Three Essays on Applied Economics

by

Chengyu Si

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
Aug 3, 2019

Keywords: land use, insurance subsidies, wage gap, social inequality

Approved by

Denis A. Nadolnyak, Chair, Professor of Department of Agricultural Economics and Rural Sociology
Valentina Hartarska, Alumni Professor of Agricultural Economics and Rural Sociology
Curtis M. Jolly, Professor of Department of Agricultural Economics and Rural Sociology
Emir Malikov, Assistant Professor of Agricultural Economics and Rural Sociology

Abstract

This dissertation is composed of three chapters examining the wage inequality in developing countries, and the effect of insurance subsidy on land use allocation.

Chapter 1 is about the distributional decomposition of gender wage gap in developing Countries. In this paper I investigate the gender wage gap across 12 developing countries. Based on the traditional Blinder-Oaxaca decomposition, I also consider the distributional effects on the wage gap decomposition. I apply the approaches which are similar to Machado and Mata (2005) and Firpo et al. (2007, 2009) to obtain both aggregate and detailed decomposition. The results show that gender wage gap exists at all wage levels and decreases as wage increase. The number of children contributes to the gender wage gap especially at low wage levels. Opening more mother friendly positions and making education and training more accessible to the low-income groups will be helpful to raise their earnings and reduce the wage inequality.

Chapter 2 investigates the effects of parental education on children's income. In wage equation, education is an important factor that affects personal income. However, education of parents also has influence on children's income, because of the intergenerational effects. In this paper, I investigate the effect of parents' educational level on children's income in 12 developing

countries. I use the maximum educational level of parents as the independent variable and estimate the equation of children's hourly income. As there is omitted ability in the wage equation, I use interruption of schooling as the instrumental variables to identify the educational years. In addition, I applied Heckman Selection Model to fix sample selection bias. The results show that high parents' educational level has positive effect on children's hourly earnings. Policy makers should consider the intergenerational effect to reduce social inequality.

Chapter 3 investigates the effect of crop insurance subsidies on agricultural land use allocation. Since the objective of crop insurance is to help farmers with risk management, the expected profit of crop production under crop insurance might be improved, leading farmers to allocate more land into crop production. In this paper, agricultural land use type is classified by irrigated/unirrigated farmland and cropland/woodland/pasture land. The data contain counties from all continental states. Considering the fractional outcome of land use share, I apply a Multinomial Fractional Logit Model to estimate the effects of insurance subsidies on land use. The results show that insurance subsidies have a significant effect on land use allocation. An increase in insurance subsidies increases farmland share, indicating insurance subsidies could be an efficient tool to adjust agricultural land use allocation.

Acknowledgements

First and foremost, I would express my deepest appreciation to my academic advisor, Dr. Denis Nadolnyak, for his patience, professional advice, and continuous support. Dr. Nadolnyak guided me every step of doing research, enlightened me to find interesting topics and advised me to solve the problems. Besides research, Dr. Nadolnyak always encourage me to do the things that might be helpful in the future. During the past few years, I have learnt a lot from Dr. Nadolnyak, not only doing research, but also rules of conduct.

I would also like to appreciate my committee members Dr. Valentina Hartarska, Dr. Curtis Jolly and Dr. Emir Malikov for their time and support to my research. Dr. Valentina Hartarska always give me very detailed suggestions and some new ideas to improve my research. Dr. Curtis Jolly shared with me his experience in doing research when I was puzzled, which encouraged me to finish the study and taught me how to write paper. Dr. Emir Malikov always help me with efficient and direct solutions when I was trapped in some academic and technical troubles.

I would like to extend my sincere thank to Dr. Robert Weigel, who supported my assistantship during the first year of Ph.D. and provided me a good opportunity to be a Teaching Assistant for Auburn Global Econ classes.

I would like to sincerely thank the stuffs in our department, especially Ms. Ashley Pangle, Ms. Ann Gulatte for their help during my study. I would also thank my officemate Rui Chen, Meijuan Wang, Wei Ge, Mo Zhou, Jingfang Zhang for helping me with school stuffs and making my school life more wonderful.

