender income gap is decreasing over time; second, the magnitude of the explained component of the gender income gap has decreased. The next section of this paper will examine whether these trends are also evident in Census data for the period 1990 to 2000.

III. Data and Sample Selection

The Canadian Census database, developed by Statistics Canada, provides information on the demographic, social and economic characteristics of the Canadian population. In this paper, I use data from the 1991 and 2001 individual microdata files of the Canadian Census. The Census is the most powerful database for the estimation of evolution of gender income gap based on human capital earnings equations; due to its huge sample size, it not only allows me to do a variety of aggregate analyses in many ways, but also increases the statistical precision of my results.[4] From my perspective, another reason for using Census data is that it contains rich detailed information about educational attainment and field of study.

However, the Census has some definite weaknesses as a source of data for the estimation of earnings equations. The most serious limitations of the census are, first of all, that it lacks information on hourly wage rates, which is the most accurate measure of workers' income, so that I have to use annual earnings instead. This might affect the reliability of my empirical analysis. Drolet (2002) used both annual earnings and hourly wage data from the 1997 SLID database to estimate the female-male income ratio; not surprisingly, the ratio of annual earnings is 61.6%, which is much smaller than the hourly wage ratio, 80.4%. The explanation for this difference is that pay ratios based on annual earnings do not accurately control for differences in hours and weeks of work. In order to overcome this problem I could use the confidential version of the SLID, instead of the Census, but applications to use this version of the SLID take too long to process.

The second problem with Census data is the measure of experience. Since there is no information on work experience in the Canadian census, I have to use something else as a

[4] In the Canadian Census Analyser, we can see that the total sample size of the individual file in the 1991 Census is 809,654. For the 2001 Census it is around 801,055.
(http://r1.chass.utoronto.ca.proxy.bib.uottawa.ca/sdaweb/html/canpumf_census_chass.htm)

proxy. The traditional proxy for experience is called "potential experience", which either equals age or the Mincer measure of experience (age-years of schooling-6). In this case, I will choose age to be the proxy and select only full-year, full-time workers. However, including only FYFT workers does not eliminate all problems, since for these workers the number of hours and months varies considerably for females and males. According to Drolet (2002), proxy measures are going to overstate women's actual work experience by ignoring interruptions in labour supply for family reasons. Marshall (1999) also found that almost 40% of working women take more than six months off work after giving birth.

The last disadvantage of Census data is that the Census is missing information about union status, which is also an important determinant of earnings. This idea was suggested by Doiron and Riddell (1994). They found that union status is not exogenous, but endogenous with other explanatory variables in the regression model. In addition, union status can affect the explained component of the gender income gap. In the union sample, wage determination can explain almost 25 percent of gender income gap, while its explained power is lower than 10 percent in the nonunion sample.

However, the SLID also has some drawbacks with respect to estimating the gender income gap. For example, although the SLID is a panel data set, some of the respondents cannot be followed over time due to attrition. Boudarbat, Lemieux and Riddell (2006) note that this attrition bias may affect sample composition in ways that are difficult to take into account. Furthermore, the number of observations in the SLID is smaller than in the Census.

Due to these drawbacks of the SLID, I still use the Canadian Census as the database for my paper. My base sample is a 20% random sample of each of the 1991 and 2001 Census microdata files for individuals, under the restriction that only individuals between the ages of 15 and 65 be included. I exclude all residents of the three Canadian territories. The reason for this is that their population is so small relative to that of the provinces and their labour market may be quite different. Also, individuals for whom some of the necessary data are missing will be excluded from my sample. The analysis also excludes all workers who are not Full Year Full Time (FYFT) in the survey year, which means all workers who worked at least 48 weeks in the last year of survey year. Workers who are self-employed are also excluded, because the basic idea of this paper is to examine the determinant of income received from employers.

For 1991, the final sample consists of 36,869 individuals, of whom 15,126 are females and 21,743 are males. The number of people selected from the 2001 census database is 37,758, which includes 21,366 males and 16,392 females.

IV. Education and Income: A Descriptive Analysis

The primary objective of this paper is to analyze the evolution of the gender income gap and the effect of educational attainment and field of study on that evolution. But before we look at the results of a regression analysis that includes many control variables, it is helpful to examine the overall tendencies in the evolution of the income gap and see if they appear to be correlated with changes in educational attainment. In this section, I present some tables constructed using my sample data that compare educational attainment and field of study, across males and females and time.

As Table 2 shows, in 1991 the three most common levels of educational attainment for both men and women were grades 9 to 13, high school graduation and college with certificate or diploma. These three levels accounted for 50.3% of women and 41.8% of men in 1991. Almost 40% of women in the 1991 sample, but only 37.5% of men, had a high school diploma or less. Clearly, fewer men than women had a low level of educational attainment. Men were much more likely than women to have trades certificate or diploma , college with a trades certificate or diploma , or university with a Ph.D. degree, However, a higher proportion of women than men had achieved the levels of college with college certificate or diploma and university with university or college certificate or diploma.

By 2001, the distributions of educational attainment for women and men had changed somewhat. First of all, the most dramatic change was that the proportion with educational attainment less than or equal to a high school diploma decreased from 39.7% in 1991 to only 30.1% for women, while for men the proportion changed from 37.5% in 1991 to 31.6% in 2001. This means that there was a greater improvement in women's education than in men's during this ten-year period, since proportionally more women than men had a higher level of education.

Besides the large decline in the proportion of women with less than a high school diploma, there is another notable change from 1991 to 2001, which is the increased number of women with educational attainment at the university level. In 2001, the proportion of women with some education at the university level was 35.8%, while in 1991 it was 28.3%.

Table 2 only gives us part of the picture with respect to changes over time in educational attainment. Tables 3 and 4 give us more insight by showing how educational attainment varies with age as well as sex. Each table shows the proportion of men and women at each level of educational attainment for each of four age groups. Table 3 contains the breakdown for the 1991 sample, while table 4 provides the same information for the 2001 sample.

Table 3 shows that in 1991, the share of higher levels of education decreased constantly with age.[5] The proportion of women aged 25-34 who have no more than a high school diploma is 32.2%, while for the 55-65 age group this share is 57.4%. Similarly, for men, the proportion with no more than a high school diploma increases from 34.2% of those aged 25-34 to 51.8% of those aged 55-65. Apparently, over time the improvement in women's educational attainment was larger than that of men; the gap between the youngest and oldest group for women is 25.2 percentage points compared with 17 percentage points for men. Table 4 shows that the distribution of educational attainment across the four age groups was similar in 2001: as age increased, the proportion of women and men with lower levels of educational attainment increased. The difference in proportions between the youngest and oldest age groups was 26.7 percentage points for women and 16 percentage points for men. The difference for women is larger than that in 1991, but only slightly. This means that in 2001, the average level of educational attainment of women still exceeded that of men.

In order to investigate the correlation between the gender income gap and changes in educational attainment, Table 5 compares average gender income ratio at different levels of education in the 1991 and 2001 Censuses.[6] The most notable observation that can be drawn from this table is that the relationship between the gender income ratio and educational attainment is positive in both 1990 and 2000: when the level of education increases, the ratio

[5] The shares of college and university without certificate, diploma or degree/ with university certificate or diploma/ with bachelor or first professional diploma for males and females decrease with age.

[6] The average gender income ratio is calculated by dividing the average annual income of women by the average annual income of men.

increases also, which means the level of educational is negatively related to the gender income gap. For example, in 1990, the ratio of female to male annual incomes is 59.5% for a person whose education level is no more than grade 5, while it is 90.8% for people with a Ph.D. The gender income ratio increased for each level of education from 1990 to 2000, except for university with a bachelor's degree, university with a master's degree, and university with a doctorate.

In addition to educational attainment, there is one more important variable related to education that must be considered: field of study. Table 6 shows that in the 1991 sample, almost 62.1% of women were concentrated in the education, commerce and business administration, secretarial science and nursing domains, while men had very different tastes in fields of study. Over 48% of males had chosen engineering/applied sciences as their field of study. By 2001 the most notable changes in women's fields of study are decreases in the importance of secretarial science and nursing, from 17% to 11.9%, and from 13.2% to 8% respectively. More women chose commerce/management/business administration as their field of post-secondary study, resulting in an increase from 15.8% to 19.8% in the proportion of women in this field.

Tables 7 and 8 show how the choice of field of study varies with age. In 1991, the oldest women in the sample (aged 55-65) were more likely to have chosen the nursing field, while commerce and business administration was more popular among younger women. In other words, women's tastes with respect to field of study have changed dramatically over the years. Meanwhile, the distribution of men's fields of study also differs across age groups; younger men are more likely than older men to have chosen engineering or commerce rather than education or fine arts.

Table 9 gives us a better understanding of the relationship between field of study and average annual income. In 1990, the largest income differential between men and women was in the secretarial science field; here the ratio of female to male income is only 67%. The highest income ratio is nursing and nursing assistance; it is 91%. For men, average earnings were highest in engineering/applied sciences in both 1990 and 2000; for women, earnings were highest in mathematics and physical sciences. From 1990 to 2000, the ratio of female to male incomes increased for all fields of study except commerce, engineering/applied science, "other health professions", and "All other" fields of study.

In summary, during the 10-year period between 1991 and 2001, the speed of improvement in average educational attainment is faster for women than for men, as evidenced by a large decrease in the share of women with a low level of educational attainment. At the same time, the positive relationship between educational attainment and income appears to have increased the female to male income ratio. On the other hand, for field of study, during this period, the tastes of women with respect to field of study have changed, with more young women choosing majors with higher future payoff instead of majors with low payoffs. In the next section, I will introduce the variables to be included in an econometric analysis that will help me to further explore these trends.

V. Explanatory Variables and Model Specification

Since the 1970s, the most popular way to analyze the gender income gap has been to estimate separate earnings equations for females and males. Although there is no universal agreement on the choice of explanatory variables, in my paper I will follow the same specification of Christie and Shannon (2001), because the basic idea of my analysis is to carry out an in-depth review and extension of this paper. The estimating equation is identical to equation (1), which is repeated here for convenience:

$$\ln W_{is} = X'_{is}\beta_s + \varepsilon_{is} \quad , \tag{1}$$

As before, W_{is} is the earnings of individual i in group s (men or women), X_{is} is a vector of explanatory variables, β_s is the vector of unknown coefficients, and ε_{is} is a random error term.

My choice of variables to include in X follows closely that of Christie and Shannon (2001). First, in analyses of the gender income gap based on human capita theory, geographical location may have pretty large effects on the earnings of workers. Coish and Hale (1995) and Christie and Shannon (2001), as well as many other papers, all include geographical variables in their specification. Therefore I include regional dummies as explanatory variables in my regression model. I constructed six dummy variables, called BC, QUE, ALTA, MAN, ONT, and SASK to represent British Columbia, Quebec, Alberta, Manitoba, Ontario and Saskatchewan respectively. The Atlantic region (Newfoundland and Labrador, Prince Edward Island, Nova Scotia and New Brunswick) is excluded in order to avoid perfect multicollinearity among the regional dummy variables.

Next, as discussed earlier, the Census does not include a measure of experience; instead I must use a proxy for experience. I follow the approach of Christie and Shannon (2001) and use age as the proxy for experience. The square of the experience measure is also included.

One advantage of using Canadian Census data is that it includes detailed information on educational attainment. In total, there are fourteen categories. I therefore include thirteen dummy variables, called EDU2 to EDU14, which indicate the highest level of education achieved, from grades 5 to 8 to a Ph.D. As for the provincial dummy variables, in order to eliminate perfect multicollinearity there is one dummy missing from the pool of educational attainment variables; this is EDU1, which represents the level of education "less than grade 5."

However, the variables EDU2 to EDU 14 do not cover all aspects of education. As discussed in section II, Brown and Corcoran (1996) found that differences in college majors account for eight percent to nine percent of the gender income gap, which means that differences in majors can affect the gap greatly. I therefore include a set of dummy variables called MAJ1 to MAJ13 to represent the field of study. However, it should be noted that the categories of fields of study in 1991 and 2001 are different. In 2001 there are twenty categories, while there are only fourteen categories in 1991. Nevertheless, after comparing the definitions of field of study in 1991 and 2001, we can see that categories 5 to 8 in 2001 are equivalent to the category 5 in 1991, while categories 12 to 15 in 2001 are equivalent to category 9 in 1991.[7] Therefore I combined categories 5 to 8 in 2001 into one single variable corresponding to commerce, management and business administration (MAJ 5), and did the same for categories 12 to 15 in 2001 to create a dummy variable (MAJ 9) for engineering and applied science with trade. The field of fine and applied arts (MAJ 2) is the excluded category.

There has always been some disagreement about whether occupation and industry

[7] In 2001 MAJ 5 is Business and commerce, MAJ 6 is Financial management, MAJ 7 is industrial and institutional management, MAJ 8 is Marketing, merchandising, retailing and sales. MAJ 12 is Building technologies, MAJ 13 is Data processing and computer technologies, MAJ 14 is Electronic and electrical technologies. MAJ 15 is other engineering technologies.

should be included in human capital earnings equations as explanatory variables. Drolet (2002) pointed out that the choice of occupation and industry may not be made by the individual employees themselves but by employers to differentiate employees by gender. If so, one should consider including occupation or industry as explanatory variables in human capital equations. Based on the results of her empirical analysis, the explanatory power of models with that include occupation and industry is much greater than that of models without them. Therefore I follow the specification of Christie and Shannon (2001) by including occupation and industry in this paper.

