THE CAUSAL EFFECT OF EDUCATION ON WAGES

Investigating the economic return of education in Ontario, Canada from 2003 to 2013

Written by

Paul Lim
1000720080

Zihan Zhao
1000749087

A research paper
presented to the University of Toronto
in fulfillment of the
essay requirement for the degree of
Master of Arts
in
Economics

Toronto, Canada

The Causal Effect of Education on Wages

Paul Lim, Zihan Zhao

University of Toronto, Canada

Excutive Summary

We study the causal effect of education on wages and examine whether the benefit of an increase in wages differ for different people in Canada in a systematic way. Previous studies have concluded that there is significant variation or heterogeneity in the returns to education on wages, and have voiced concerns that this heterogeneity in the benefits of education on wage is contributing to increasing wage inequality in Canada. Because wage inequality is correlated with undesirable phenomena such as macroeconomic instability, and because the Canadian government spends huge sums of money to subsidize education, it is important to see if the benefit of education only accrues to certain groups or people with certain characteristics.

We use the Labour Force Survey data from 2003-2013 (for Ontario) to examine the true benefits of education on wages and how the benefits vary in the population. We also examine how other characteristics such as experience, sex, membership in a trade union, and what industry the individual is working in affects the hourly wage of the individual. Our results suggest that the benefits of education are higher for people with higher ability and since people have varying ability in the population, the benefits of education also vary within the population. Therefore to improve the wage outcomes of people with low ability more is needed than simply further education. In addition, we find that other characteristics such as sex and experience are all highly significant in determining an individual's hourly wage. In particular, Men earn more than women on average, people in certain industries earn more than others, and as the size of the firm measured as the number of workers on site increase so does an individual's wage in that firm. In measuring these effects we not only examine the average as is common in the literature, but also how the effects change for low versus high wage-earners. Our study is the first to study the changing effects of explanatory variables such as education, sex, firm size, and industry for Canada, and we examine these effects using a novel econometric technique called Quantile-IV estimation based on Chernozhukov and Hansen's (2005, 2006).

1. Introduction

A recent Globe and Mail article has voiced concerns that increasing wage inequality will "hurt social inclusion in Canada" (Eggleton, 2013). There are other concerns that economists voice about wage inequality, such as correlation with macroeconomic instability. International studies have suggested that in recent years education has increased wage inequality around the world. They

Preprint submitted to Department of Economics, University of Toronto *December 3, 2013*

suggest that the potential mechanisms through which education increases wage inequality are: a) the complementarity of education and ability, b) the evolving conditions of the supply and demand of jobs causing heterogeneity in returns to education, and c) asymmetric information leading to too many students entering high-skills programs, thereby depressing returns to education for the lowest quantiles. In this paper we study the causal effect of education on wages, and whether education could have contributed to recent increases in wage inequality in Canada. Our original contributions to the literature are to first study the heterogeneity of the causal effect of education on wage in Canada through the use of the Labour Force Survey data, and the use of spouse's education as instruments in IV-Quantile regression estimation technique.

There are two potential beneficiaries of our study. First, prospective students will benefit in their education decisions from the dissemination of not only true causal effects of education on wages but the distribution of these effects. We mentioned that some economists suspect asymmetries of information on the part of the students, that they do not truly understand the costs and benefits of continuing education. The results of our study will help these students understand the costs, benefits, and risks they face post-graduation. Second, governments can benefit in their decisions regarding education subsidies from understanding the true causal effects education on wages. Governments provide huge subsidies to education through funding, scholarships, and student loans and bursaries. It would greatly help governments to discover not only the average causal effect of education on wages but the distribution of these effects and the distributional outcome of wages of individuals benefited by government education subsidies.

2. Literature

Our paper combines two streams of literature on the effect of education on wages. First, there is an extensive literature on using IV estimation techniques to find the true causal effect of wages (Card, 1999), which generally find that returns to education via IV estimation is higher than estimated Mincer returns to Education. Second, there is a smaller literature studying the distribution of the non-causal effect of education on wages (Martins& Pereira, 2004) which does not include Canada. They find that there is evidence of heterogeneity in returns to education in the 16 countries they examine. We combine the two methodologies to study the distribution of the causal effect of education on wages using an IV-Quantile regression estimation approach. Our data is the Labour Force Survey, a monthly rotating panel survey conducted by Statistics Canada. We are the first to study heterogeneity of returns to education in Canada, and the first to use spouse's education as instruments in quantile-IV regression. There is a direct precedent study which uses twins' education as instruments to study heterogeneity in returns to education using quantile-IV for a specialized data set on wage outcomes for twins in US (Arias, Hallock,& Sosa-Escudero, 2002).

3. Data and Empirical Strategy

3.1. ***Data***

We use the Labour Force Survey micro data from 2003-2013 to study the effect of education on wages. The Labour Force Survey is monthly survey with a rotating panel design that collects information on topics such as labour, unemployment, education, and demographic characteristics. It has a rotating panel design such that once an individual is picked for a monthly survey that individual is surveyed again for 5 further months, and every month 1/6 of the individuals surveyed has not been surveyed the previous month. Each monthly survey has around 100,000 observations. The Labour Force Survey is conducted by Statistics Canada and covers all of Canada as its population. In our study, we focus solely on individuals surveyed in Ontario due to the sheer size of the total data set putting pressures on our computational power. When we form our data set we do not take advantage of the unique rotating panel structure of the survey data. We obtain an annual time series of cross section data set from the raw Labour Force Survey data by only taking observations from the January monthly survey of each year from 1993 to 2013. Because each panel is only surveyed for six months, we could in theory take two monthly data sets from each year and be sure that no individual appears in our data set twice. However, there might be concerns that there could be a systematic difference between data taken in January and July due to unknown seasonal factors. To avoid this concern we simply take data only from the January monthly survey of each year. Another reason we only take January monthly surveys and in fact only focus on Ontario is that because quantile regression is estimated through linear programming numerically, not through closed form equations as in OLS, running quantile regression on the whole data set takes literally dozens of hours to compute. Therefore, for seasonality and computational concerns, we focus on the returns to education in Ontario from 2003-2013.