I am thankful to my friends Yanfei Sun, Sheng Dong, Guyu Qin, Songyan Yu for their accompanying and making my daily life in Auburn memorable. I would also like to thank my boyfriend Luye He. It was his company and understanding that brought me a joyful life.

Finally, I would express my deepest gratitude to my family in China. They always give me endless love and support which encourage me keep moving and finish my Ph.D. study.

Table of Contents

List of Figures

List of Tables

Chapter 1 Distributional Decomposition of Gender Wage Gap in Developing Countries

1.1. Introduction

Gender wage gap exists widely all over the world. Most of the studies attribute the gender wage gap to the difference in educational level, experience, industry, and gender discrimination. However, household status is also an important factor causing gender wage gap. Due to the role specialization in our society, married women may allocate more time to housework rather than the employment. Especially for the families with children under 6, women should put more effort on looking after their children. The role specialization also liberates men from doing housework. However, individuals with more children bear more pressure to make money and to support the family. As the role specialization is different between male and female, the effect of household status on the wage might be different. The wage structure also varies across different wage levels. Exploring the distributional effect of household status on gender wage gap in developing countries is important. First, it helps us to find the source of social inequality in developing countries. Second, policy makers could get information from the wage equation and decompose it to find the efficient way to improve the earnings, such as establishing some training programs; opening some positions which are friendly to the special groups like mothers. Finally, analyze the effect at different wage distributional levels help the policy makers to formulate customized policies to the different social class.

The objective of this paper is to investigate the effect of children on gender wage gap in developing countries. To explore this problem, we need to know the wage distribution and other personal characteristics in developing countries. Using STEP data of 12 developing countries, we estimate the wage equation based on Mincer (1974)'s theory. The distributional effect could be obtained with the application of quantile regression. Then we apply the estimated results with Blinder-Oaxaca wage decomposition. The quantile decomposition follows the approaches in Machado and Mata (2005) and Firpo et al. (2007, 2009). The contribution of this paper is to apply the classic method in wage decomposition studies and get aggregate decomposition at different quantiles; another contribution is to apply the methods to explore the gender wage gap in 12 developing countries.

The rest of the paper is organized as follows. The following part is the literature review, which lists some research on gender wage gap and some popular methods. Next we describe the data of interest. Then we talk about the model used in the empirical analysis. After the estimation, we discuss about the results and the conclusion.

1.2. Literature review

Gender wage discrimination has been studied for several decades. In the early study, Altonji and Blank (1999) summarize the research in wage gap by race and gender. In the review, they discuss the causes of gender wage gap, including pre-market human capital in education and family background; different return to experience and seniority; job characteristics; and "unexplained gap", which indicate the gender discrimination. In the recent years, gender wage discrimination is still a popular topic. Blau and Kahn (2017) investigate the explanations of gender wage gap using Blinder-Oaxaca decomposition. They claim that occupation and industry are important. Moreover, there is evidence that number of children has negative effect on women's wage. Mihăilă (2016) finds that international trade plays an important role in gender wage gap and female labor force participation. Amaya & Mougenot (2019) explore the high-paid gender wage gap of Peruvian physicians. They find the likelihood of male physicians to earn high salary is 81% higher than that of female, and the main reason is the unexplained component, which is associated to gender discrimination.

What household status brings to women in the labor market is the motherhood penalty. Budig and England (2001) claim that motherhood is associated with lower hourly pay for reasons such as lose job experience, lower productivity at work, trade of for mother-friendly jobs or other discriminations by employers. They use 1982-1993 National Longitudinal Survey of Youth data with fixed effect model find a wage penalty of 7% per child. Korenman and Neumark (1990) examine the effects of marriage and motherhood on wage. They find a negative effect of marriage and motherhood on wage. However, the authors claim that this result should be understated, since the marriage and motherhood might be associated with lower experience and tenure. Glauber (2007) focuses on the motherhood penalty between different races and marital status. They find that motherhood penalty changes among different races, even for women with similar marital status. For example, for African Americans, wage penalty exists if mothers with more than 2 children, while for Hispanic women there is no motherhood wage penalty. Glauber (2018) investigates the trend of motherhood penalty using data from the Current Population Survey. He finds that from 1980 to 2014, the motherhood penalty decreased. By the early 2010s, the motherhood penalty of high-earning female was eliminated, but in the low-earning level it still exists. However, using data from National Longitudinal Survey of Youth, England et al. (2016) find that during 1979 to 2010, motherhood penalty exists among the women with high skill and high incomes.