For industries, the categories are the same in both the 1991 and 2001 Censuses. I include fifteen dummy variables for industries, with "other primary industries" treated as the reference industry. Based on the 2001 Census Individuals File User Documentation and Standard Industrial Classification - Establishments (SIC-E) 1980, the "other primary industries" included fishing and trapping industries; logging and forestry industries; mining (including milling), quarrying and oil well industries.[8] For occupations, the available variables for 1991 and 2001 are different. Unfortunately, as Appendix A shows, it is easy to see that it would be difficult to combine occupational categories of 2000 into the same sub-groups in 1990 based on strong different definitions. I can't combine the occupation variables in 2000 into the same categories used in 1990, because the definitions are so different. Rather than aggregating some of the 2001 categories to correspond to those for 1991, I define thirteen occupational dummies for 1991 and twenty-four occupational

---

8

2001 Census Individuals File User Documentation:
http://p8090-prod.library.utoronto.ca.proxy.bib.uottawa.ca/datalib/codebooks/c/cc01/pumfs/ind/documentation_e_april.pdf

Standard Industrial Classification - Establishments (SIC-E) 1980:
http://www.statcan.gc.ca/subjects-sujets/standard-norme/sic-cti/sice-ctie80_menu-eng.htm

dummies for 2001. For both years "senior management" is the excluded occupation.

There are also other important explanatory variables besides those I mentioned above. First of all, marital status is represented by the dummy variable MARRIAGE, which equals one if individuals are married and not separated, and zero if individuals are single, divorced, separated or widowed. Secondly, language skills are measured by two dummy variables, ONELANG and BOTHEF. The variable ONELANG is used to identify individuals who can only speak one of Canada's official languages, while BOTHEF identifies individuals who can speak both of them. Third, some indicator of the presence of children should ideally be included, but unfortunately the Census variable related to the presence of children at home in 1991 is restricted to females only, while the 2001 variable is for both females and males. Therefore it is meaningless for me to include a dummy variable for the presence of children since it is not applicable for males in 1991. Finally, the explanatory variables also include an indicator of immigration status (IMMIGRANT) and an indicator of visible minority status (MINORITY).

The dependent variables in this model are natural log of women's or men's annual labour income in 1990 and 2000. One thing I should mention here is that in the Census, the earnings data are for the reference year, which means the incomes I compare here are for 1990 and 2000. At the same time, the effect of inflation between 1990 and 2000 should be removed from my data before making intertemporal comparisons, so I convert the 1990 incomes to 2000 dollars by multiplying them by the factor 1.2103. [9]

A complete list of variable names and definitions is provided in Appendix A. Table 10

[9] This inflation factor is calculated by the inflation calculator provided by The Bank of Canada, which can be found at http://www.bankofcanada.ca/en/rates/inflation_calc.html .

presents descriptive statistics for all variables in my human capital earnings equation excluding educational variables (which were discussed in the previous section). The mean of log of annual income increased for women, but decreased for men from 1990 to 2000, which means the female to male income ratio has shrunk during this period. The average woman's age is smaller than that of the average man in both 1990 and 2000. In my sample, women were less likely to be married than men in both 1990 and 2000, while they have exactly same probability of being immigrants in both years. Over 75 percent of women and men are concentrated in Ontario, Quebec and British Columbia in both 1990 and 2000. The distributions of language skills of men and women are very similar; the proportion of people who can speak only one official language is around 79 percent for both genders. Among occupational categories, women are highly concentrated in "business and finance", "nurses", "teacher and professor" and "art, culture, recreation and sport," over 55 percent. However, besides "business and finance", men are highly concentrated in "other management", "social science" and "sales", which account for almost 50 percent of men in 1990. By 2000, the taste of women in occupations had apparently changed; over 40 percent of women were concentrated in business and administration, while over 40 percent of men were in "other management", "natural and applied sciences", "other trades occupation" and "operators". With respect to industries, women are concentrated in "health and social services", "educational service", "finance and real estate" and "manufacturing" in 1990. Manufacturing accounts for a large proportion of men--around 24 percent, while the shares of other industries for men are equally distributed. The distribution of industries for both genders in 2000 is very close to the distribution of 1990.

In summary, the industrial and occupational distributions of men and women are quite different, while for other explanatory variables there are no differences for women and men. In the next section, the results of OLS estimation will give us a more accurate analysis of the correlation between the explanatory variables and income for both women and men.

VI. Econometric analysis

In order to isolate the effect of each explanatory variable on the gender income gap, I adopt the traditional log earnings equation and the decomposition techniques of Oaxaca (1973) and Cotton (1988) in this paper. In this section, the OLS estimates of human capital earnings equations for men and women will be discussed. Separate equations are estimated for men and women for both 1990 and 2000 earnings. First, however, I discuss the results of a number of diagnostic tests of the underlying assumptions of the linear regression models estimated.

Because the distributions of t-statistics and F-statistics are both based on the assumption that the random error term is normally distributed, it is important to test whether the random error is normally distributed. I do this using the Jarque-Bera test. In table 11, the value of the Jarque-Bera test statistic is 1458460 for women and 2995623 for men in 1990, and 1072747 for women and 1130000 for men in 2000. The p-values of these statistics are all equal to 0.000, which means I must reject the null hypothesis that the random error term is normally distributed. Generally speaking, since the most popular conservative solution to the problem of a non-normally distributed error term is to use the same t and F tests anyway, so I still use these t and F tests for hypotheses about the coefficients.

Although I delete one dummy variable from each group of variables in order to avoid perfect multicollinearity, it is still possible that near multicollinearity may pose a problem for the interpretation of the results. Therefore I use two tests here to examine whether there is multicollinearity. Firstly, the condition index of the matrix of explanatory variables is 94.046 for women and 92.295 for men in 1991, and 103.31 for women and 98.964 for men in 2001. These numbers show that there exist some strong correlations between the explanatory variables. Secondly, I also used the Auxiliary R^2 s to judge how serious the multicollinearity is. In table 11, the values of the auxiliary R^2 s of many explanatory variables are greater than 0.8, which is consistent will the high value of the condition index.[10] However, this may be because omitted category for some groups of variables has a very low frequency. In Table 12, I select the groups of variables with the most seriously multicollinearity and re-do the multicollinearity tests using a different reference category. One can easily see that after changing the reference category, the multicollinearity disappears.[11]

Heteroskedasticity is also a very common problem in cross-section data, in the presence of heteroskedasticity the OLS estimator is inefficient and the t and F tests are not valid. Therefore it's important to test whether this error presents in my data. Since the tests for normality indicate that the error term in my data is not normally distributed, I use the modified version of the Breusch-Pagan-Godfrey (BPG) test proposed by Koenker (1981) to test for heteroskedasticity. As shown in Table 11, for 1990, the value of the modified BPG statistic is 16402 for women and 21743 for men, while the corresponding p-value is 0.000 for

---

[10] Almost all auxiliary R^2 s are greater than 0.8 for both women and men for the educational attainment variables. In the case of the industry dummy variables, all auxiliary R^2 s are greater than 0.8 for women.

[11] In Table 12, not only the condition index, but also the all auxiliary R^2 s decline dramatically when EDU and IND are excluded instead of EDU1 and IND2.

both. Therefore the null hypothesis should be rejected at any level of significance. Thus there is a heteroskedasticity problem in the equations for women and men estimated using 1991 Census data. Similarly, Table 11 shows that there is also a heteroskedasticity problem in the 2001 Census data for both women and men. In order to correct for heteroskedasticity, I construct all t-values and p-values, as well as all F values, using White's heteroskedasticity-consistent covariance matrix estimator.

In order to test for specification error, I used Ramsey's RESET test and the FRESETL and FRESETS tests of DeBenedictis and Giles (1998). The FRESETS and FRESETL tests are based on using Fourier transforms instead of powers of the predicted values to proxy of the mean of ε .The results of these tests are listed in table 11. For the 1991 sample, the p-values of FRESET (3), for women and FRESETL (1) and all the FRESETS statistics for men are less than 0.05, so I can conclude that at the 5% significance level, the null hypothesis that the specification is correct should be rejected. Using the 2001 Census data, the p-values of all tests for men are smaller than the significance level, so I should reject the null hypothesis as well. These apparent specification errors may be attributed to the omission of some explanatory variables; however, I have already included all the most important variables in this model so it is not obvious that which new variables should be included.

Tables 13 and 14 report the estimated coefficients together with the corresponding t-ratio and p-values for 1990 and 2000 respectively. One point that should be mentioned here is that due to the potential heteroskedasticity problem, all t-values and p-values as well as all F values are constructed using White's heteroskedasticity-consistent covariance matrix.

The R^2 s, adjusted R^2 s and F values of the joint significance tests all provide information about the explanatory power of model. For both 1990 and 2000, all values of R^2s and the adjusted R^2s are around 0.2, which is not big, but is a fairly standard number for cross-section data. The tables also show that the p-values for the F-test of overall significance are all 0.000 for both years and both genders, which means the null hypothesis that the coefficients of all the explanatory variables are zero simultaneously should be rejected immediately at any level of significance. Again, this implies that the model has a certain amount of explanatory power.

In the remainder of this section, I divide my discussion of the estimated coefficients into several subsections based on categories of variables. Before I begin the discussion, all reference categories for groups of explanatory variables are repeated for convenience: "neither English nor French" for Language; "Atlantic region" for Province; "less than grade 5" for Educational Attainment; "fine and applied arts" for Field of study; " senior management" for Occupation; and "other primary industries" for Industry.

Experience

As the most standard variable in human capital equations, experience will be discussed first. AGE and AGESQUARE are used as a proxies for experience. Table 13 shows that in 1990, the p-values of both AGE and AGESQUARE for both women and men are equal to zero, which means that at any reasonable level of significance the null hypothesis that the estimated coefficients of AGE and AGESQUARE are equal to zero should be rejected. For women, the estimated coefficient is 0.0307 for AGE and -0.0003 for AGESQUARE, while the coefficients are 0.0571 for AGE and -0.000581 for AGESQUARE in the case of men.

From an economic point of view, the rate of return to age for women and men is calculated as (coefficient of AGE + coefficient of AGESQUARE times AGE). In order to compare the return to age for women and men, I take 30 as the value of AGE, which means the return to an additional year of age for women is 0.0212 and for men is 0.0397. Therefore, women have a lower return to age than men have, while the gap in returns to age is 0.0185.

Table 14 shows that the p-values of the estimated coefficients of AGE and AGESQUARE are still zero for both genders in 2000; therefore the same conclusion that their effect on earnings is strongly statistically significant can be drawn. The magnitudes of the estimated coefficients for women are 0.0534 and -0.0005, while the estimated coefficients are 0.0587 and -0.0006 for men. Compared with the 1990 estimates, for both men and women the coefficient of AGE has increased from 1990 to 2000, but the coefficient of AGESQUARE has decreased from 1990 to 2000. At age 30, women have a rate of return to age of 0.0372, while men have a rate of return 0.0408. The gap in rates of return is 0.0036, which is smaller than the return gap in 1990. This result means that over this ten-year period, the return to age for women has increased relative to that of men.

Other explanatory variables

This group includes the variables ONELANG, BOTHEF (language dummy variables), MINORITY (visible minority), IMMIGRATION (immigration status) and MARRIAGE (marital status). According to Table 13, only the estimated coefficient of MINORITY is significant at the 1% significance level for both genders. This means that visible minority status has a negative effect on labour income for both genders; however, the effect is larger for men than for women in 1990. The estimated coefficient of MARRIAGE for men, which is 0.1442, is significantly different from zero at the 1% level of significance, but it is not significant even at the 10% level for women. This means that holding other variables constant, married men have 14.42% higher incomes than non-married men, while the marital status does not have an important effect on women's income. The coefficient of BOTHEF for women is 0.1540 and is significantly different from zero at the 10% level, while the coefficient of IMMIGRATION is -0.0272 for men, which means a male immigrant's income is 2.72% lower than that of a native-born male, and women who can speak both English and French earned 15.4% more than those who speak neither English nor French in 1990.

Table 14 shows that the coefficient of MARRIAGE is still significantly different from zero for men in 2000, but it decreases from 0.1442 to 0.1379, which means holding other constant the advantage of married men in income over those non-married men has decreased. In addition, the estimated coefficients of MINORITY and IMMIGRANT for both genders, as well as those of ONELANG and BOTHEF for men, are significantly different from zero. For women, the estimated coefficient of IMMIGRANT is -0.0624, while that of MINORITY decreases from -0.0871 in 1990 to -0.1214 in 2000. This means female immigrants have 6.24% lower income than non-immigrant women, and visible minority women have 12.14% lower income than non-minority women. Furthermore, the negative effect of belonging to a visible minority has increased from 1990 to 2000 for women. For men the estimated coefficient of MINORITY decreases from -0.2043 in 1990 to -0.2401 in 2000, while the coefficient of IMMIGRANT decreases from -0.0272 in 1990 to -0.0628 in 2000. These results imply that the negative effects of immigration status and visible minority status for both genders' labour income are getting worse, while the extent of this effect is greater for men than for women.