The Labour Force Survey measures 79 variables on various demographic, labour, and education characteristics of each individual surveyed. Our key variables of interest are variables that measure wage and education. Wage is measured hourly and includes tips and commissions. It is the dependent variable for all of our regressions and it is transformed by taking the natural logarithm of each value. This is done because it follows the original theoretical model's specification, and because natural log of wages yields a close approximation of the normal distribution.In Figure 1 we show the histograms of log hourly wage for different levels of education and for our data set as a whole. It shows that the log of hourly wage is approximately normal. In Figure 2 we combine the histograms of men and women's log hourly wage. While both histograms are similar in shape, the men's histogram is slightly centered to the right of the women's histogram. This is consistent with the fact that in general men earn more than women.

Education is measured for both the respondent of the survey and the respondent's spouse. For both, education is measured by asking for the highest degree, certificate or diploma the respondent has obtained and matching the answer to one of seven categories from: 0 to 8 years, some secondary, grade 11-13 (high school) graduated, some post-secondary, post-secondary diploma, university bachelor degree, and university graduate degree. The respondent chooses which of the above

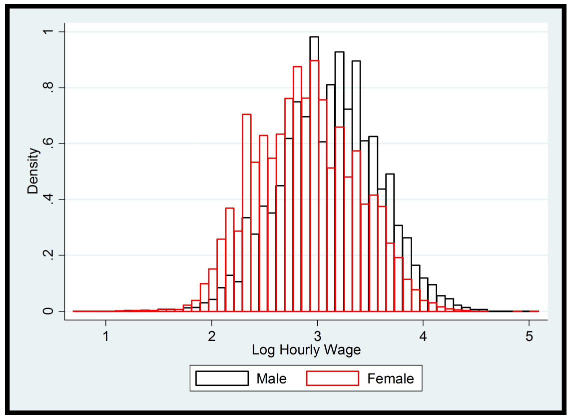

Figure 1: Genders and Wage Disparity

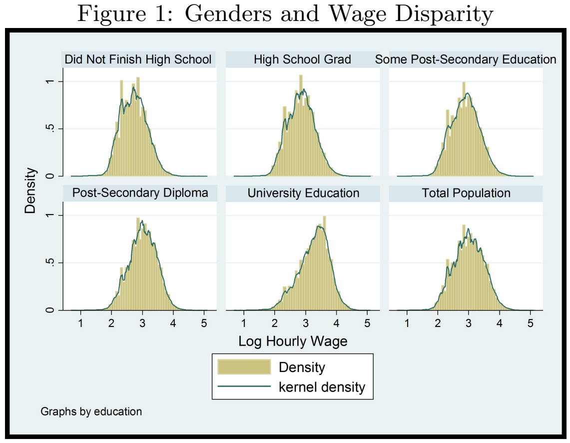

Figure 2: Education Levels and Wage Disparity

categories best fits his educational attainment. The respondent is also asked what her spouse's highest educational attainment is if applicable. Education of spouse is measured slightly differently through 6 categories: the difference is where for spouse's education, there is only one category for university degree whereas the respondent can respond either university bachelor or graduate degree. Using these two variables that measure the respondent's and respondent's spouse's highest educational attainment, we construct a set of four dummies for the respondent's and respondent's spouse's education respectively. The four dummies represent university, post-secondary diploma, some post-secondary, and high school graduate respectively. They equal "1" if the highest educational attainment is university (bachelor or graduate), post-secondary, some post-secondary, or high school graduate is the highest educational attainment respectively. The key explanatory variables of interest are these four dummy variables representing an individual's highest educational attainment, and our instrument is the set of four dummies representing the respondent's spouse's highest educational attainment.

Note that our dummy variable representation of education is different from most of the literature's use of a discrete variable representing years of education. There has been some debate about whether the use of dummy variables for certain levels of education such as university improves the fit of OLS regressions of the Mincer equation and whether this is evidence for "sheepskin" effect

(Lemieux, 2006). We are not directly contributing to this debate but driven by necessity due to the nature of our data and as long as our assumptions for our IV estimation hold, our dummy variables will uncover the causal effect of various levels of education on wages.

3.2. ***Econometric Model***

3.2.1. ***Identification and the Theoretical Model of Schooling Decision***

A standard ordinary least squares (OLS) regression of wages on pertinent and relevant explanatory variables including years or levels of education will fail to identify the true causal effect of education on wages due to the violation of a key assumption for OLS estimation. It is likely that some significant part of a person's wage is explained by that person's ability. By ability we mean some innate set of characteristics of a person that affects the productivity of that person at her job. However, there is no agreement on what exactly the set of characteristics is that affect one's productivity or how to measure it. Therefore, since we have a variable that is omitted from the right-hand side of the regression equation but likely to be key in explaining wages and is correlated with other explanatory variables in our regression equation (such as education), if we estimate a standard OLS equation of wages on a set of explanatory variables such as education and experience, we will have omitted variable bias that biases our OLS regression estimates of the effect of education and experience from their true causal values, leading us to fraudulent information.

There are two ways to address this issue. The first way is to try measure the omitted variables such as ability in an accurate way. Some studies have estimated OLS regressions using various proxies for the missing variable ability with measures like IQ. However, there is much debate on whether measures like IQ scores accurately capture the set of characteristics that are most relevant to determining one's productivity, and how reliable the results of studies that rely on these proxies are (Card, 1999). The second method is to use alternative econometric estimation techniques such as instrumental variables (IV) or fixed effects (FE) estimation which, if additional assumptions are satisfied, may solve the omitted variable bias problem and estimate the true causal effects of education.