The reason of wage inequality might be different between developing countries and developed countries. Gonzalez and Miles (2001) explore the cause of increasing wage inequality in Uruguay. They find that not like the case in developed countries that the fall in the real minimum wage cause the wage inequality, the wage inequality in Uruguay is explained by the increase in return to education. Wood (1997) find that since mid-1980s, increased openness Latin America has widened the wage differential between skilled and unskilled workers. Panagides and Patrinos (1994) first study the union-nonunion wage differential in developing country. They use a household survey in 1989 of Mexico, and find that overall gap is 10.4%, where union has positive effect on women and indigenous people.

In the studies of wage discrimination, Blinder-Oaxaca decomposition is a popular method （Blinder, 1973; Oaxaca, 1973）. It decomposes the difference in the mean wage of male and female into different components including unexplained component, where the unexplained part is usually regarded as gender discrimination. Bartlett and Callahan (1984) analyze the explanations of wage gap in a sample of older white men. They find that the explanations of the marriage premium of men are role specialization and perceived needs. Biltagy (2014) applied Blinder-Oaxaca decomposition with Heckman Selection model to estimate wage difference in Egypt. The result shows that the 25% wage difference between males and females is attributed to education, experience and discrimination against women. However, traditional Oaxaca-Blinder decomposition has several limitations. Goraus et al (2015) compare several popular methods in estimating gender wage gap. Based on current methods such as Oaxaca-Blinder decomposition, they claim that a perfect method needs to address selection issue; account for characteristics of unmatched men and women; and allow to accounting for different wage distribution. However, there seems no literature that meets those requirements simultaneously. Nopo (2008)'s method is based on matching, but lack of distributional analysis.

Recently, researchers tend to focus on the wage inequality at different distributional level. Some of them decompose the wage difference across different wage level (Melly, 2005; Albrecht et al., 2003); others analyze the wage gap at different distributional level. The basic idea of wage decomposition with distributional effects is quantile regression. Machado and Mata (2005) develop a counterfactual decomposition of changes in wage, using Portuguese data for the period 1986-1995. They find the increase in educational level leads greater wage inequality. Then the approach in Machado and Mata (2005) is widely applied in the distributional decomposition research. Heinze (2006) use this method to analyze gender wage gap in Germany. He finds the firm characteristics contribute the largest part of the gender wage gap, and the decomposed parts of gender wage gap vary across different wage levels. Nguyen et al. (2007) investigate the welfare inequality between urban and rural area in Vietnam. By using Machado and Mata (2005) decomposition technique and 1998 survey, they find the gap is due to the difference in education, ethnicity and age for the lower quantiles, and the difference between urban and rural sectors for the higher quantiles. Gardeazabal and Ugidos (2005) measure the gender wage gap in Spain, and find that gender wage discrimination reaches the highest point at ninth percentile. Montes-Rojas et al. (2017) investigate the caste wage differentials in Nepal. They find that the effect of occupation and firm size are uniform across quantiles. But for the low quantiles, education has large effect on the wage gap.

1.3. Data

In this paper, I use the STEP (Skills Toward Employment and Productivity) Measurement data. The dataset are household survey data from the World Bank STEP Measurement Program. This survey provides information on the supply distribution of skills and the demand for skills in labor market of developing countries. The scope of the survey includes household demographic characteristics, education and training, employment, job skill requirements, personality, behavior and preferences, family background, and some other characteristics. In this paper, I use data from 12 developing countries, including Colombia, Ghana, Kenya, Sri Lanka, Ukraine, Armenia, Lao PDR, Macedonia, Vietnam, China, Bolivia, and Georgia. Those 12 countries cover the regions of Central America, Africa Sub-Saharan, Eastern & Southern Europe, Southern & Eastern Asia, which are representative regions of the developing world. They also have different culture and developing level, which could be considered as a good sample for us to learn the wage and household status of the developing world. This household survey data were collected in either 2012 or 2013, where we have Armenia, Georgia, Ghana, Kenya, Macedonia, Ukraine data in 2013, and others in 2012.