Provinces

For the provincial dummy variables, the Atlantic area is the reference region. Table 13 shows that, for women in 1990, the estimated coefficients of BC, QUE, ALTA and ONT are significantly different from zero at the 1% level of significance. These values are 0.1488, 0.0973, 0.1311 and 0.1905 respectively, which implies that holding all else equal, average income is lowest in the Atlantic area and highest in Ontario. However, the estimated coefficients of MAN and SASK are not significantly different from zero. This implies that in 1990, the average income of women in Manitoba and Saskatchewan was not significantly different from that of the Atlantic region. For men in 1990, the estimated coefficients of BC, ALTA and ONT are significantly different from zero at the 1% significance level. The values of the estimated coefficients imply that the highest income region for men is Ontario, followed by British Columbia and Alberta. However, average incomes in Manitoba, Quebec and Saskatchewan are not significantly different from those in the Atlantic area for men. This result is very similar to that of Christie and Shannon (2001), who found that average incomes were highest in BC and ONT in 1990,

Table 14 shows that in 2000 the coefficients of all the provincial dummies are significant at the 1% level in both the male and female equations, with the exception of the coefficients of SASK for women and MAN for men. However, the coefficients of these two variables are significant at the 5% or 10% levels. The region with the highest average income for women is British Columbia, followed by Ontario, Quebec, Alberta, Manitoba and Quebec. For men, Ontario and BC are still the two most profitable provinces in 2000.

Industry

In this group of variables, the reference industry is "other primary industries". Table 13 shows that for women in 1990, the only industry variables with coefficients are significant at the 1% level are IND1, IND3, IND4, IND7 to IND10 and IND12 to IND16, while the coefficient of IND5 is significant at the 5% level. It is easy to see that all the estimated coefficients of these variables are negative, which implies the returns of these industries for women are all lower than the reference industry "other primary industries". The most profitable industry for women is "other primary industries" followed by IND5 (Transportation and storage); its coefficient implies that holding all else constant, women working in transportation earned 13.06% less than those working in "other primary industries". The lowest average income is IND1 (Agriculture) as it has the lowest coefficient (-0.7979). For men, the estimated coefficients of all the industry dummies are significantly different from zero. Men who worked in the federal government earned only 4.13% less than those who worked in senior manager. As for women, all the estimated coefficients are negative. Therefore, the "food/beverage service" industry with the lowest coefficient of -0.7151 is the lowest income industry compared to "other primary industries".

The results also give rise to other interesting findings with respect to differences between industries. For example, the average income earned by both genders in the agriculture industry, where most jobs are blue collar jobs, is much lower than the average income earned in the federal government, where most jobs are white collar jobs. The average income of women in the federal government is 75.74% higher than the average income earned in agriculture, while for men the differential is about 43.21%.

Table 14 shows that for women in 2000, the estimated coefficients of the variables IND1, IND4, IND8, and IND13 to IND16 are significantly different from zero at the 1% significance level, while those of IND7 and IND9 are also significant at at least the 5% or 10% level of significance. This implies that the transportation and storage; communication; business service; federal government; and other government industries; have the same average income compared to other primary industry for women in 2000. Similar to 1990, all estimated coefficients are negative. However, the magnitudes of the coefficients have increased. This means that although "other primary industries" is still the most profitable industry for women, the average income earned from other industries has improved relative to this reference industry. The new result here is that the "food/beverage service" industry with a coefficient of -0.4877 is now the industry with the lowest income, instead of the "agriculture" industry in 1990. For men, all the estimated coefficients are significantly different from zero at the 1% level. With the exception of IND1 (agriculture), IND4 (construction), and IND6 (communication), coefficients of all other industry dummies decrease in 2000 compared to 1990. This result means that over the ten-year period, the returns to agriculture, construction and communication have improved relative to other primary industries, while the returns to other industries have been getting worse. The differential in incomes between the agricultural and federal government industries for women and men is still large in 2000, although it falls to 44.9% for women and 33.1% for men.

Occupation

In this category, "senior manager" was selected to be the reference occupation. Table 13 shows that for women in 1990, almost all the estimated coefficients of dummy variables in this category are significantly different from zero at the 1% level, while the coefficient of OCP2 is also significant at 5%. All the occupational variables have negative coefficients, which means that the "senior manager" occupation was the best type of job in 1990, while OCP14 (other manual workers) with smallest estimated coefficient did the worst job. For men, all the coefficients are significantly different from zero at the 1% significance level. As for women, the "senior manager" occupation is the most profitable occupation for men in 1990, while OCP14 (other manual workers) earned lowest income relative to the reference occupation for men.

Table 14 shows that the number of explanatory variables for occupation has been increased from 14 to 24 in 2000. For women, twenty-two of the estimated coefficients are significantly different from zero at the 1% significance level, while those of OCP7 (registered nurse) and OCP9 (science, government service) are also significant at the 5% and 10% significance levels respectively. "Senior manger" still is the most profitable occupation for women due to the negative sign of all the estimated coefficients, while OCP16 (childcare), OCP22 (trader helper), OCP25 (labours in processing) are worst occupations for females. For men, only the estimated coefficient of OCP7 is non-significant in 2000. Due to the negative sign of all coefficients, the occupation "senior manager" earned the highest income for men in 2000, while OCP17 (occupations in travel), OCP23 (occupations in primary industries) and OCP25 (labours in processing) were the worst occupation in 2000 for men.

Educational attainment

In this group of variables, the reference level of educational attainment is an education below the grade 5 level. Table 13 shows that for women in 1990 the only educational attainment variables with coefficients that are significant at the 1% level are EDU3, EDU4, and EDU9 to EDU14, while the coefficient of EDU8 is significant at the 5% level. For men, the estimated coefficients of all the educational attainment variables are significant at the 1% level. The parameter estimates for the educational attainment variables are pretty consistent with human capital theory, which predicts that the more education a person has, the higher their earned income. Most of the estimated coefficients increase with the level of educational attainment, especially for men, although there are a few exceptions. The highest return is to a doctoral degree, while the lowest one is for less than grade 5 for both women and men due to the positive sign of all the estimated coefficients.

Looking more closely at the estimated coefficients, one can see that they imply that men with a bachelor's degree earned 54.43% more income than men with less than grade 5, while men who graduated from high school earned 32.83% more than the reference group. Thus the incomes of men with a bachelor's degree are on average 21.6% higher than those of men with a high school diploma. However, for women the income difference between those with a bachelor's degree and those with a high school diploma is only 7.32%. Similarly, the income difference between men with a master's degree and men with a high school diploma is on average 24.11%, while the difference for women is 20.56%. In addition, men with a doctoral degree earned 35.29% more income than those men who graduated from high school, while for women the income difference between those with a doctoral degree and those with a high school diploma is 46.87%. These differences in income across educational levels are both economically important and consistent with prior expectations.

For both women and men, the returns to education increase as the level of education increases in most cases. For example, 8-10 additional years of education after high school leading to a doctoral degree raises women's income by on average 46.87%, while the increase for men is 35.29%. These large returns may provide a strong incentive for people to want to spend many years to achieve a higher level of education. It is interesting to note that a doctoral degree appeared to yield higher returns to women than men in 1990.

However, there are some exceptional cases where income does not increase when the educational level rises. On average, the percentage difference in income between women with various levels of college education and women with a high school diploma are 1.82% (EDU 6--without college diploma); -15.98% (EDU 7--with trades certificate); and -11.6% (EDU 8--with college diploma). The corresponding differences for men are -1.73%, -1.78% and 3.46% respectively. These results imply that on average, people with a college education earned almost the same or less income than people with only a high school diploma for both genders. These results are very odd in comparison to those for other levels of education and inconsistent with human capital theory.

The estimated coefficients for men are generally greater than those for women, except for a doctoral degree and for the level "university without certificate, diploma or degree," which implies that the returns to most levels of educational attainment are higher for men than those of women. For example, as mentioned previously, women with a bachelor's degree earn only 7.32% more than those with a high school diploma, while men with a bachelor's degree earn 21.6% more on average than men with a high school diploma. These results imply that a university degree was more attractive to men than to women in 1990. However, women earn higher returns to EDU9 (university: without certificate, diploma or degree) and EDU14 (university: with earned doctorate) than men. Together with the gender differences in average educational attainment I discussed in the previous section, these gender differences in returns to educational attainment may also help to explain the gender income gap. The magnitude of their contribution is discussed in the next section.

Table 14 shows that for women in 2000, the estimated coefficients of EDU8 to EDU14 are significantly different from zero at the 1% significance level, while those of EDU3 to EDU7 are also significant at at least the 10% level of significance. Again, the magnitudes of the coefficients tend to increase with the level of educational attainment. With the exception of the doctorate (EDU14) and "college without trades certificate or diploma" (EDU6), the estimated coefficients appear to have increased when compared to those of 1990 for women. However, the magnitude of the difference in average incomes between high school graduates and those with bachelors' and doctoral degrees doesn't change much. The average income of a woman with a bachelor's degree is 28.03% higher than the income of a woman with a high school diploma. Similarly, the incomes of women with master's and doctoral degrees are 39.45% and 41.74% higher than those of women with a high school diploma. A college education is still a not good choice for women in 2000, due to the very small income difference between the women with college education and women with a high school diploma.

For men, the estimated coefficients of EDU4, EDU6 and EDU9 to EDU14 are highly significant at the 1% level, while that of EDU8 is significantly different from zero at the 5% level and the coefficients of EDU3 and EDU7 are significant at the 10% level. As for women, the estimated coefficients for men progressively increase with the level of educational attainment, except in the cases of EDU7, EDU8, and EDU10. However, in contrast to the findings for women, the significant coefficient estimates for the educational attainment variables decline between 1990 and 2000. Similarly, the economic importance of the difference in average income between high school graduates and those with bachelor's and doctoral degrees doesn't change much. Men with a bachelor's degree earned 15.03% more income than men with a high school diploma. Similarly, men with master's and doctoral degrees earned 23.95% and 30.3% less than men with a high school diploma which is slightly less than the gap in 1990. These results imply that the returns to university degrees declined somewhat from 1990 to 2000 for men.

Field of study

In this category, the "fine and applied arts" field is the reference field of study. Table 13 shows that in 1990, for women the most highly significant estimated coefficients are those of MAJ1, MAJ5, MAJ6 and MAJ8 to MAJ13, although the coefficient of MAJ4 is also significant at the 5% level. The positive sign of all the significant coefficients for the field of study variables implies that the reference major, "fine and applied arts," yields the lowest return to women in 1990, while the MAJ11 (other health profession/science/technology) with a coefficient of 0.3408 is the most profitable major for women, which means the average income of women whose major is other health profession/science/technology exceeds that of those women whose major is fine and applied arts by 34.08%. For men, only the coefficients of MAJ1, MAJ3, MAJ6, MAJ7, MAJ9 and MAJ11 are statistically significant, mostly at the 5% and 10% levels of significance. Among the significant coefficients for men, those of MAJ6 (Secretarial science) and MAJ3 (Humanities and related) imply that on average these fields earned incomes 18.73% and 18.02% lower than the average income of those men who majored in fine and applied arts in 1990, while MAJ11 (other health profession/scedience/technology) is the most profitable for men, yielding an average income 12.08% higher than a major in fine and applied arts.

In summary, the major "other health profession/science/technology" is the most profitable for both genders in 1990, while "fine and applied arts" is the least profitable for women and "secretarial science" and "humanities and related" are the least profitable for men, based on their low average incomes in 1990.

Table 14 shows that for women in 2000, the most highly significant estimated coefficients for field of study variables are those of MAJ10 (nursing and nursing assistant) and MAJ12 (mathematics/physical science). The values of the coefficients decrease to 0.1363 and 0.1478 in 2000, as compared to 0.1697 and 0.2911 in 1990. This implies that the advantages in returns of these two majors have decreased relative to the reference major (fine and applied arts) in 2000, holding all else constant. In other words, the attractiveness of these two majors declined from 1990 to 2000 compared to the reference major. Only two other estimated coefficients, those of MAJ5 (commerce/management) and MAJ9 (engineering/applied science) are significant at the 5% level. The values of these two coefficients have also decreased relative to 1990. Among the fields with significant coefficients, the major "fine and applied arts" still has the lowest return for women, while the major "Mathematics/physical science" is the most profitable major in 2000 instead of "other health profession/science/technology." For men, the estimated coefficients of MAJ 4, MAJ5 and MAJ8 to MAJ13 are significantly different from zero; of these, only the coefficients of MAJ9 and MAJ11 increase compared with 1990. For men, the major "all other" has the lowest return, while the major "other health profession/science/technology" is still the most profitable major in 2000.

The returns to different fields of study have changed for women in 2000. The "mathematics/physical science" field is now the most attractive major for women, while the most profitable major for men is still same as in 1990. In 2000, the estimated returns to many majors are no longer different from those to the reference major; this is true for MAJ1, MAJ3, MAJ4, MAJ6, MAJ7, MAJ8 and MAJ11 for women and MAJ1, MAJ3, MAJ6, and MAJ13 for men. These gender differences in returns to major field of study provide another possible tool to explain the evolution of the gender income gap from 1990 to 2000.

However, for 1990 my estimated coefficients are somewhat different from those of Christie and Shannon (2001). I think that the difference between my results and theirs may be due to sampling variability. Gordon, Osberg, and Phipps (2005) demonstrated that coefficient estimates can vary considerably across different samples of data; they found that almost 50% of the variation in men's estimated wage elasticities can be explained by sampling variability. Thus the differences between my 1990 results and those of Christie and Shannon may well be due to the fact that we have used different samples.

In summary, the regression results show several things: firstly, for both sexes and both years, income is positive related to experience, but yields larger returns to men than to women, although the return in experience for both genders increases from 1990 to 2000. Secondly, for most industries, the return for women appears to have increased from 1900 to 2000, while there was a decline in the return for men. Thirdly, for both genders, income is positive related to educational attainment, as the coefficients of most of the educational attainment variables increase progressively with the level of education. At the same time, the returns to educational attainment have increased for women, but decreased for men from 1990 to 2000. Lastly, the returns to women's major fields of study decreased from 1990 to 2000.