We focus on the IV estimation method. The idea of IV estimation is to find a variable or a set of variables that is correlated with the endogenous variables in our regression (right-hand side explanatory variables that may be correlated with omitted variables or are caused by the dependent variable we are trying to explain) but uncorrelated with the stochastic error term including the potentially omitted variables such as ability. The variable or set of variables are called instruments. Through the use of these instruments, if the assumption that our instruments are not correlated with omitted variables holds, we can obtain the causal effects of our endogenous variables on the dependent variable. The assumption that instruments are not correlated with omitted variables is critical, and referred to as the exclusion restriction (Wooldridge, 2012). Note that this assumption cannot be tested directly because to do so we need to somehow measure the omitted variables to test if they are correlated with our instruments, but the reason that they are omitted is that we cannot find an accurate measure for them. We have to rely on indirect evidence and economic theory to support the claim that an instrument satisfies the exogeneity condition.

There is an extensive literature on the use of IV estimation to study the causal effect of education on wages. Previous examples of instruments include distance of residence from campus and tuition of state colleges (Kane& Rouse, 1993), smoking behavior (Dickson, 2012), and education of parents (Card, 1993). Our proposed instrument is spouse's highest attained education level. In our data set our instruments are significantly correlated to the endogenous variables checked by F-tests on first-stage regressions. In addition, we believe that our instrument improves on the previously used instrument of education level of parents in terms of satisfying the exogeneity condition.

Here we briefly refer to a theoretical model of investment in education based on Card (1994), and illustrated in Arias, Hallock,& Sosa-Escudero (2001), which will help highlight some key issues in our study. This model describes the optimal schooling decision of individuals through the concept of human capital. Human capital refers to the combination of skill and knowledge levels that ultimately determine the productivity of an individual as a worker. Human capital is composed of education and ability. Ability is fixed at some early point in an individual's lifetime and human capital can only be augmented through further education. Based on this model, individuals choose the level of schooling that will set equal the marginal benefit and marginal cost to education. The marginal benefit of education is determined by a homogeneous effect of education as well as the effect of ability on the rate of human capital accumulation through education. If this latter effect is positive, that a higher ability allows higher gains of human capital through education, then education and ability are said to be complements. The marginal cost of education is determined by factors such as wealth and taste for education. In this setting, under some model assumptions it can be derived that if education and ability are complements then on average a person with more ability will also be more educated, since on average the marginal benefit of a person with higher ability will be higher than a person with low ability.

According to this result, our instrument of spouse's education for a person's education would violate the exogeneity assumption since in our data individuals and their spouses have similar education levels, and on average they will have similar ability. However, we believe that the assumptions required to derive this result are too strong to hold in real life and that the verisimilitude of the theoretical result is suspect since the theoretical model does not account for factors that dominate ability's effect on schooling decisions. These assumptions include each individual objectively knowing: her ability and its future effect on human capital accumulation through education, and the future compensation of their human capital in the job market. When we add uncertainty for each individual regarding their ability and other factors, it is no longer the case in general that higher educated people must on average have higher ability than lower educated people. In addition, the model also assumes perfect credit markets and household wealth or lack of is no factor in schooling decisions. Empirical evidence suggests this is an unrealistic assumption and that the capability of an individual's household to fund her schooling is a significant determinant of schooling decisions (Filmer, 2005). Therefore, given these considerations there is little reason to suspect that a spouse's education level is correlated to an individual's ability. It must be noted that for spouse's education to satisfy the exogeneity condition we need an additional assumption that an individual's ability is not affected by household wealth, taste, or other factors that determine schooling decisions. If this assumption did not hold then spouse's education would be correlated with an individual's ability. For example, if "ability" includes how big an individual's network is and how adept they are at getting good jobs through the use of her network, and if this network is on average bigger for people from wealthier households, then our instrument would no longer be valid. We will refer to this assumption again when we interpret the results of our quantile regressions.

Finally, as mentioned above, we believe our instrument improves upon the previously suggested instrument of parents' education in terms of satisfying the exogeneity condition. This is because parents' education levels may influence a person's ability in her childhood. One example of such an effect is if parents with higher education are better able to care for a child with better knowledge about nutrition and health so there is less risk of under-development due to malnutrition or disease. A less extreme example is if parents with higher education are better able to instill good habits or values such as time management and high work ethic. With our instrument there is significantly less concern of a direct causal link since most married people meet their spouses after coming into adulthood with their physiology and ability completely matured, so it is much less likely that a spouse's education level affects a person's ability.

There is empirical evidence that demographic and family characteristics such as race and sex affect returns to education (Card, 1999) and further that even when controlling for these characteristics, there still persists significant heterogeneity of returns to education on wages (Arias et al, 2002). The heterogeneity of returns to education could be due to the heterogeneity of ability for people with same level of education, the job market conditions for graduates of various education levels, or over-supply of highly qualified graduates due to asymmetric information. We apply quantile regression analysis to parsimoniously characterize the heterogeneity of returns to education and study what may be causing this heterogeneity. Quantile regression is similar to OLS regression in that you are drawing a line of best fit through the data, but in quantile regression there is the added constraint that for a specified percentile, the percentage of points that lie under the quantile line of best fit is equal to that percentile. Quantile regression allows the study of how the effects of the explanatory variables differ for different percentiles of the dependent variable. Just like OLS, quantile regression estimates of explanatory variable coefficients are biased from the true causal values if there is omitted variable bias, and IV estimation method allows us to obtain true causal estimates given the usual IV assumptions as well as new assumptions specific to quantile-IV estimation are met.

If returns to education are heterogeneous, there are many potential benefits for policy makers and prospective students to know the exact distribution of this heterogeneity as well as understand the process that generates it. One specific example is where increasing the level of education is proposed by some as the solution to wage inequality (Ashenfelter& Rouse, 1998). If returns to education are heterogeneous then increasing education may exacerbate wage inequality. Furthermore, efforts to understand the process generating heterogeneity of returns to education may shed further light on the workings of the process in which education increases wage. Non risk-neutral students will benefit from knowing the distribution of wage outcomes after education. For these reasons and more, studying not only the average effects of education on wage but the distribution of effects is important and we apply quantile regressions to study it.