Table 1.1 shows the definition of the data that will be used in the estimation.

Table 1.1. goes about here

Figure 1.1 shows the kernel density plot of hourly earnings. From the figure, the distribution of male and female are similar. However, at lower level, the density of female's wage is larger than male's, while at high level, the density of male's earnings is larger. Overall, female's hourly earnings are slightly lower than males.

Table 1.2 shows the data summary of both male and female. The mean value of log hourly earnings of male is 1.0328, and that of female is 0.7596. The dollar values are adjusted for PPP. The male's educational year is also larger than female, where the mean values are 11.06 and 10.7 respectively. The average age of male is 36.5, and female is 37.46. Tenure denoted the number of months in current job, which can be considered as the index of experience. The average values are close to 96.2 for both male and female. Training and certificate are dummy variables, where 1 denotes "has training/certificate" and 0 denotes no. The average values show the percentages that have training/certificate. 9.7% and 7.5% of male and female participate in a training course in last 12 months. The proportion of male that have certificate is 11%, and female 9.8%. The average number of children of male is 0.33, and female is 0.43.

1.4. Model

1.4.1. Wage Equation

Firstly, I do regression on wage equation for male and female separately. The wage equation could be written as follows:

$$\ln(hourly\ earning)_{gi}$$
$$= \beta_{g0} + \beta_{g1}edu_{gi} + \beta_{g2}age_{gi} + \beta_{g3}tenure_{gi} + \beta_{g4}training_{gi}$$
$$+ \beta_{g5}certificate_{gi} + \beta_{g6}children_{gi} + \varepsilon_{gi}$$

Where g denotes gender, valued as $m = male, f = female$. The dependent variable on the left-hand side is log of hourly earnings. In this case, the estimated coefficients provide the percentage change of wage that caused by the change in the independent variables.

1.4.2. Quantile regression

The idea of quantile regression is first developed by Koenker and Bassett (1978). Compare with OLS which is based on conditional mean, quantile regression allows for distributional effect of regressors, and more robust to outliers. In quantile regression, we estimate conditional quantile of y, which could be written as:

$$Q_t(y|x) = X'\beta_t$$

Where q_t is the t^{th} quantile of y such that

$$\mathrm{t} = \Pr[\,\mathrm{y} \le q_t]$$

and the marginal effect is calculated by

$$\beta_{jt} = \frac{\partial Q_t(y|t)}{\partial x_j}$$

1.4.3. Wage decomposition

Then I applied the Blinder-Oaxaca decomposition, which is based on the wage equation that estimated on conditional mean:

(1) $Y_m = X_m\beta_m + u_m$

(2) $Y_f = X_f\beta_f + u_f$

(3) $\bar{Y}_m - \bar{Y}_f = \hat{\beta}_m\bar{X}_m - \hat{\beta}_f\bar{X}_f = \hat{\beta}_m\left(\bar{X}_m - \bar{X}_f\right) + \bar{X}_f(\hat{\beta}_m - \hat{\beta}_f)$ or

(4) $\bar{Y}_m - \bar{Y}_f = \hat{\beta}_m\bar{X}_m - \hat{\beta}_f\bar{X}_f = \hat{\beta}_f\left(\bar{X}_m - \bar{X}_f\right) + \bar{X}_m(\hat{\beta}_m - \hat{\beta}_f)$

Where B_m and B_f are estimated coefficients from male/female's wage equation. In equation (4) we assume that the wage of male is non-discriminatory, while in equation (5) wage of female is non-discriminatory. The estimations based on those equations might be different. However, in either equation, the first term in the left-hand side denotes the unexplained part, usually considered as gender discrimination; the second term in the right-hand side denotes the gender wage difference in explanatory variables.