This discussion of the estimated coefficients can only provide a brief overview of the correlations between the explanatory variables and the dependent variables. However, the most important task of this paper is to examine how the gender income gap has evolved and what contributes to it. Thus, in the next section, I will decompose the gender income gap into various components to try to answer these questions.

VII. Decomposition of the Gender Earnings Gap

The statistical technique called decomposition can divide the total gender income gap into an explained component and an unexplained component. The former represents the gap in income due to differences between men and women in productivity-determining characteristics, while the second one measures the income differential arising from gender discrimination in the form of different returns to the same characteristics. Two types of decomposition technique will be used in this paper: that of Oaxaca (1973), and that of Cotton (1988).

Oaxaca (1973) decomposition

Oaxaca (1973) showed how the difference in the mean values of the logs of male and female earnings can be divided into an explained and an unexplained component. As explained in section II, the formula for this decomposition is:

$$\ln W^{m} - \ln W^{f} = (X^{m} - X^{f}) b^{m} + X^{f} (b^{m} - b^{f}),$$

where the variables are defined in section II. As shown in Table 15, the overall gender income gap, measured as the difference between the logs of male and female earnings, was 0.376 in 1990. Only 31% of the gap can be explained by differences in the average characteristics of men and women, while 69% of the gap is due to differences in the returns to productivity-determining characteristics. This finding is close to those of Drolet (2001, 2002), Baker Desaulniers, and Grant (1995) and Christie and Shannon (2001). The decomposition of the explained component of the 1990 gender income gap indicates that the men had higher average levels than women of the explanatory variables in categories such as field of study, family, occupation, industry, age and other personal variables. The negative share of educational attainment in the decomposition of the explained component shows that women have a slightly higher average level of educational attainment than men have.

The decomposition of the unexplained component shows that men had an advantage over women to returns to educational attainment, family, and age, while women enjoyed higher returns than men to occupation, industry, and field of study. The large positive contribution of age, the proxy for work experience, to the unexplained component in 1990 particularly stands out. In addition, the sum of the shares of educational attainment in the explained and unexplained components is 25.5% of the gender income gap, while the sum for field of study is -20.3% of the total gap.

By 2000 the overall gap had shrunk to 0.313, consisting of a slightly larger explained share of 31.9% and a slightly smaller unexplained share of 68.1%. This change implies that over this ten-year period, the average income of women improved relative to men's and wage discrimination decreased. The largest change in the explained component is the change in the sign of the contribution of family factors and the increase in the contribution of industry and age. Industry of employment plays the greatest role in the explained component of the gender income gap in both 1990 and 2000. The contribution of differences in the educational attainment of males and females is not important in 1990; in the 1991 Census they account for -0.5% of the gender income gap, but they are more important in the 2001 Census, where the contribution of educational attainment is -3.2%. The contribution of field of study to the explained component is about 6.24% in 1990 and 4.5% in 2000. These results show that the women's progress in terms of educational attainment and field of study all contribute more explanatory power to the gender income gap from 1991 to 2001. Furthermore, the sum of the shares of educational attainment in the explained and unexplained components decreases to -54.3% from 25.5% of the gender income gap in 2000, while the sum for field of study increased to 14.3% from -20.3% of the total gap.

In comparison to the changes in the explained component, the changes in the composition of the unexplained component are more dramatic. The higher returns to women to field of study, location and other explanatory variables in 1900 disappeared in 2000, changing to an advantage for men. There is a huge decline the contribution of men's advantage over women in the returns to age, from 162% in 1990 to 37.3% in 2000, which means that the importance of gender discrimination in the returns to work experience decreased dramatically during this period. The advantage men enjoyed in returns to educational attainment in 1990 also disappears in 2000, changing from 26% in 1990 to -51.1% in 2000.

Cotton (1988) decomposition

As explained in section II, the formula for Cotton's (1988) decomposition of the gender difference in the mean of log income is:

$$\ln W^{m} - \ln W^{f} = (X^{m} - X^{f}) b^{c} + X^{m}(b^{m} - b^{c}) + X^{f}(b^{c} - b^{f}) .$$

Table 16 shows this decomposition for both 1990 and 2000. According to this decomposition method, the share of the explained component of the gender income gap is only 17% in 1990, compared to 31% using the Oaxaca (1973) decomposition in 1990. The unexplained component due to male advantages in returns is around 54.8%, while the contribution of female disadvantages in returns is 28.5% in 1990. Within each broad category, the outcome of the Cotton decomposition is very similar to that of the Oaxaca decomposition; for example, the decomposition of the explained component of the 1900 gender income gap indicates that the most important category is industry, which accounts for 11.7%, followed by field of study with 5.25%. Men had higher average levels of characteristics than women in categories such as field of study, family, industry, age and other variables, while women had higher average levels in educational attainment.

Looking at the decomposition of the unexplained component of the gender income gap, it's clear that the sign of each category of male's advantage component and female's disadvantage component is the same, as equation (3) implies. As for the Oaxaca decomposition, the share of male's advantage or female's disadvantage in returns to age has the largest weight in gender income gap; which are 95.6% for men's advantage and 66.3% for women's disadvantage in 1990. This implies that men received higher returns to work experience than did women, which accounted for a substantial share of the gender income gap. In addition, men also have an advantage in the returns to educational attainment and family, while women have the advantage in other categories. The decomposition of the male advantage component indicates that the return of educational attainment for men is higher than for women, with a positive share of 15.6%. The sum of the shares of male advantage and female disadvantage in educational attainment is 26.2% of the gender income gap, while the sum of the shares of male advantage and female disadvantage with respect to field of study amounted to -25.4% of the gap in 1990.

By 2000, the share of the explained component of the gender income gap increased to 22.8%, compared to 31.9% using the Oaxaca (1973) decomposition in 1990. The unexplained component due to male advantages in returns decreased to 38.45% from 54.8%, and the ratio of the contribution of female disadvantages in returns increased to 38.71% from 28.5% relative to 1990. The increase in the explained component is due to increases in the shares of location, family, occupation and industry from 1990 to 2000. Industry still plays the most important role, accounting for 20% of the gap. The results also show that educational attainment played an important role in decreasing the gender income gap, based on its share of -4.2% in 2000, compared to -0.8% in 1990.

For the unexplained component, the share of the male advantage in returns apparently decreased, while the extent of the female disadvantage is getting seriously worse.[12] The contributions of several categories change in sign. The decrease in the share of male's advantage in returns can be attributed to decreases in the contributions of educational attainment, family, age, other variables, occupation and industry from 1990 to 2000. In other words, the increase in the share of women's disadvantage in returns is due to increases in the contributions of field of study, location, family, and other variables. The contribution of returns to educational attainment is -21.1% in the form of male advantage and -29.1% in the form of female disadvantage, while the contribution of returns to field of study is 5.07% to male advantage and 5.5% to female disadvantage.

In summary, based on the previous discussion of the decomposition of the gender income gap, several conclusions can be drawn. Firstly, the gender income gap decreased from 0.376 to 0.313 from 1990 to 2000. Secondly, between 1990 and 2000, the share of the explained component of the gender income gap increased consistently, but the two different decomposition methods attribute these changes to slightly different factors. For example, according to the Oaxaca decomposition occupation makes a positive contribution to the explained component, while according to Cotton's decomposition its contribution is negative. However, the contributions of educational attainment and field of study decreased in both cases. This means that during this ten year-period, the improvement in women's average

[12] This is an odd result, in that the increase in women's disadvantage in returns is accompanied by an decrease in men's advantage.

educational levels and choice of majors has reduced the gender income gap progressively. Thirdly, for each decomposition, the unexplained component has declined over the ten years. Discrimination in the returns to explanatory variables in favour of men decreased for educational attainment, family, occupation, industry, and age. For both decompositions, the share of educational attainment in the unexplained component decrease from a positive number to a negative number, which means a decrease in discrimination in the returns to educational attainment has been an important contributing factor to the decrease in the gender income gap, while women have a higher return to educational attainment than men. However, field of study's contribution changed in the opposite direction, leading to an increase in male's advantage in this category. These conflicting results highlight the difficulty of reducing the gender income gap.

VIII. Conclusion

This paper used Canadian Census data for 1991 and 2001 to examine the evolution of differentials in the annual income of women and men during this 10-year period, and the effect of educational attainment and field of study on this evolution. Using the standard Oaxaca (1973) and Cotton (1988) decomposition techniques, I find that firstly, women still suffering form a lower income compared with men, but the gender income gap has been reduced during this period. The gap in log incomes has decreased from 0.376 in 1990 to 0.313 in 2000. Secondly, field of study has some power to explain this reduction in the gender income gap in both 1990 and 2000 its contribution decreases from 6.2%-5.3% in 1990 to 4.5%-3.4% in 2000 depending on the decomposition used. This means that the income gender gap has decreased as more women enter the fields of study in which men have been concentrated. Lastly, the share of educational attainment rose from 1990 to 2000. During this period, the share of educational attainment in the gender income gap has increased from -0.5%- -0.8% in 1991 to -3.2%- -4.2% in 2000, depending on the decomposition. The improvement in the average educational attainment of women has influenced the decline in the gender income gap.

Together my paper and previous studies have one thing in common: in the future, there is no doubt that the gender income gap is going to decrease over time unless something unexpected happens. Educational attainment and field of study will play a more important role in reducing the gender income gap. Policies that not only increase the average educational level of female, but also encourage more women to pursuer the same majors as men will be useful in decreasing the gender income gap. At the same time, policies that reduce the differentials in returns to the characteristics of men and women would also help to reduce the gap.

Although I have devoted all effort to this paper, potential data limitations associated with the Canadian Census may raise questions about the quality of this paper. In future research I would like to use better data (e.g., the SLID) to do the same analysis again and include more important explanatory variables in order to obtain a more accurate as well as consistent and comparable estimate of the gender income gap in Canada.

Reference

Baker, Michael, and Andree Desaulniers, and Mary Grant (1995) "The distribution of the male/female earnings differential, 1970-1990." *Canadian Journal of Economics*, Vol.28, No.3, pp.479-501.

Boudarbat, Brahim, and Lemieux Thomas, and Riddell W. Craig (2006) "Recent Trends in Wage Inequality and the Wage Structure in Canada." *Dimensions of Inequality in Canada*, pp. 273-306.

Christie, Pamela, and Michael. Shannon (2001) "Educational attainment and he gender wage gap: evidence from the 1986 and 1991 Canadian Censuses." *Economics of Education Review*, Vol. 20, No. 2, pp. 165-180.

Coish, David, and Alison Hale (1995) "The wage gap between men and women: an update." *SLID Research Paper Series*, (Statistics Canada, Catalogue no. 75F0002M), Ottawa, pp. 95-114.

Cotton, Jeremiah (1988) "On the decomposition of wage differentials." *Review of Economics and Statistics,* Vol. 70, No. 2 pp. 236-342.

DeBenedictis, L. F., and D. E. A. Giles (1998) "Diagnostic testing in econometrics: variable addition, FRESET, and Fourier approximations." In A. Ullah and D. E. A. Giles, *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, 383-417.

Doiron, Denise J. and W. Craig Riddell (1994) "The impact of unionization on male-female earnings differences in Canada." *Journal of Human Resources,* Vol. 29, No. 2, pp. 504-534.

Drolet, Marie (2001) "The persistent gap: new evidence on the Canadian gender wage gap." Statistics Canada, Analytical Studies Branch, Analytical Studies Branch Research Paper Series.

Drolet, Marie (2002) "The male-female wage gap." *Perspectives on Labour and Income* Statistics Canada catalogue no. 75-001-XPE, Ottawa, pp. 29-37.

Finnie, Ross, and Ted Wannell (2004) "The evolution of the gender earnings gap amongst Canadian university graduates." *Analytical Studies Branch Research Paper Series* Statistics Canada, catalogue no. 11F0019MIE No. 235, Ottawa.

Gordon, Daniel V. and Lars Osberg, and Shelley Phipps (2005) "Sampling variability: some observations from a labour supply equation." *Applied Economics,* Vol. 37, No. 18, pp.

2167-2175.

Gunderson, Morley (1979) "Decomposition of the male/female earnings differential: Canada 1970" *Canadian Journal of Economics,* Vol.12, No.3, pp.479-485.

Holmes, R. A. (1976) "Male-female earnings differentials in Canada." *Journal of Human Resources*, Vol. 11, No. 1, pp. 1-21.

Koenker, R. (1981) "A note on studentizing a test for heteroscedasticity." *Journal of Econometrics*, Vol. 17, No. 1, pp. 107-112.

Marshall, K. (1999) "Employment after childbirth." *Perspectives on Labour and Income* Statistics Canada, Catalogue no. 75-001-XPE No. 3, pp. 18-25.

Miller, Paul W. (1987) "Gender differences in observed and offered wages in Canada, 1980." *Canadian Journal of Economics*, Vol. 20, No. 2, pp. 225-244.

Oaxaca, Ronald (1973) "Male-female wage differentials in urban labor markets." *International Economic Review,* Vol. 14, No. 3 pp. 693-709.

Shapiro, D.M, and M. Stelcner (1987) "The persistence of the male-female earnings gap in Canada, 1970-1980: the impact of equal pay laws and language policies" *Canadian Public Policy*, Vol. 13, No. 4, pp. 462-472.

The inflation calculator of The Back of Canada can be found at: http://www.bankofcanada.ca/en/rates/inflation_calc.html.