3.2.2. ***Estimation of Returns to Education***

We estimate five different models of returns to education based on various estimation technique mentioned above:

- a) OLS regression of the classical Mincer equation
- b) OLS regression of fully-specified model
- c) IV regression with probit first-stage
- d) Quantile regression
- e) Quantile-IV regression with probit first-stage

Regressions a) to c) are points of reference. Previous studies of returns to education on wage have primarily focused on OLS and IV regressions. Our main regressions of interest are the results of the quantile regressions, especially the results of the quantile-IV regression with spouse's education dummies as instruments. The quantile-IV regressions will allow us to analyze heterogeneity in the causal effect of education on wage.

$$\log Wage_i = \beta_0 + \beta_1 Education_i + \beta_2 Experience_i + \beta_3 Experience_i^2 + u_i \tag{1}$$

$$\begin{aligned} \log Wage_{it} = \beta_0 + \beta_1 University_{it} + \beta_2 Diploma_{it} + \beta_3 SomePostSec_{it} + \beta_4 HighSchool_{it} \\ + \beta_5 Experience_{it} + \beta_6 Experience_{it}^2 + u_{it} \end{aligned} \tag{2}$$

The Mincer OLS regression is estimated mainly for reasons of historical perspective. Variables for education, experience, and experience squared appear as the only right-hand side explanatory variables. Many of the early studies in returns to education used the Mincer equation and it is interesting to note how our full specification estimates differs from Mincer estimates. All of the following regressions based on OLS, IV, quantile, and quantile-IV methods estimate our fully specified model , which contain other variables besides education and experience we believe determine an individual's hourly wage. These variables include a dummy variable equaling "1" if person is male and "0" if female, a dummy indicating trade union membership, a set of three dummies indicating size of company ("More than 500 employees at the location", "100 to 500", "20 to 99", and "less than 20"), and a set of dummies indicating industry of occupation such as agriculture, utilities, construction, and educational service.Although we do not have information on the types of degrees such as engineering or medicine that individuals with post-secondary or university education get, we believe the set of industry dummies alleviate any omitted variable bias arising from lack of information on the type of degrees each individual obtains. We also include a set of time dummies for the year of the survey.

3.2.3. ***Quantile-IV Regression***

After estimating OLS and IV estimates of our fully-specified model, we turn to our main regressions of interest the quantile and quantile-IV regressions analyzing heterogeneity of returns to education. Quantile regression is in many respects similar to OLS regression. Where OLS regression draws a line of best fit through the data such that the sum of squared residuals is minimized, quantile regression draws a line of best fit through a specified quantile (between 0 and 1) such that the sum of absolute value of residuals is minimized. In other words, for a specified quantile, say 0.1, quantile regression line minimizes the absolute value of residuals such that 10% of the data lie under the line of best fit. With OLS methods one can only test for the existence of heterogeneity in the conditional distribution of a dependent variable through heteroskedasticity tests (Wooldridge, 2012), in our case the conditional distribution of wage controlling for education and other explanatory variables; there is no way to characterize the heterogeneity except testing for the severity of the heterogeneity. Quantile regression characterizes the heterogeneity in a simple and intuitive way by fitting lines of best fit through the conditional distributions for a specified quantile. For example, if a researcher wanted to know whether the returns to education for the bottom 10% of the data were different from returns to education for individuals in the median, she could run quantile regressions for quantiles 0.1 and 0.5 and compare the coefficient estimates between the two quantile regressions.

$$\begin{aligned}\log Wage_{it} = \beta_0 + \beta_1 University_{it} + \beta_2 Diploma_{it} + \beta_3 SomePostSec_{it} + \beta_4 HighSchool_{it} \\ +\beta_5 Experience_{it} + \beta_6 Experience_{it}^2 + \beta_7 Sex_{it} + \beta_8 Union_{it} + \beta_9 HugeFirm_{it} \\ +\beta_{10} LargeFirm_{it} + \beta_{11} MediumFirm_{it} + \beta_{12}\sum_{i=1}^{17} industries_i + \beta_{13}\sum_{t=1}^{9} years_t + u_{it}\end{aligned} \tag{3}$$

Just as with OLS regression, omitting relevant explanatory variables causes omitted variable bias in quantile regression. Although this omitted variable bias can be addressed for quantile regression through IV methods as in OLS, for the quantile regression case it requires additional assumptions on top of the standard IV assumptions for OLS (Chernozhukov& Hansen, 2006). The key additional assumption is called rank invariance, which requires that conditional on other explanatory variables, an individual's rank in the conditional distribution of the dependent variable is the same regardless of which treatment the individual receives. In our application this means for example if an individual received a university degree and earned a wage higher than 90% of others who also received a university degree, then if that same individual only had a high school education he would still earn a wage higher than 90% of others with only a high school degree. Note we can never empirically test the validity of this assumption for any data set since we can never simultaneously observe an individual's wage under two different education levels holding all other variables constant. This assumption seems unreasonable in our case but this assumption can be replaced by another assumption called rank similarity that is less stringent. Rank similarity states that the rank of an individual in the conditional distribution for different treatment values can only vary stochastically with mean zero. We believe that this is a reasonable assumption for our data set. Therefore, given the standard IV assumptions and the rank similarity assumption holds, we can estimate the causal effect of education at different quantiles of wage and analyze the heterogeneity of returns to education.