2001 Census Individuals File User Documentation: http://p8090-prod.library.utoronto.ca.proxy.bib.uottawa.ca/datalib/codebooks/c/cc01/pumfs/ind/documentation_e_april.pdf

Standard Industrial Classification - Establishments (SIC-E) 1980: http://www.statcan.gc.ca/subjects-sujets/standard-norme/sic-cti/sice-ctie80_menu-eng.htm

Table 1

Summary of Decomposition Results in Some Previous Papers

Study	Year	Database	Decomposition	Gap	Explained Component
Gunderson (1979)	1970	Census	Oaxaca (1973)	0.51	0.187 (36.7%)
Baker, Desaulniers, and Grant (1995)	1970	Census	Oaxaca (1973)	0.508	0.134 (26.4%)
Miller (1987)	1980	Census	Oaxaca (1973)	0.351	0.148 (41.2%)
Shapiro and Stelcner (1987)	1980	Census	Oaxaca (1973)	0.421	0.145 (34.4%)
Coish and Hale (1995)	1980	Census	Oaxaca (1973)	0.439	0.118 (26.9%)
Baker, Desaulniers, and Grant (1995)	1985	Census	Oaxaca (1973)	0.422	0.103 (24.4%)
Baker, Desaulniers, and Grant (1995)	1985	SCF	Oaxaca (1973)	0.441	0.092 (20.9%)
Christie and Shannon (2001)	1985	Census	Oaxaca (1973) and Cotton (1988)	0.433	0.079 (18.2%%)
Christie and Shannon (2001)	1990	SCF	Oaxaca (1973)	0.402	0.066 (16.4%)
Baker, Desaulniers, and Grant (1995)	1990	Census	Oaxaca (1973) and Cotton (1988)	0.386	0.062 (16%)
Baker, Desaulniers, and Grant (1995)	1993	SLID	Cotton (1988)	N/A	12%

Table 2

Higest Level of Education Attained by Full-time, Full-year Workers, 1991 and 2001

	1991 Census				2001 Census			
	Women Share (%)	Number	Men Share (%)	Number	Women Share (%)	Number	Men Share (%)	Number
Less than Grade 5	0.8	126	1	208	0.7	108	0.7	157
Grades 5 to 8	4.5	682	5.7	1237	1.8	301	2.8	614
Grades 9 to 13	16.3	2468	16.9	3673	11.8	1936	14	3120
Secondary - high school graduation certificate	18.3	2769	14	3037	15.8	2592	13.6	2949
Trades certificate or diploma	3.5	532	6.3	1375	2.7	446	4.8	1031
College: Without trades or college certificate or diplma	7.3	1116	5.7	1250	6.5	1060	5.8	1245
College:With trades certificate or diploma	5.1	776	10.5	2273	5.3	874	11	2369
College: with college certificate or diploma	15.7	2375	11	2386	19.6	3218	14.3	3083
University: without certificate, diploma or degree	3.7	553	4.1	886	3.3	542	3.5	755
University: with university or college certificate or diploma	8	1217	6	1307	8.4	1380	7	1513
University: With bachelor or first professional degree	11.9	1800	12.1	2634	16.9	2768	14.4	3122
University: With certificate or diploma above bachelor level	2	295	2	435	2.8	462	2	434
University: With master degree	2.4	370	3.8	816	3.9	635	4.5	976
University: With earned doctorate	0.3	47	1.1	226	0.5	80	1.2	259
Total	100	15126	100	21743	100	16402	100	21627

Table 3

Educational Attainment Shares by Age and Sex , 1991 Census*

	Age Group			
	25-34	35-44	45-54	55-65
Women				
Less than Grade 5	0.3	0.6	1.3	2.7
Grades 5 to 8	0.9	3.6	8.6	14.6
Grades 9 to 13	13.3	14.9	19.5	25.9
Secondary - high school graduation certificate	17.8	20.7	17.0	14.1
Trades certificate or diploma	3.6	3.5	3.2	3.6
College: Without trades or college certificate or diploma	8.7	6.7	6.3	7.3
College: With trades certificate or diploma	5.9	4.6	4.9	4.9
College: With college certificate or diploma	18.4	14.9	14.2	11.8
University: Without certificate, diploma or degree	4.4	4.1	2.7	1.3
University: With university or college certificate or diploma	7.7	8.7	8.7	5.4
University: With bachelor or first professional degree	15.6	12.2	8.5	4.8
University: With certificate or diploma above bachelor level	1.5	2.4	2.3	1.3
University: With master degree	1.8	3.2	2.4	2.2
University: With earned doctorate	0.1	0.4	0.5	0.1
Sample size: Women	5275	5255	3199	1345
Men				
Less than Grade 5	0.4	0.4	1.4	3.3
Grades 5 to 8	1.6	3.2	9.5	17.6
Grades 9 to 13	17.3	15.0	17.5	19.7
Secondary - high school graduation certificate	15.5	13.9	13.3	11.2
Trades certificate or diploma	5.4	6.0	7.1	8.8
College: Without trades or college certificate or diploma	7.3	5.8	4.1	3.7
College: With trades certificate or diploma	10.8	10.7	10.0	9.4
College: With college certificate or diploma	13.8	11.3	8.5	6.5
University: Without certificate, diploma or degree	4.2	5.2	3.0	2.8
University: With university or college certificate or diploma	6.0	6.6	6.4	3.9
University: With bachelor or first professional degree	14.0	13.8	9.6	6.5
University: With certificate or diploma above bachelor level	1.3	2.3	2.6	1.8
University: With master degree	2.3	4.7	5.1	3.0
University: With earned doctorate	0.2	1.0	1.9	1.9
Sample size: Men	7422	7035	4783	2430

* All values in the table other than the sample size are percentages of the sample for a particular age group.

Table 4

Educational Attainment Shares by Age and Sex , 2001 Census*

	Age group			
	25-34	35-44	45-54	55-65
Women				
Less than Grade 5	0.3	0.5	0.7	2.0
Grades 5 to 8	0.5	0.8	2.4	7.1
Grades 9 to 13	7.4	12.0	13.5	17.3
Secondary - high school graduation certificate	9.6	16.6	19.5	18.3
Trades certificate or diploma	1.5	3.0	3.2	3.5
College: Without trades or college certificate or diploma	6.9	6.9	5.7	6.3
College: With trades certificate or diploma	6.0	5.6	4.7	4.7
College: With college certificate or diploma	21.6	20.9	17.7	15.7
University: Without certificate, diploma or degree	4.0	3.1	3.3	2.3
University: With university or college certificate or diplomε	9.4	8.6	8.2	6.1
University: With bachelor or first professional degree	25.9	15.4	13.3	9.3
University: With certificate or diploma above bachelor leve	3.0	2.4	3.2	2.6
University: With master degree	3.7	3.8	4.2	3.9
University: With earned doctorate	0.2	0.6	0.4	1.1
Sample size: Women	4195	5690	4981	1500
Men				
Less than Grade 5	0.5	0.5	0.7	1.7
Grades 5 to 8	1.1	1.6	2.8	10.8
Grades 9 to 13	11.8	15.9	14.0	17.3
Secondary - high school graduation certificate	13.0	13.5	14.7	12.7
Trades certificate or diploma	2.7	4.8	5.7	7.1
College: Without trades or college certificate or diploma	7.6	5.8	5.0	3.3
College: With trades certificate or diploma	11.6	11.4	10.1	10.1
College: With college certificate or diploma	16.1	15.7	12.7	9.4
University: Without certificate, diploma or degree	4.1	3.1	3.4	3.5
University: With university or college certificate or diplomε	6.9	6.9	7.5	6.2
University: With bachelor or first professional degree	18.9	13.7	13.3	9.4
University: With certificate or diploma above bachelor leve	2.0	1.7	2.6	1.6
University: With master degree	3.4	4.1	5.9	4.6
University: With earned doctorate	0.3	1.2	1.6	2.4
Sample size: Men	5596	7471	6160	2340

* All values in the table other than the sample size are percentages of the sample for a particular age group

Table 5

Average Wage and Salary Income by Educational Atainment

Educational Attainment	1990 Women	1990 Men	1990 Ratio (%)	2000 Women	2000 Men	2000 Ratio (%)
Less than Grade 5	19903	33433	59.5	21328	32199	66.2
Grades 5 to 8	21108	36849	57.3	22007	34896	63.1
Grades 9 to 13	26074	39899	65.4	26335	39187	67.2
Secondary - high school graduation certificate	29232	43666	66.9	30759	42637	72.1
Trades certificate or diploma	28711	45586	63.0	30941	45800	67.6
College: Without trades or college certificate or diplma	30442	43716	69.6	31285	42835	73.0
College:With trades certificate or diploma	28043	45269	61.9	28712	45814	62.7
College: with college certificate or diploma	33502	48994	68.4	34590	50237	68.9
University: without certificate, diploma or degree	34524	50311	68.6	36228	51604	70.2
University: with university or college certificate or diploma	38000	52205	72.8	38682	51828	74.6
University: With bachelor or first professional degree	44629	60119	74.2	45434	63059	72.0
University: With certificate or diploma above bachelor level	50191	66099	75.9	52195	64604	80.8
University: With master degree	55623	67244	82.7	54605	72820	75.0
University: With earned doctorate	69500	76509	90.8	60657	74973	80.9

* All incomes are reported in 2000 dollars

Table 6

Field of study of post-secondary graduates

Variable	Definition	1991 census				2001 census			
			Women		Men		Women		Men
		Share (%)	Number	Share (%)	Number	Share (%)	Number	Share (%)	Number
Maj1	Education/recreation/counselling	16.0	1184	6.4	731	15.8	1563	5.6	714
Maj2	Fine and applied arts	5.0	369	2.7	310	4.5	444	2.6	327
Maj3	Humanities and related	7.2	533	4.5	518	6.8	672	5.1	648
Maj4	Social sciences and related	9.0	665	8.8	1013	11.6	1149	9.5	1219
Maj5	Commerce/managt./business administration	15.8	1171	16.5	1885	19.8	1954	17.3	2216
Maj6	Secretarial science	17.0	1262	0.9	105	11.9	1176	0.8	105
Maj7	Agricultural, biological, nutritional, and food sciences	3.6	265	4.0	461	3.6	356	4.0	517
Maj8	Engineering/applied sciences	0.7	54	7.6	876	1.4	134	8.6	1099
Maj9	Engi/app. science tech/trades	4.2	314	41.0	4690	4.9	484	37.9	4848
Maj10	Nursing and nursing assistance	13.3	983	0.6	69	7.7	760	0.4	57
Maj11	Other health prof/sci/tech	5.6	417	2.1	235	8.8	872	2.5	326
Maj12	Mathematics/physical sciences	2.5	182	4.7	541	2.9	283	5.4	691
Maj13	All other	0.2	13	0.2	18	0.2	16	0.2	20
Total		100	7412	100	11452	100	9863	100	12787

Note: The share of male (female) post-secondary graduates in field i is the number of men (women) who chose field I divided by the total number of men (women) with a post-secondary qualification.

Table 7

Field of Study Shares by Age and Sex, 1991 Census*

	Age group			
	25-34	35-44	45-54	55-65
Women				
Education/recreation/counselling	6.1	8.4	11.0	5.2
Fine and applied arts	3.3	2.4	1.6	1.3
Humanities and related	3.4	4.0	3.1	3.0
Social sciences and related	6.3	4.3	2.6	1.7
Commerce/managt./business administration	10.2	7.5	5.8	3.9
Secretarial science	9.8	8.1	6.4	8.0
Agricultural, biological, nutritional, and food sciences	2.2	1.5	1.6	1.2
Engineering/applied sciences	0.6	0.4	0.1	0.1
Engi/app. science tech/trades	2.9	2.1	1.3	0.7
Nursing and nursing assistance	4.6	7.3	8.3	7.0
Other health prof/sci/tech	3.4	2.9	2.3	1.3
Mathematics/physical sciences	1.7	1.2	0.7	0.6
All other	0.1	0.1	0.0	0.1
Total	54.6	50.0	44.7	34.1
Men				
Education/recreation/counselling	1.9	4.2	5.0	2.4
Fine and applied arts	1.4	1.4	1.5	1.6
Humanities and related	1.9	2.7	2.9	2.2
Social sciences and related	5.0	5.5	4.1	2.2
Commerce/managt./business administration	9.5	9.0	8.2	6.3
Secretarial science	0.3	0.5	0.6	0.9
Agricultural, biological, nutritional, and food sciences	2.2	2.5	1.7	1.6
Engineering/applied sciences	4.2	4.2	3.7	3.8
Engi/app. science tech/trades	23.3	22.4	19.7	18.0
Nursing and nursing assistance	0.3	0.4	0.2	0.2
Other health prof/sci/tech	1.1	1.0	1.2	0.9
Mathematics/physical sciences	2.7	2.7	2.4	1.5
All other	0.1	0.1	0.1	0.1
Total	53.8	56.4	51.2	41.8

* All values in the table other than the sample size are percentages of the sample for a particular age group.

Table 8

Field of Study Shares by Age and Sex, 2001 Census*

	Age group			
	25-34	35-44	45-54	55-65
Women				
Education/recreation/counselling	10.7	8.1	10.1	9.9
Fine and applied arts	3.8	2.8	2.1	1.3
Humanities and related	5.2	3.5	4.0	3.8
Social sciences and related	11.0	6.5	5.1	4.1
Commerce/managt./business administration	14.7	13.6	9.5	6.2
Secretarial science	6.9	7.7	6.8	7.1
Agricultural, biological, nutritional, and food sciences	2.4	2.5	1.8	1.7
Engineering/applied sciences	1.4	1.0	0.3	0.3
Engi/app. science tech/trades	3.2	3.5	2.5	1.6
Nursing and nursing assistance	1.5	4.1	5.9	5.5
Other health prof/sci/tech	6.3	5.0	5.1	4.3
Mathematics/physical sciences	2.1	1.8	1.6	0.0
All other	0.0	0.1	0.2	1.0
Total	69.3	60.2	54.7	46.8
Men				
Education/recreation/counselling	3.4	2.4	4.5	3.0
Fine and applied arts	1.8	1.4	1.4	1.5
Humanities and related	3.2	2.5	3.2	3.2
Social sciences and related	6.8	5.0	5.9	4.3
Commerce/managt./business administration	11.3	10.1	10.3	7.9
Secretarial science	0.4	0.4	0.6	0.7
Agricultural, biological, nutritional, and food sciences	2.9	2.5	2.0	1.8
Engineering/applied sciences	5.3	5.3	5.1	3.8
Engi/app. science tech/trades	21.4	24.3	21.9	20.2
Nursing and nursing assistance	0.3	0.3	0.3	0.2
Other health prof/sci/tech	1.5	1.6	1.4	1.7
Mathematics/physical sciences	3.6	3.5	2.7	2.4
All other	0.1	0.1	0.1	49.3
	0.0			
Total	61.9	59.5	59.4	50.7

* All values in the table other than the sample size are percentages of the sample for a particular age group.

Table 9

Average Wage and Salary Income by Field of Study

Field of Pos-Secondary Study	1990 Women	1990 Men	1990 Ratio (%)	2000 Women	2000 Men	2000 Ratio (%)
Education/recreation/counselling	43428	56411	77.0	42863	51332	83.5
Fine and applied arts	27936	47515	58.8	29786	41547	71.7
Humanities and related	41053	50386	81.5	41907	50068	83.7
Social sciences and related	41956	60527	69.3	41836	59214	70.7
Commerce/managt./business administration	39019	57453	67.9	41736	62890	66.4
Secretarial science	30307	45226	67.0	31082	44225	70.3
Agricultural, biological, nutritional, and food sciences	32546	45688	71.2	36681	46399	79.1
Engineering/applied sciences	48128	67252	71.6	46011	67826	67.8
Engi/app. science tech/trades	32913	47567	69.2	37147	49141	75.6
Nursing and nursing assistance	38573	42158	91.5	44174	49858	88.6
Other health prof/sci/tech	42415	60026	70.7	36656	58069	63.1
Mathematics/physical sciences	48350	60602	79.8	53558	66674	80.3
All other	32501	38227	85.0	33525	48409	69.3
No Post-Secondary Qualification	26832	40674	66.0	28284	40129	70.5

* All incomes are reported in 2000 dollars.

Table 10

Descriptive Statistics

	1990				2000			
	Women		Men		Women		Men	
NAME	MEAN	ST. DEV	MEAN	ST. DEV	MEAN	ST. DEV	MEAN	ST. DEV
LY	10.243	0.69561	10.620	0.70772	10.293	7.91E-01	10.607	0.80175
MARRIAGE	0.608	0.48813	0.715	0.45152	0.559	0.49657	0.629	0.48305
AGE	39.790	9.7977	40.477	10.125	41.709	9.4331	41.886	9.6797
AGESQUAR	1679.200	828.38	1740.900	869.67	1828.600	802.1	1848.100	832.8
IMMIGRAT	0.197	0.3974	0.197	0.39762	0.206	0.40427	0.206	0.40443
BC	0.105	0.30705	0.113	0.31655	0.119	3.23E-01	0.116	0.32058
QUE	0.242	0.42813	0.246	0.43053	0.235	4.24E-01	0.229	0.42036
ALTA	0.096	0.29405	0.095	0.29306	0.098	0.29787	0.104	0.30508
MAN	0.040	0.19611	0.039	0.19338	0.037	1.89E-01	0.037	0.18783
ONT	0.420	0.49354	0.409	0.49172	0.413	4.92E-01	0.420	0.49359
SASK	0.030	0.16954	0.027	0.16328	0.028	0.16406	0.029	0.16648
MINORITY	0.088	0.28301	0.077	0.26702	0.125	0.33016	0.115	0.31919
ONELANG	0.799	0.40055	0.785	0.41084	0.785	0.41056	0.788	0.409
BOTHEF	0.195	0.39583	0.211	0.40774	0.208	0.40617	0.208	0.40567
OCP2	0.078	0.2682	0.139	0.34647	0.096	2.94E-01	0.135	0.34205
OCP3	0.186	0.38931	0.147	0.35424	0.039	1.93E-01	0.027	0.1629
OCP4	0.060	0.2384	0.065	0.2461	0.125	3.30E-01	0.019	0.13654
OCP5	0.028	0.16412	0.016	0.12638	0.188	3.91E-01	0.057	0.23231
OCP6	0.008	8.84E-02	0.044	0.20605	0.038	1.90E-01	0.117	0.32174
OCP7	0.162	0.36886	0.020	0.13955	0.048	0.21399	0.007	8.35E-02
OCP8	0.036	0.18604	0.057	0.23138	0.046	2.10E-01	0.010	9.81E-02
OCP9	0.008	9.02E-02	0.137	0.3443	0.055	2.28E-01	0.025	0.15462
OCP10	0.199	0.39961	0.061	0.2399	0.078	0.26757	0.035	0.18414
OCP11	0.105	0.30637	0.067	0.25084	0.019	1.35E-01	0.015	0.12331
OCP12	0.047	0.21236	0.143	0.35054	0.022	1.45E-01	0.035	0.18367
OCP13	0.061	0.23853	0.047	0.21136	0.045	2.07E-01	0.027	0.1625
OCP14	0.016	0.12368	0.038	0.19228	0.025	0.15482	0.015	0.12074
OCP15					0.007	8.49E-02	0.034	0.18024
OCP16					0.017	1.29E-01	0.002	4.19E-02
OCP17					0.061	2.39E-01	0.048	0.21366
OCP18					0.002	3.98E-02	0.016	0.12618
OCP19					0.000	1.91E-02	0.024	0.1516
OCP20					0.008	8.87E-02	0.113	0.31695

OCP21					0.004	6.52E-02	0.065	0.24712
OCP22					0.005	7.14E-02	0.028	0.16424
OCP23					0.008	8.87E-02	0.022	0.14596
OCP24					0.044	2.04E-01	0.083	0.27623
OCP25					0.014	1.16E-01	0.019	0.1354
IND1	0.010	0.10103	0.010	0.10187	0.011	1.05E-01	0.011	0.1028
IND3	0.120	0.32481	0.233	0.42275	0.118	3.22E-01	0.245	0.42989
IND4	0.015	0.12211	0.057	0.23226	0.011	0.10617	0.058	0.23398
IND5	0.021	0.14258	0.068	0.25179	0.019	0.13574	0.068	0.25178
IND6	0.041	0.19842	0.061	0.23897	0.033	0.17748	0.051	0.22038
IND7	0.035	0.18321	0.067	0.2506	0.046	0.20903	0.081	0.27229
IND8	0.101	0.30189	0.091	0.28731	0.093	0.29066	0.088	0.28389
IND9	0.112	0.31577	0.052	0.22244	0.102	0.30218	0.050	0.21868
IND10	0.059	0.23607	0.051	0.22086	0.077	0.26593	0.075	0.26347
IND11	0.041	0.19919	0.051	0.21973	0.034	0.1805	0.038	0.19111
IND12	0.060	0.23743	0.071	0.25685	0.052	0.22143	0.051	0.21925
IND13	0.109	0.31142	0.068	0.25132	0.114	0.31835	0.057	0.23266
IND14	0.172	0.3773	0.030	0.17119	0.184	0.3873	0.033	0.17988
IND15	0.044	0.20502	0.024	0.15193	0.045	0.2077	0.029	0.16779
IND16	0.050	0.21777	0.034	0.18132	0.056	0.22951	0.039	0.19322

Table11

Diagnostic Tests for Regression Models

	Women		Men	
	1990	**2000**	**1990**	**2000**
Test for normality				
Jarque-Bera test	1458460 (0.000)	1072747 (0.000)	2995623 (0.000)	1129400 (0.000)
Test for multicollinearity				
Condition Index	94.046	103.31	92.295	98.964
R-SQUARE OF MARRIAGE	0.0635	0.0785	0.1214	0.1503
R-SQUARE OF AGE	0.9851	0.986	0.9863	0.9863
R-SQUARE OF AGESQUAR	0.985	0.9859	0.9861	0.9862
R-SQUARE OF IMMIGRAT	0.36	0.4483	0.3273	0.4336
R-SQUARE OF BC	0.5754	0.5979	0.5788	0.6061
R-SQUARE OF QUE	0.733	0.7261	0.7302	0.7384
R-SQUARE OF ALTA	0.5551	0.5558	0.5396	0.5802
R-SQUARE OF MAN	0.3537	0.3347	0.3362	0.342
R-SQUARE OF ONT	0.7713	0.7699	0.7612	0.7793
R-SQUARE OF SASK	0.2921	0.2734	0.2647	0.2908
R-SQUARE OF MINORITY	0.3129	0.4307	0.2719	0.4061
R-SQUARE OF ONELANG	0.9658	0.9667	0.9755	0.9738
R-SQUARE OF BOTHEF	0.9665	0.9676	0.9759	0.9743
R-SQUARE OF EDU2	0.8444	0.7339	0.8513	0.7931
R-SQUARE OF EDU3	0.947	0.9429	0.9389	0.9455
R-SQUARE OF EDU4	0.9521	0.9552	0.9301	0.9432
R-SQUARE OF EDU5	0.8522	0.8343	0.9129	0.9036
R-SQUARE OF EDU6	0.9017	0.9071	0.8581	0.8851
R-SQUARE OF EDU7	0.8899	0.9058	0.9438	0.9533
R-SQUARE OF EDU8	0.9591	0.9692	0.9463	0.9622
R-SQUARE OF EDU9	0.8265	0.8387	0.8153	0.8283
R-SQUARE OF EDU10	0.9299	0.9392	0.9119	0.9315
R-SQUARE OF EDU11	0.9512	0.9665	0.9538	0.9641
R-SQUARE OF EDU12	0.7856	0.8527	0.792	0.812
R-SQUARE OF EDU13	0.818	0.8854	0.8766	0.9043
R-SQUARE OF EDU14	0.3846	0.5107	0.6789	0.7333
R-SQUARE OF MAJ1	0.7915	0.8042	0.7346	0.718
R-SQUARE OF MAJ3	0.6275	0.6286	0.6418	0.6717
R-SQUARE OF MAJ4	0.6777	0.7455	0.7745	0.7949
R-SQUARE OF MAJ5	0.7644	0.8135	0.8545	0.8659
R-SQUARE OF MAJ6	0.7716	0.727	0.2599	0.2526
R-SQUARE OF MAJ7	0.4326	0.468	0.5986	0.6184
R-SQUARE OF MAJ8	0.1664	0.2952	0.7607	0.7871
R-SQUARE OF MAJ9	0.4693	0.5376	0.9231	0.9225
R-SQUARE OF MAJ10	0.7584	0.7343	0.2292	0.2723
R-SQUARE OF MAJ11	0.5673	0.7043	0.4674	0.5979
R-SQUARE OF MAJ12	0.3784	0.4445	0.661	0.6985
R-SQUARE OF MAJ13	0.0385	0.0444	0.058	0.0609
R-SQUARE OF OCP2	0.9309	0.9035	0.8794	0.8414
R-SQUARE OF OCP3	0.9662	0.8014	0.8878	0.5558
R-SQUARE OF OCP4	0.9151	0.9223	0.7934	0.4629

R-SQUARE OF OCP5	0.8352	0.9432	0.4994	0.7203
R-SQUARE OF OCP6	0.6223	0.8026	0.7336	0.832
R-SQUARE OF OCP7	0.9624	0.861	0.5496	0.4666
R-SQUARE OF OCP8	0.8683	0.8448	0.7738	0.4636
R-SQUARE OF OCP9	0.6135	0.8524	0.8855	0.5447
R-SQUARE OF OCP10	0.9678	0.9072	0.7843	0.75
R-SQUARE OF OCP11	0.9471	0.6721	0.7997	0.4415
R-SQUARE OF OCP12	0.8994	0.6995	0.8893	0.6201
R-SQUARE OF OCP13	0.9163	0.8353	0.7466	0.5771
R-SQUARE OF OCP14	0.7523	0.7633	0.7088	0.5747
R-SQUARE OF OCP15		0.4496		0.64
R-SQUARE OF OCP16		0.6599		0.1156
R-SQUARE OF OCP17		0.8673		0.7006
R-SQUARE OF OCP18		0.1554		0.4451
R-SQUARE OF OCP19		0.0444		0.5623
R-SQUARE OF OCP20		0.4772		0.8352
R-SQUARE OF OCP21		0.3706		0.7674
R-SQUARE OF OCP22		0.3681		0.5767
R-SQUARE OF OCP23		0.6426		0.6183
R-SQUARE OF OCP24		0.8367		0.7996
R-SQUARE OF OCP25		0.6201		0.4967
R-SQUARE OF IND1	0.5696	0.7202	0.2704	0.3779
R-SQUARE OF IND3	0.9233	0.9433	0.8536	0.8964
R-SQUARE OF IND4	0.6251	0.6366	0.6349	0.7244
R-SQUARE OF IND5	0.6932	0.7475	0.6729	0.7531
R-SQUARE OF IND6	0.8144	0.8297	0.6519	0.6914
R-SQUARE OF IND7	0.7889	0.8719	0.6771	0.7771
R-SQUARE OF IND8	0.9119	0.9313	0.738	0.7963
R-SQUARE OF IND9	0.9176	0.934	0.6459	0.7031
R-SQUARE OF IND10	0.8611	0.916	0.6306	0.7662
R-SQUARE OF IND11	0.8165	0.835	0.6258	0.6443
R-SQUARE OF IND12	0.8627	0.8837	0.695	0.7037
R-SQUARE OF IND13	0.9193	0.9459	0.7304	0.7947
R-SQUARE OF IND14	0.9425	0.9604	0.5419	0.6666
R-SQUARE OF IND15	0.8294	0.8815	0.4745	0.6689
R-SQUARE OF IND16	0.8438	0.8933	0.5358	0.6436
R-SQUARE OF CONSTANT	0.0000	0.0000	0.0000	0.0000