Given these assumptions Chernozhukov and Hansen (2006) detail a process to estimate quantile-IV regressions in an efficient way. This process is an analog of two-stage least squares (2SLS) estimation of IV. In our study, however, we do not follow this process but approximate it by estimating four first-stage regressions for our four dummy instruments of spouse's education with probit and using the fitted values from the first stage regressions to replace respondents education dummies in the second-stage structural regression. This method should yield coefficient estimates very similar to Chernozhukov and Hansen estimates, and although there is concern about the standard errors being too small, a recent paper has documented that this 2SLS analog for quantile-IV yields estimates nearly identical to Chernozhukov and Hansen estimation (Autor, Houseman,& Kerr, 2012).

4. Results and Analysis

The results of the Mincer equation, OLS, and IV of fully-specified model are reported first. The coefficient estimates of the education dummies indicate the approximate average percentage change in hourly wage if an individual has higher education than the baseline education of not finishing high school. IV regression estimates of the effects of education on wages are the highest, followed by Mincer estimates and then OLS estimates. IV estimates that the average causal effect of university education on hourly wage compared to not finishing high school is an approximately 80.5% increase in hourly wage. IV estimates of the average causal effect of diploma, some post-secondary, and high school education compared to not finishing high school are approximately 48.7%, 46.1%, and 47.1% respectively. Note that for IV estimates there is no significant difference in the average causal effects of diploma, some post-secondary education, and high school education with 95% confidence level. This is an unintuitive result which states that there is almost no net benefit on hourly wage for furthering education from high school to post-secondary diploma programs, but it would be consistent with the story that there is asymmetric information for prospective students where they are not fully aware of the costs and benefits of furthering education. It could also be that while on average there is no significant difference, for different quantiles there is a significant difference in how high school and diploma education affect hourly wage.

Coefficient estimates of all other explanatory variables have signs that are consistent with intuition and past empirical evidence, and are all significant with 95% confidence level. The sex dummy has a positive sign, indicating men earn more than women. Tenure is positive and Tenure squared is negative, which is consistent with the Mincer literature. Firm Size dummies are positive and increasing with the size of the firm. Union dummy is positive, indicating membership in a union on average increases hourly wage. Industry dummies are all positive and significant, with industries utilities, construction, professional/scientific/technical, health care/social assistance, and public administration having the highest coefficient estimates. We selected the accommodation/food services industry as our baseline case since it has the lowest average hourly wage. Industries with the lowest coefficient estimates include agriculture, retail trade, and management/administration/other support. Finally, year dummies with baseline case the year 2003 are all positive and increasing with years. This is consistent with GDP growth and inflation.

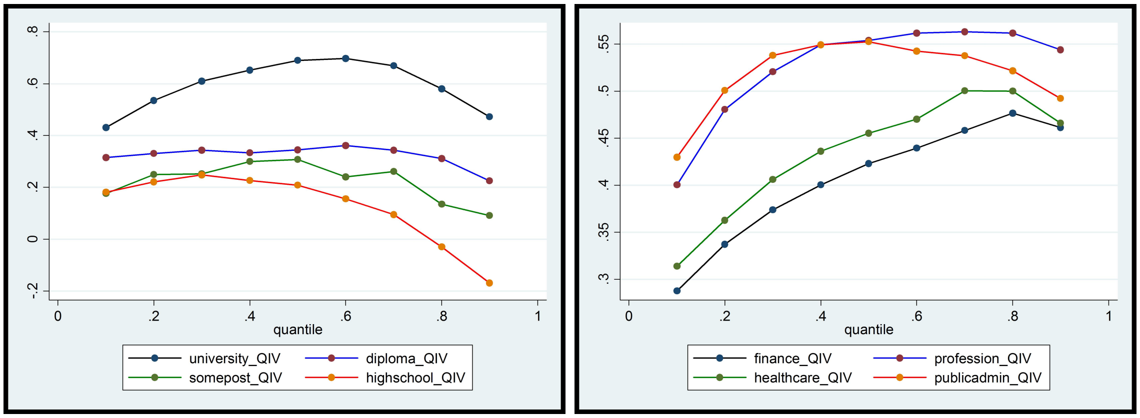

Figure 3: Coefficients of Variables for Different Quantiles

Now we turn to the quantile and quantile-IV results. We estimated quantile and quantile-IV regressions with equidistant deciles from 10% to 90%. Quantile-IV estimates of the causal effect of education on hourly wage are higher than quantile estimates except for the the 80th and 90th percentile where the coefficients are similar. This would be consistent with the general finding that IV estimates of causal effects are higher than OLS estimates of returns to education. At the lower deciles quantile-IV estimates of some post-secondary and high school dummies are similar but for higher deciles some post-secondary coefficients are higher than high school coefficients before becoming similar again for the two highest deciles. There seems to be significant differences in returns to high school and some post-secondary education in the middle deciles but not at the lowest and highest deciles. For university education, the coefficient estimate steadily increases from 0.43058 in 10th percentile to as high as 0.69805 in the 60th percentile. Thereafter the coefficient decreases until it is 0.4725945 in the 90th percentile (Refer to Figure 3). It is intuitive that the effect of university education on wages is increasing in quantiles but it is puzzling to see that from the 80th to 90th percentile it decreases from 0.5802347 to 0.4725945. This could be due to a few reasons. First, this could be due to the effect of taking the natural logarithm of hourly wages instead of taking the raw data as the dependent variable. Taking the natural logarithm of hourly wage will distort the original distribution of hourly wage, especially on the far right-hand side of the distribution or in our case for the highest wage-earners. This distortion of the original distribution may be what is causing this puzzling result. Second, it may be that for low-median to high income it makes sense that there is tangible and increasing benefit of university education, for the highest 10% hourly wages education has a far smaller impact than other intangibles such as luck or extraordinary circumstance. This would be consistent with our results that as you go from low to high deciles the benefit of university education increases but for the highest decile university education does not matter so much as the other intangible, unexplainable factors such as luck. What is for certain is that there is a statistically significant heterogeneity of returns to education at 95% confidence level (Refer to Table 3).