Test for heteroskedasticity error				
Modified Breusch-Pagan-Godfrey (BPG) test	16402 (0.000)	16645 (0.000)	21743 (0.000)	21627 (0.000)
Test for specification error				
Ramssry's RESET test				
RESET(2)	1.1751 (0.278)	4.7685 (0.029)	0.0098536 (0.921)	16.995 (0.000)
RESET(3)	0.7823 (0.457)	2.5485 (0.078)	2.7919 (0.061)	8.5858 (0.000)
RESET(4)	1.5696 (0.194)	2.3156 (0.074)	1.8857 (0.130)	5.8638 (0.001)
FRESETL test				
FRESET(1)	1.4003 (0.247)	2.9551 (0.052)	3.5404 (0.029)	8.8167 (0.000)
FRESET(2)	0.72056 (0.578)	1.6117 (0.168)	2.1882 (0.068)	4.7609 (0.001)
FRESET(3)	3.0339 (0.006)	1.8669 (0.082)	4.1984 (0.000)	3.4648 (0.002)
FRESETS test				
FRESET(1)	0.48565 (0.615)	1.1645 (0.312)	3.5392 (0.029)	4.4205 (0.012)
FRESET(2)	2.3129 (0.055)	1.4369 (0.219)	4.1407 (0.002)	3.0963 (0.015)
FRESET(3)	1.8696 (0.082)	1.2702 (0.267)	2.9851 (0.006)	2.4908 (0.021)

Notes:

1. the numbers of bracket following RESET means the highest power of additional variables stared on 2, which means RESET (3) has two more additional variables with power 2 and 3.

2. the numbers of bracket following FRESETL and FRESETS mean the model has additional variables in number of two times them, for examaple, FRESET (2) has 4 more additional variables into the model.

Table12

Auxiliary R-square EDU3 and IND3 are Excluded from the Model

	Women		Men	
	1990	2000	1990	2000
Test for multicollinearity				
Condition Index	88.597	92.04	87.602	91.115
R-SQUARE OF EDU1	0.1244	0.0918	0.0944	0.067
R-SQUARE OF EDU2	0.2416	0.1745	0.2545	0.174
R-SQUARE OF EDU4	0.4539	0.5104	0.3772	0.4138
R-SQUARE OF EDU5	0.5985	0.5115	0.8112	0.766
R-SQUARE OF EDU6	0.2888	0.3326	0.2279	0.26
R-SQUARE OF EDU7	0.6588	0.6624	0.8746	0.8789
R-SQUARE OF EDU8	0.8685	0.8822	0.8796	0.899
R-SQUARE OF EDU9	0.1917	0.2228	0.1987	0.1989
R-SQUARE OF EDU10	0.7945	0.7895	0.8143	0.8277
R-SQUARE OF EDU11	0.8596	0.8829	0.9029	0.9103
R-SQUARE OF EDU12	0.552	0.6179	0.6346	0.625
R-SQUARE OF EDU13	0.5947	0.6759	0.7671	0.7847
R-SQUARE OF EDU14	0.1909	0.2377	0.5104	0.5387
R-SQUARE OF IND1	0.1911	0.5211	0.0696	0.3403
R-SQUARE OF IND2	0.1108	0.1	0.1355	0.2315
R-SQUARE OF IND4	0.1349	0.1369	0.1819	0.3051
R-SQUARE OF IND5	0.1713	0.264	0.1914	0.3842
R-SQUARE OF IND6	0.3021	0.2997	0.2017	0.215
R-SQUARE OF IND7	0.2444	0.3625	0.2074	0.3151
R-SQUARE OF IND8	0.5224	0.5999	0.2807	0.3655
R-SQUARE OF IND9	0.531	0.5615	0.2632	0.3018
R-SQUARE OF IND10	0.3909	0.4938	0.2312	0.3481
R-SQUARE OF IND11	0.3177	0.3328	0.2216	0.2812
R-SQUARE OF IND12	0.3961	0.4266	0.2706	0.3484
R-SQUARE OF IND13	0.6343	0.7621	0.4771	0.6347
R-SQUARE OF IND14	0.6684	0.7591	0.2463	0.4145
R-SQUARE OF IND15	0.3788	0.5662	0.1659	0.4661
R-SQUARE OF IND16	0.3722	0.4702	0.1561	0.2448

Table 13

OLS Estimates for1990 for Full-Time, Full-year workers Aged 25-65

Dependent variable: ly (log of annual wage and salary income)

	Women			Men		
Variable	**Coefficient**	**t-statistic**	**p-value**	**Coefficient**	**t-statistic**	**p-value**
AGE	0.0307	7.344	0.000	0.0571	14.780	0.000
AGESQUARE	-0.0003	-6.378	0.000	-0.0006	-12.800	0.000
MARRIAGE	-0.0141	-1.326	0.185	0.1442	13.460	0.000
IMMIGRANT	0.0029	0.188	0.851	-0.0272	-1.921	0.055
BC	0.1488	6.034	0.000	0.1652	8.714	0.000
QUE	0.0973	4.198	0.000	0.0138	0.745	0.457
ALTA	0.1311	5.261	0.000	0.0729	3.593	0.001
MAN	0.0247	0.697	0.486	-0.0352	-1.272	0.203
ONT	0.1905	8.662	0.000	0.1532	9.440	0.000
SASK	0.0253	0.767	0.443	-0.0402	-1.228	0.219
MINORITY	-0.0871	-3.984	0.000	-0.2043	-9.399	0.000
ONELANG	0.1150	1.393	0.164	0.0391	0.495	0.620
BOTHEF	0.1540	1.851	0.064	0.0778	0.974	0.330
EDU2	0.0945	1.029	0.304	0.1790	1.975	0.048
EDU3	0.2338	2.598	0.009	0.2579	2.843	0.004
EDU4	0.3283	3.662	0.000	0.3744	4.137	0.000
EDU5	0.1610	1.580	0.114	0.3285	3.416	0.001
EDU6	0.3465	3.807	0.000	0.3571	3.877	0.000
EDU7	0.1685	1.635	0.102	0.3562	3.697	0.000
EDU8	0.2123	2.034	0.042	0.4090	4.243	0.000
EDU9	0.4262	4.607	0.000	0.4144	4.470	0.000
EDU10	0.2765	2.623	0.009	0.4428	4.582	0.000
EDU11	0.4015	3.829	0.000	0.5443	5.578	0.000
EDU12	0.4835	4.494	0.000	0.6168	6.218	0.000
EDU13	0.5379	4.926	0.000	0.6155	6.255	0.000
EDU14	0.7970	6.993	0.000	0.7273	6.905	0.000
MAJ1	0.1535	2.831	0.005	-0.0775	-1.789	0.074
MAJ3	0.0336	0.561	0.575	-0.1800	-4.279	0.000
MAJ4	0.1185	2.129	0.033	0.0332	0.935	0.350
MAJ5	0.2077	3.892	0.000	0.0154	0.447	0.655
MAJ6	0.1531	2.966	0.003	-0.1873	-1.771	0.077
MAJ7	0.0476	0.752	0.452	-0.1208	-2.503	0.012
MAJ8	0.2366	2.952	0.003	0.0598	1.595	0.111
MAJ9	0.1939	3.483	0.000	0.0271	0.841	0.400
MAJ10	0.1697	2.911	0.004	0.0245	0.502	0.616
MAJ11	0.3408	6.143	0.000	0.1208	2.173	0.030
MAJ12	0.2911	5.004	0.000	-0.0397	-0.902	0.367
MAJ13	0.2410	2.855	0.004	-0.1340	-0.944	0.345
OCP2	-0.1435	-2.395	0.017	-0.2030	-4.794	0.000
OCP3	-0.1580	-2.685	0.007	-0.2977	-7.072	0.000
OCP4	-0.3633	-5.890	0.000	-0.4413	-9.474	0.000
OCP5	-0.2455	-3.926	0.000	-0.4068	-7.430	0.000
OCP6	-0.4955	-5.035	0.000	-0.3527	-7.686	0.000
OCP7	-0.3964	-6.683	0.000	-0.3981	-8.573	0.000
OCP8	-0.4235	-6.357	0.000	-0.3441	-7.299	0.000
OCP9	-0.5261	-5.715	0.000	-0.4279	-9.613	0.000

OCP10	-0.4362	-7.410	0.000	-0.5866	-13.170	0.000
OCP11	-0.5550	-8.998	0.000	-0.4290	-9.473	0.000
OCP12	-0.6223	-8.984	0.000	-0.5600	-12.510	0.000
OCP13	-0.5636	-8.924	0.000	-0.6679	-13.860	0.000
OCP14	-0.6420	-7.697	0.000	-0.6133	-12.040	0.000
IND1	-0.7979	-9.327	0.000	-0.6236	-11.190	0.000
IND3	-0.1867	-3.956	0.000	-0.1988	-8.614	0.000
IND4	-0.2658	-4.259	0.000	-0.3112	-10.610	0.000
IND5	-0.1306	-2.471	0.013	-0.1627	-5.988	0.000
IND6	-0.0450	-0.895	0.371	-0.0981	-3.829	0.000
IND7	-0.2405	-4.812	0.000	-0.2695	-10.110	0.000
IND8	-0.4426	-9.160	0.000	-0.4284	-16.400	0.000
IND9	-0.2230	-4.769	0.000	-0.2581	-8.584	0.000
IND10	-0.2286	-4.782	0.000	-0.2573	-9.195	0.000
IND11	-0.0413	-0.843	0.399	-0.1915	-6.811	0.000
IND12	-0.1390	-2.984	0.003	-0.1940	-7.834	0.000
IND13	-0.2263	-4.730	0.000	-0.2701	-9.838	0.000
IND14	-0.2622	-5.572	0.000	-0.3885	-10.430	0.000
IND15	-0.6581	-12.230	0.000	-0.7151	-15.880	0.000
IND16	-0.4905	-9.300	0.000	-0.4694	-12.410	0.000
CONSTANT	9.5901	62.920	0.000	9.3844	64.410	0.000
R-square	0.2358			0.2038		
Ajusted R-square	0.2325			0.2014		
F statistic	78.5173		0.000	82.325379		0.000
Observations	15126			21743		

Table 14

OLS Estimates for 2000 for Full-Time, Full-year workers Aged 25-65
Dependent variable: ly (log of annual wage and salary income)

Variables	Women Coefficient	t-statistic	p-value	Men Coefficient	t-statistic	p-value
MARRIAGE	0.0104	0.879	0.380	0.1379	12.180	0.000
AGE	0.0534	11.080	0.000	0.0587	13.040	0.000
AGESQUAR	-0.0005	-9.470	0.000	-0.0006	-11.400	0.000
IMMIGRANT	-0.0624	-3.177	0.001	-0.0628	-3.825	0.000
BC	0.2778	10.220	0.000	0.2612	11.010	0.000
QUE	0.0868	3.352	0.001	0.1091	4.700	0.000
ALTA	0.1757	6.566	0.000	0.2301	9.175	0.000
MAN	0.0961	2.701	0.007	0.0719	2.370	0.018
ONT	0.2627	11.280	0.000	0.2651	12.440	0.000
SASK	0.0679	1.749	0.080	0.1539	4.733	0.000
MINORITY	-0.1215	-4.735	0.000	-0.2407	-10.940	0.000
ONELANG	0.0674	0.593	0.553	0.2592	2.907	0.004
BOTHEF	0.1242	1.082	0.279	0.2874	3.171	0.002
EDU2	0.2110	1.474	0.141	0.0984	1.062	0.288
EDU3	0.2486	1.750	0.080	0.1624	1.826	0.068
EDU4	0.3460	2.436	0.015	0.2485	2.801	0.005
EDU5	0.2485	1.673	0.094	0.1232	1.264	0.206
EDU6	0.3374	2.354	0.019	0.2792	3.103	0.002
EDU7	0.2856	1.947	0.052	0.1622	1.691	0.091
EDU8	0.3795	2.619	0.009	0.2140	2.246	0.025
EDU9	0.4770	3.304	0.001	0.3579	3.902	0.000
EDU10	0.4639	3.179	0.001	0.1975	2.052	0.040
EDU11	0.6263	4.308	0.000	0.3911	4.085	0.000
EDU12	0.7362	5.036	0.000	0.4184	4.145	0.000
EDU13	0.7405	5.056	0.000	0.4784	4.939	0.000
EDU14	0.7634	4.310	0.000	0.5510	4.978	0.000
MAJ1	0.0162	0.398	0.690	-0.0079	-0.166	0.869
MAJ3	-0.0302	-0.690	0.490	-0.0531	-1.097	0.273
MAJ4	0.0246	0.623	0.533	0.1064	2.551	0.011
MAJ5	0.0936	2.512	0.012	0.1821	4.713	0.000
MAJ6	0.0160	0.404	0.686	0.0730	0.989	0.323
MAJ7	-0.0287	-0.536	0.592	0.0748	1.562	0.118
MAJ8	-0.0633	-0.701	0.483	0.1545	3.611	0.000
MAJ9	0.1063	2.419	0.016	0.1427	3.769	0.000