Table 1: Mincer, OLS and IV Regression Results

	Mincer		OLS		IV	
Explanatory Variables	Est. Values	S.E.	Est. Values	S.E.	Est. Values	S.E.
$University$	0.551	(0.005)	0.427	(0.005)	0.805	(0.026)
$Diploma$	0.287	(0.005)	0.218	(0.005)	0.487	(0.025)
$SomePostSec$	0.180	(0.007)	0.153	(0.006)	0.461	(0.054)
$HighSchool$	0.104	(0.005)	0.098	(0.005)	0.471	(0.041)
$Experience$	0.003	(0.000)	0.000	(0.000)	0.003	(0.000)
$Experience^2$	−0.000	(0.000)	−0.000	(0.000)	−0.000	(0.000)
Sex			0.210	(0.003)	0.211	(0.003)
$Agriculture$			0.134	(0.016)	0.149	(0.017)
$Forestry, fishing$			0.254	(0.006)	0.257	(0.007)
$Utilities$			0.601	(0.012)	0.560	(0.013)
$Construction$			0.490	(0.008)	0.491	(0.009)
$ManufacturerDurables$			0.400	(0.007)	0.386	(0.008)
$ManufactNondurables$			0.337	(0.008)	0.327	(0.008)
$WholesaleTrade$			0.378	(0.009)	0.352	(0.009)
$RetailTrade$			0.104	(0.007)	0.879	(0.008)
$Transport, warehousing$			0.282	(0.008)	0.279	(0.009)
$Finance, insurance$			0.417	(0.008)	0.372	(0.009)
$Professional, scientific$			0.557	(0.008)	0.498	(0.009)
$Management$			0.148	(0.009)	0.135	(0.010)
$EducationalServices$			0.461	(0.008)	0.387	(0.009)
$HealthCare$			0.460	(0.007)	0.426	(0.009)
$Info, culture$			0.368	(0.009)	0.328	(0.010)
$OtherServices$			0.274	(0.009)	0.253	(0.010)
$PublicAdministration$			0.539	(0.008)	0.488	(0.009)

Source: Labour Force Survey, Statistics Canada

There is heterogeneity for some other explanatory variables. The coefficient of the sex dummy is increasing in deciles and the difference between deciles is significant. For example, the difference between the increase in hourly wages due to being a male is approximately 5% higher at the 90th percentile than the 10th percentile at 95% confidence level. Tenure is increasing and then decreasing as deciles increase. The firm size dummies are increasing in deciles. The increase in wages from being in a firm with more than 100 workers is approximately 8.2% higher at the 90th percentile than the 10th percentile. This is an intuitive result. The union dummy is decreasing in deciles. Finally, for some industries the coefficient estimates increase significantly as deciles increase. This is true for professional/scientific/technical, management/admin/other support, and health care/social assistance industries (Refer to Figure 3). This makes sense as we expect that for those individuals in the high percentiles of hourly wage have occupations that typically pay wages such as doctors, lawyers, etc So we expect the effect of being in these industries to increase as the deciles of log hourly wage increase.

Table 2: Quantile Regression and Quantile IV Regression Results

	Sex		*University*		*Diploma*		*SomePostSec*		*HighSchool*		*Experience*		*Experience*^{2}	
Quantiles	Quantile	QIV	Quantile	QIV	Quantile	QIV	Quantile	QIV	Quantile	QIV	Quantile	QIV	Quantile	QIV
q10	0.148	0.145	0.247	0.431	0.156	0.315	0.0990	0.176	0.079	0.181	0.003	0.003	−0.000	−0.000
	(0.003)	(0.005)	(0.006)	(0.022)	(0.004)	(0.262)	(0.010)	(0.067)	(0.005)	(0.040)	(0.000)	(0.000)	(0.000)	(0.000)
q20	0.175	0.172	0.321	0.535	0.170	0.331	0.104	0.250	0.078	0.221	0.003	0.003	−0.000	−0.000
	(0.004)	(0.004)	(0.007)	(0.013)	(0.005)	(0.018)	(0.007)	(0.043)	(0.004)	(0.025)	(0.000)	(0.000)	(0.000)	(0.000)
q30	0.195	0.194	0.383	0.610	0.187	0.344	0.115	0.252	0.079	0.248	0.003	0.003	−0.000	−0.000
	(0.003)	(0.004)	(0.006)	(0.012)	(0.004)	(0.015)	(0.007)	(0.041)	(0.004)	(0.020	(0.000)	(0.000)	(0.000)	(0.000)
q40	0.208	0.211	0.430	0.653	0.198	0.333	0.124	0.299	0.080	0.226	0.003	0.003	−0.000	−0.000
	(0.003)	(0.004)	(0.006)	(0.011)	(0.004)	(0.016)	(0.006)	(0.035)	(0.005)	(0.022)	(0.000)	(0.000)	(0.000)	(0.000)
q50	0.216	0.219	0.471	0.689	0.215	0.345	0.147	0.308	0.088	0.208	0.003	0.003	−0.000	−0.000
	(0.003)	(0.005)	(0.006)	(0.018)	(0.005)	(0.019)	(0.006)	(0.060)	(0.006)	(0.023)	(0.000)	(0.000)	(0.000)	(0.000)
q60	0.224	0.219	0.504	0.698	0.235	0.361	0.160	0.240	0.096	0.156	0.003	0.003	−0.000	−0.000
	(0.003)	(0.005)	(0.008)	(0.021)	(0.007)	(0.024)	(0.006)	(0.052)	(0.006)	(0.027)	(0.000)	(0.000)	(0.000)	(0.000)
q70	0.227	0.217	0.517	0.670	0.243	0.343	0.169	0.261	0.101	0.095	0.002	0.002	−0.000	−0.000
	(0.003)	(0.005)	(0.008)	(0.022)	(0.007)	(0.027)	(0.009)	(0.056)	(0.007)	(0.033)	(0.000)	(0.000)	(0.000)	(0.000)
q80	0.224	0.208	0.515	0.580	0.253	0.312	0.183	0.136	0.104	−0.030	0.002	0.002	−0.000	−0.000
	(0.003)	(0.004)	(0.009)	(0.025)	(0.007)	(0.028)	(0.007)	(0.058)	(0.009)	(0.041)	(0.006)	(0.000)	(0.000)	(0.000)
q90	0.228	0.204	0.506	0.473	0.258	0.225	0.213	0.091	0.112	−0.169	0.002	0.002	−0.000	−0.000
	(0.005)	(0.006)	(0.011)	(0.021)	(0.007)	(0.037)	(0.010)	(0.071)	(0.007)	(0.037)	(0.000)	(0.000)	(0.000)	(0.000)