MAJ10	0.1363	2.698	0.007	0.1512	2.134	0.033
MAJ11	0.0698	1.377	0.168	0.2083	3.630	0.000
MAJ12	0.1478	2.865	0.004	0.1596	3.686	0.000
MAJ13	-0.1795	-1.292	0.196	-0.1353	-0.339	0.735
OCP2	-0.1535	-3.064	0.002	-0.2491	-7.581	0.000
OCP3	-0.2666	-4.907	0.000	-0.3174	-7.555	0.000
OCP4	-0.4238	-8.373	0.000	-0.4844	-10.620	0.000
OCP5	-0.4800	-9.611	0.000	-0.6412	-17.450	0.000
OCP6	-0.2138	-3.918	0.000	-0.3452	-10.150	0.000
OCP7	-0.1462	-2.268	0.023	-0.0788	-1.153	0.249
OCP8	-0.4075	-6.348	0.000	-0.4409	-6.616	0.000
OCP9	-0.3641	-6.928	0.000	-0.3925	-9.238	0.000
OCP10	-0.2257	-3.845	0.000	-0.3840	-7.942	0.000
OCP11	-0.3913	-6.345	0.000	-0.5141	-9.253	0.000
OCP12	-0.2431	-4.264	0.000	-0.4640	-10.900	0.000
OCP13	-0.5830	-10.150	0.000	-0.5859	-13.000	0.000
OCP14	-0.6531	-8.776	0.000	-0.4648	-6.683	0.000
OCP15	-0.2944	-4.339	0.000	-0.4067	-9.777	0.000
OCP16	-0.7978	-11.160	0.000	-1.0086	-3.529	0.000
OCP17	-0.5804	-10.040	0.000	-0.7849	-18.000	0.000
OCP18	-0.2496	-2.589	0.010	-0.3463	-7.498	0.000
OCP19	-0.3206	-1.716	0.086	-0.6191	-12.130	0.000
OCP20	-0.4815	-6.025	0.000	-0.5019	-13.880	0.000
OCP21	-0.6384	-4.314	0.000	-0.6364	-16.640	0.000
OCP22	-0.7566	-5.483	0.000	-0.6893	-14.220	0.000
OCP23	-0.6847	-7.023	0.000	-0.7197	-13.220	0.000
OCP24	-0.5947	-10.010	0.000	-0.5965	-15.360	0.000
OCP25	-0.7303	-8.862	0.000	-0.7018	-12.300	0.000
IND1	-0.4487	-5.024	0.000	-0.6030	-7.883	0.000
IND3	-0.0936	-1.435	0.151	-0.2310	-6.697	0.000
IND4	-0.2459	-3.119	0.002	-0.2934	-7.417	0.000
IND5	-0.0833	-1.096	0.273	-0.2564	-6.598	0.000
IND6	0.0296	0.434	0.664	-0.1929	-5.120	0.000
IND7	-0.1288	-1.930	0.054	-0.3293	-8.784	0.000
IND8	-0.3585	-5.457	0.000	-0.5081	-13.320	0.000
IND9	-0.1296	-2.009	0.045	-0.2826	-7.049	0.000
IND10	-0.1008	-1.562	0.118	-0.3363	-8.818	0.000
IND11	0.0660	0.979	0.328	-0.2718	-7.125	0.000

IND12	-0.0246	-0.379	0.705	-0.2956	-7.428	0.000
IND13	-0.2033	-2.916	0.004	-0.3566	-7.730	0.000
IND14	-0.2531	-3.895	0.000	-0.5662	-11.260	0.000
IND15	-0.4877	-6.492	0.000	-0.7917	-13.270	0.000
IND16	-0.2970	-4.318	0.000	-0.4840	-11.550	0.000
CONSTANT	8.9429	44.660	0.000	9.2127	59.910	0.000
R-squared	0.2114			0.2024		
Adjusted R-squared	0.2076			0.1996		
F statistic	63.795836		0.000	72.350844		0.000
Observations	16402			21627		

Table 15

Oaxaca Decomposition of the Gender Earnings Gap*

	1990	2000
Total Gap	0.376	0.313
Explained Component (%):	31	31.9
Educational attainment	-0.5	-3.2
Field of study	6.2	4.5
Location	-0.1	0.6
Marital/family	4.1	-3.7
Occupation	5.8	3.23
Industry	13.8	23.2
Age	0.9	6.42
Other Variables	0.8	0.8
Constant	0	0
Unexplained Component (%):	69	68.1
Educational attainment	26	-51.1
Field of study	-26.5	9.8
Location	-11.7	3.55
Marital/family	25.6	22.7
Occupation	-14.6	-27.3
Industry	-12.2	-67.2
Age	162	37.3
Other Variables	-24.4	54.2
Constant	-54.7	86.2

Notes:

* The total gap is the difference in the sample means of the log earnings between men and women.

1. Total gap is real number, while others are percentage of their shares

Table 16

Cotton Decomposition of the Gender Earnings Gap*

	1990	2000
Total Gap	0.376	0.313
Explained Component (%):	17.0	22.8
Educational attainment	-0.8	-4.2
Field of study	5.25	3.4
Location	-0.1	0.6
Marital/family	1.44	1.9
Occupation	-1.65	7
Industry	11.7	20
Aag	0.6	-0.4
Other Variables	0.6	0.6
Constant	0	0
Male advantage Component (%):	54.8	38.5
Educational attainment	15.6	-21.1
Field of study	-14.6	5.07
Location	-6.84	1.57
Marital/family	17.8	11.1
Occupation	-1.22	-10.5
Industry	-5.01	-24.5
Aag	95.6	16.1
Other Variables	-14.2	23.6
Constant	-32.2	37.2
Female disadvantage Component (%):	28.5	38.7
Educational attainment	10.6	-29.1
Field of study	-10.9	5.5
Location	-4.8	2
Marital/family	10.5	12.9
Occupation	-6	-15.5
Industry	-5	-38.2
Aag	66.3	21.2
Other Variables	-10	30.8
Constant	-22.4	49

Notes:

* The total gap is the difference between the logs of male and female earnings

1. Total gap is real number, while others are percentage of their shares

Appendix A

Variable Name	Definition
LY	equals log of women's or men's income in 2000 constant dollars *
AGE	the number of actual age of individual, the range will be from 25 to 65
AGESQUARE	age squared
MARRIAGE	equals 1 if the person is marride and not separate
IMMIGRANT	equals 1 if the person is immigration
MINORITY	equals 1 if the person is visible minority
Language	
ONELANG	equals 1 if the person can only speak one of official language
BOTHEF	equals 1 if the person can speak both of official language
Province	
PRO24	equals 1 if the person is located in Quebec
PRO35	equals 1 if the person is located in Ontario
PRO46	equals 1 if the person is located in Manitoba
PRO47	equals 1 if the person is located in Saskatchewan
PRO48	equals 1 if the person is located in Alberta
PRO59	equals 1 if the person is located in British Columbia
Educational attaiment	
EDU2	equals 1 if the person's educational attainment is grades 5 to 8
EDU3	equals 1 if the person's educational attainment is grades 9 to 13
EDU4	equals 1 if the person's educational attainment is high school graduation certificate
EDU5	equals 1 if the person's educational attainment is trades certificate or diploma
EDU6	equals 1 if the person's educational attainment is college without trades certificate or diploma
EDU7	equals 1 if the person's educational attainment is College: With trades certificate or diploma
EDU8	equals 1 if the person's educational attainment College: With college certificate or diploma
EDU9	equals 1 if the person's educational attainment is University: Without certificate, diploma or degree
EDU10	equals 1 if the person's educational attainment is University: With university or college certificate or diploma
EDU11	equals 1 if the person's educational attainment is University: With bachelor or first professional degree
EDU12	equals 1 if the person's educational attainment is University: With certificate or diploma above bachelor level
EDU13	equals 1 if the person's educational attainment is University: With master degree

EDU14	equals 1 if the person's educational attainment is University: With earned doctorate

Field of study

MAJ1	equals 1 if the person's field of study is Educational, recreational and counselling services
MAJ3	equals 1 if the person's field of study is Humanities and related
MAJ4	equals 1 if the person's field of study is Social sciences and related
MAJ5	equals 1 if the person's field of study is Commerce/management/business administration.
MAJ6	equals 1 if the person's field of study is Secretarial science
MAJ7	equals 1 if the person's field of study is Agricultural, biological, nutritional, and food sciences
MAJ8	equals 1 if the person's field of study is Engineering/applied sciences
MAJ9	equals 1 if the person's field of study is Engi/app. science tech/trades
MAJ10	equals 1 if the person's field of study is Nursing and nursing assistance
MAJ11	equals 1 if the person's field of study is Other health profession/science/technology
MAJ12	equals 1 if the person's field of study is Mathematics/physical sciences
MAJ13	equals 1 if the person's field of study is All other

Occupation (1991 Census)

OCP2	equals 1 if the person's occupation is Middle and other management
OCP3	equals 1 if the person's occupation is Professionals
OCP4	equals 1 if the person's occupation is semi-professionals
OCP5	equals 1 if the person's occupation is Supervisors
OCP6	equals 1 if the person's occupation is Foremen/women
OCP7	equals 1 if the person's occupation is Administration. & Secretarial and clerical
OCP8	equals 1 if the person's occupation is Sales/Service
OCP9	equals 1 if the person's occupation is Skilled crafts/trades
OCP10	equals 1 if the person's occupation is Clerical workers
OCP11	equals 1 if the person's occupation is Sales and Service
OCP12	equals 1 if the person's occupation is Semi-skilled manual
OCP13	equals 1 if the person's occupation is Sales/Service
OCP14	equals 1 if the person's occupation is Other manual workers

Occupation (2001 Census)

OCP2	equals 1 if the person's occupation is Other management occupations
OCP3	equals 1 if the person's occupation is Professional occupations in business and finance
OCP4	equals 1 if the person's occupation is Financial, secretarial and administrative occupations
OCP5	equals 1 if the person's occupation is Clerical occupations and clerical supervisors
OCP6	equals 1 if the person's occupation is Occupations in natural and applied sciences
OCP7	equals 1 if the person's occupation is Professional occupations in health, registered nurses and supervisors
OCP8	equals 1 if the person's occupation is Technical, assisting and related occupations in health
OCP9	equals 1 if the person's occupation is Occupations in social science, government services and religion
OCP10	equals 1 if the person's occupation is Teachers and professors

OCP11	equals 1 if the person's occupation is Occupations in art, culture, recreation and sport
OCP12	equals 1 if the person's occupation is Wholesale, technical, insurance, real estate sales specialists, and retail, wholesale and grain buyers
OCP13	equals 1 if the person's occupation is Retail trade supervisors, salespersons, sales clerks and cashiers
OCP14	equals 1 if the person's occupation is Chefs and cooks, supervisors, and other occupations in food and beverage service
OCP15	equals 1 if the person's occupation is Occupations in protective services
OCP16	equals 1 if the person's occupation is Childcare and home support workers
OCP17	equals 1 if the person's occupation is Service supervisors, occupations in travel and accommodation, attendants in recreation and sport and sales
OCP18	equals 1 if the person's occupation is Contractors and supervisors in trades and transportation
OCP19	equals 1 if the person's occupation is Construction trades
OCP20	equals 1 if the person's occupation is Other trades occupations
OCP21	equals 1 if the person's occupation is Transport and equipment operators
OCP22	equals 1 if the person's occupation is Trades helpers, construction, and transportation labourers and related occupations
OCP23	equals 1 if the person's occupation is Occupations unique to primary industries
OCP24	equals 1 if the person's occupation is Supervisors, machine operators and assemblers in manufacturing
OCP25	equals 1 if the person's occupation is Labourers in processing, manufacturing and utilities
Industry	
IND1	equals 1 if the person's industry is Agriculture
IND3	equals 1 if the person's industry is Manufacturing
IND4	equals 1 if the person's industry is Construction
IND5	equals 1 if the person's industry is Transportation and storage
IND6	equals 1 if the person's industry is Communication/other utilities
IND7	equals 1 if the person's industry is Wholesale trade
IND8	equals 1 if the person's industry is Retail trade
IND9	equals 1 if the person's industry is Finance, ins. and real estate
IND10	equals 1 if the person's industry is Business services
IND11	equals 1 if the person's industry is Government services: Federal
IND12	equals 1 if the person's industry is Governmentt. services: Other
IND13	equals 1 if the person's industry is Educational services
IND14	equals 1 if the person's industry is Health and social services
IND15	equals 1 if the person's industry is Accom/food/beverage services
IND16	equals 1 if the person's industry is Other services

Notes:

The reference individual: Language: neither english nor franch; Province: Atlantic region; Educational attainment: less than grade 5; Field of study: fine and arts; Occupation: senior management; Industry: other primary industries.

* In order to transfer the 1990 income to 2000 constant dollars, I use the inflation factor 1.2103 multiplying with 1990 income.