Source: Labour Force Survey, Statistics Canada

Table 3: Inter-Quantile Results

	University		*Diploma*		*SomePostSec*		*HighSchool*	
Inter-Quantiles	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
q20-q10	0.104	0.000	0.016	0.320	0.074	0.112	0.039	0.078
q30-q10	0.180	0.000	0.029	0.100	0.076	0.241	0.066	0.008
q40-q10	0.222	0.000	0.018	0.467	0.123	0.094	0.044	0.277
q50-q10	0.259	0.000	0.030	0.255	0.131	0.126	0.026	0.439
q60-q10	0.267	0.000	0.047	0.116	0.064	0.438	−0.025	0.509
q70-q10	0.239	0.000	0.029	0.280	0.084	0.397	−0.086	0.020
q80-q10	0.150	0.000	−0.003	0.920	−0.041	0.600	−0.211	0.000
q90-q10	0.042	0.265	−0.090	0.030	−0.085	0.406	−0.351	0.000
q40-q30	0.042	0.000	−0.011	0.316	0.047	0.167	−0.022	0.183
q50-q30	0.079	0.000	0.001	0.968	0.055	0.324	−0.040	0.040
q60-q30	0.088	0.000	0.017	0.348	−0.012	0.819	−0.092	0.000
q70-q30	0.060	0.001	−0.001	0.981	0.008	0.901	−0.153	0.000
q80-q30	−0.030	0.229	−0.032	0.223	−0.117	0.195	−0.278	0.000
q90-q30	−0.138	0.000	−0.119	0.000	−0.161	0.090	−0.417	0.000
q60-q50	0.009	0.378	0.017	0.048	−0.068	0.040	−0.052	0.002
q70-q50	−0.019	0.241	−0.001	0.945	−0.047	0.516	−0.113	0.000
q80-q50	−0.109	0.000	−0.033	0.064	−0.172	0.005	−0.237	0.000
q90-q50	−0.217	0.000	−0.120	0.000	−0.216	0.001	−0.377	0.000
q80-q70	−0.090	0.000	−0.032	0.078	−0.125	0.023	−0.125	0.000
q90-q70	−0.197	0.000	−0.119	0.000	−0.169	0.016	−0.264	0.000
q90-q80	−0.108	0.000	−0.087	0.000	−0.044	0.370	−0.139	0.000

Source: Labour Force Survey, Statistics Canada

There are three ways of interpreting the heterogeneity in the causal effect of returns to education. The first explanation is that there are simply too many graduates of high education that leads to many students taking jobs they are over-qualified and their education is not properly rewarded with what should be higher wages, since there is simply not enough jobs that require their level of education. Unfortunately with our given data set we have little means of testing this explanation and we refer the reader to Green and Zhus (2010) study, which finds that there is some evidence of asymmetric information. The second explanation is that the evolving labour market causes heterogeneity in returns to education by heterogeneity in the wage graduates of similar education earn in different times and industries. In our regressions we added industry and time dummies to control for this source of heterogeneity but heterogeneity persists. Therefore we believe that the third potential source of heterogeneity, that education and ability are complements, in combination with the potential over-qualification story in Green and Zhu (2010), is the best explanation for the heterogeneity of returns to education. If education and ability are complements, then those with higher ability would have higher returns to education since they earn higher wages post-graduation and a heterogeneous distribution of ability in the population would cause heterogeneous returns to education no matter what level of education for the population. This is indeed what we see in our results. For all levels of education we have evidence of heterogeneity in returns to education. Therefore we believe that the heterogeneous distribution of ability in the population is causing heterogeneous returns to education.

5. Conclusion

We have estimated quantile-IV regressions of log hourly wage on a set of explanatory variables including education and found that there is significant heterogeneity in the effect of education on hourly wage in Canada using the Labour Force Survey micro data from 2003-2013. Returns to education are higher for individuals with higher hourly wages than those with low wages. We also find that what industry an individual works in matters in determining her hourly wage, and that for some industries the effect of working in that industry on her hourly wage increases or decreases depending on whether she is high or low wage earner. Finally, we find that sex is significant in determining wage and that there is some the benefit of being male increases for higher wages. Technically speaking our results only apply to married individuals in Ontario but we do not suspect any reasons why the basic pattern of our results could not represent the population of Canada. Our results indicate that because there is significant heterogeneity in returns to education at all levels of education wage inequality cannot be reduced through education. In order to reduce wage inequality the heterogeneity in the distribution of ability must be reduced. In order to do so, further studies must explore the nature of ability and its interaction with education. In terms of improving estimates of heterogeneity in the causal effect of education on wages, future studies could try to find instruments that are more likely to satisfy the IV conditions than spouse's education and also find a quantile-IV estimation technique that does not require additional assumptions such as rank invariance or rank similarity.

Reference

Arias, O., Hallock, K. F.,& Sosa-Escudero, W. (2002). Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. *Economic Applications of Quantile Regression*, 7-40.

Ashenfelter, O.,& Rouse, C. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *The Quarterly Journal of Economics*,113(1), 253-284.

Autor, D. H., Houseman, S. N.,& Kerr, S. P. (2012). The effect of work first job placements on the distribution of earnings: An instrumental variable quantile regression approach (No. w17972). *National Bureau of Economic Research.*

Bennell, P. (1996). Rates of return to education: Does the conventional pattern prevail in sub-Saharan Africa?. *World Development*, 24(1), 183-199.

Biagetti, M.,& Scicchitano, S. (2011). Education and wage inequality in Europe. *Economics Bulletin*, 31(3), 2620-2628.

Bingley, P., Zhu, Y.,& Walker, I. (2005). Education, Work and Wages in the UK. *German Economic Review*, 6(3), 395-414.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling (No. w4483). *National Bureau of Economic Research.*

Card, D. (1994). Earnings, schooling, and ability revisited (No. w4832). *National Bureau of Economic Research.*

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3, 1801-1863.

Chernozhukov, V.,& Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1), 245-261.

Chernozhukov, V.,& Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*,132(2), 491-525.

Chernozhukov, V.,& Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1), 379-398.

Dickson, M. (2012). The Causal Effect of Education on Wages Revisited.*Oxford Bulletin of Economics and Statistics*, 75, 4, 477-498.

Eggleton, A. (2013). *Income inequality is hurting social inclusion in Canada* Retrieved from http://www.theglobeandmail.com/report-on-business/economy/economy-lab/inequality-is-hurting-social-inclusion-in-canada/article15449465/.

Filmer, D. (2005). Gender and wealth disparities in schooling: Evidence from 44 countries. *International Journal of Educational Research*, 43(6), 351-369.

Fitzenberger, B., Hujer, R., MaCurdy, T. E.,& Schnabel, R. (2002). Testing for uniform wage trends in West-Germany: A cohort analysis using quantile regressions for censored data. *Economic Applications of Quantile Regression*,41-86.

Frlich, M.,& Melly, B. (2010). Estimation of quantile treatment effects with Stata. *Stata Journal*, 10(3), 423.

Galvao Jr, A. F. (2011). Quantile regression for dynamic panel data with fixed effects. *Journal of Econometrics*, 164(1), 142-157.

Green, F.,& Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62(4), 740-763.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go?. *Economics of education review*, 19(2), 131-147.

Heywood, J. S. (1994). How widespread are sheepskin returns to education in the US?. *Economics of Education Review*, 13(3), 227-234.

Kane, T. J.,& Rouse, C. E. (1993). Labor market returns to two-and four-year colleges: is a credit a credit and do degrees matter? (No. w4268). *National Bureau of Economic Research.*

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89.

Koenker, R. (2005). *Quantile regression (No. 38).* Cambridge university press.

Lemieux, T. (2006). *The Mincer equation thirty years after schooling, experience, and earnings* (pp. 127-145). Springer US.

Lumpe, C.,& Weigert, B. (2010). Immigration, Education and Wage Inequality.*Journal of Economic Integration*, 25(3), 592-612.

Martins, P. S.,& Pereira, P. T. (2004). Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics*, 11(3), 355-371.

Psacharopoulos, G.,& Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics*, 12(2), 111-134.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World development*, 22(9), 1325-1343.

Appendix

```
************************Mincer
reg lnhrlyear university diploma somepost high tenure tenure2 year2004 year2005 year2006 year2007
year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 & survmnth==1
******************************OLS
reg lnhrlyear sex  university diploma somepost high tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 &
survmnth==1
*********************************IV
///First Stage
probit university sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
test sbachelor sdiploma ssomepost shigh
probit diploma sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
test sbachelor sdiploma ssomepost shigh
probit somepost sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
test sbachelor sdiploma ssomepost shigh
probit high sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
test sbachelor sdiploma ssomepost shigh
ivreg2 lnhrlyear sex  tenure tenure2 uniond huge large medium industry01 industry02 industry03
industry04 industry05 industry06 industry07 industry08 industry09 industry10 industry11 industry12
industry13 industry14 industry15 industry17 industry18 year2004 year2005 year2006 year2007 year2008
year2009 year2010 year2011 year2012 year2013 (university diploma somepost high = university_15
diploma_15 somepost_15 high_15)
///First Stage
probit university sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
```

```
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
predict university_15
test sbachelor sdiploma ssomepost shigh
probit diploma sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
predict diploma_15
test sbachelor sdiploma ssomepost shigh
probit somepost sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
predict somepost_15
test sbachelor sdiploma ssomepost shigh
probit high sex  sbachelor sdiploma ssomepost shigh tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1
predict high_15
test sbachelor sdiploma ssomepost shigh
ivreg2 lnhrlyear sex   tenure tenure2 uniond huge large medium industry01 industry02 industry03
industry04 industry05 industry06 industry07 industry08 industry09 industry10 industry11 industry12
industry13 industry14 industry15 industry17 industry18 year2004 year2005 year2006 year2007 year2008
year2009 year2010 year2011 year2012 year2013 (university diploma somepost high = university_15
diploma_15 somepost_15 high_15)
***************Quantile
sqreg lnhrlyear sex  university diploma somepost high tenure tenure2 uniond huge large medium
industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08 industry09
industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18 year2004
year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if prov==35 ^
survmnth==1, quantile(0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9)
***************Quantile IV
sqreg lnhrlyear sex  university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9)
***************Inter Quantile-IV
iqreg lnhrlyear sex  university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
```

```
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.3)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.4)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.5)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.6)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.7)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.8)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.1 0.9)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.4)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.5)
```

```
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.6)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.7)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.8)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.3 0.9)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.5 0.6)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.5 0.7)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.5 0.8)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
prov==35 ^ survmnth==1, quantile(0.5 0.9)
iqreg lnhrlyear sex   university_15 diploma_15 somepost_15 high_15 tenure tenure2 uniond huge large
medium industry01 industry02 industry03 industry04 industry05 industry06 industry07 industry08
industry09 industry10 industry11 industry12 industry13 industry14 industry15 industry17 industry18
year2004 year2005 year2006 year2007 year2008 year2009 year2010 year2011 year2012 year2013 if
```