In the distributional wage decomposition, I apply the Machado and Mata (2005) approach (Machado and Mata, 2005; Nguyen and Albrecht, 2007). The MM approach allows us to decompose the gender wage gap into two factors; one is the contribution of the difference in characteristics of male and female, which is x in the equation, and another is gender discrimination, also called coefficients effect, which is caused by the difference in the coefficients of estimation. The idea of MM approach is the counterfactual analysis that based on quantile regression. And the key factor is the distribution of female's wage if they have male's return rates, which are the coefficients in male's wage equation. The procedure of MM decomposition could be described as follows: 1) generate a random sample $u_1, u_2, \ldots\ u_m$ from uniform distribution $\mathrm{U}[0, 1]$ with sample size m; 2)using data of male, run regression at $\tau = u_i$ to get m estimates of QR the coefficients $\hat{\beta}_\tau^m$ for each u_i; 3)generate a m size random sample with replacement from the rows of Z^f from female data; 4) construct the counterfactual distribution y_τ^c with the coefficients from step 2 and the sample from step 3, which could be written as $y_\tau^c = Z^f \hat{\beta}_\tau^m$. Then the aggregate decomposition at each quantile could be written as:

$$y_\tau^m - y_\tau^f = (y_\tau^m - y_\tau^c) + \left(y_\tau^c - y_\tau^f\right)$$

Where the first term on the right-hand side denotes covariates effect, which indicate the difference in characteristics, and the second term denotes the coefficients effect.

1.5. Results

Table 1.3 shows the result of the regressions on wage equation of both male and female. The regressions include country fixed effects. The first two columns are the results from OLS regression, and the rest of the table shows the quantile regression results.

Table 1.3 goes about here

From the OLS estimation result, we find that generally education has positive effect on hourly earnings for both male and female. For male, one more year's education increases their earnings by 6.90%, and 7.45% for female. Age, training and certificate have positive effect on the earnings of both male and female. Training increases male and female's earnings by 15.5% and 19.0% respectively, and certificate increases the earnings by 9.18% and 12.1% respectively. Effect of children on earnings is not significant.

Figure 1.1, Figure 1.2 and Figure 1.3 go about here

The results of quantile regression and figures show that the effect on hourly earnings varies across different distribution. The effects of educational years on both male's and female's earnings at low quantile are larger than the high quantiles. Certificate has positive effect on both male and female earnings from 10 to 40 quantile. But from 50 to 90 quantiles, certificate only has positive effects on female's earnings. The effect of number of children also varies with different quantiles. From quantile 10 to 80, number of children only shows positive effect on male's earnings, and negative effect on female's earnings at quantile 10. At low earning level, which is 10th quantile, one more child decreases female's earnings by 5.86%. At the 20th quantiles, children only have positive effect on male's earnings but no effect on female. From quantiles 50 to 90 results, the effect of number of children might be associated since higher earnings may lead to more children.

Analyzing the separate regression results will be helpful to find the different status of different regions. The estimated results of regression by region are shown in the appendix. In the Central America, which includes Colombia and Bolivia, the effect of education on earnings is significant. For male, one-year increase in education increases hourly earnings by 6.73%, and 7.23% for female. Besides education, training also has positive on the hourly earnings of both male and female, while the effect on male is slightly greater than female. The difference of educational effect is the largest at the 29th quantile. One-year increase in education raises 5.78% of male's earning, but 7.14% of female'. In the high quantiles of 60 to 80, the effect of education on earnings goes larger than other quantiles. At the top quantile of 90th, the effect of education goes smaller, but the effect on male is greater than female by 1.02%. The effect of training is significant at low and middle quantile levels, but insignificant at high quantiles. At quantile 30, one more children decrease female's earnings by 10.2%. In the Africa Sub-Saharan, which includes Ghana and Kenya, the overall effect if education is slightly larger than Central America, but the difference between male and female is only 0.01%. Besides education, working experience and training also have significant effect on hourly earnings. One-month increase in experience increases hourly earnings by 1.49% of male, and 1.22 of female. The effect of training of earnings of female is larger than male. Certificate also has positive effect on male's earnings. The effect of education is larger at middle levels than others. At 40th quantile, the difference of educational effect between is the largest, which is 1.07%. The effect of training is significant at all levels. Especially, as quantile increases the effect of training on female's earning goes much larger than that of male. At quantile 10, one more children decreases female's wage by 12.1%, while at quantile 89 and 90, the effect of children on female is positive. In the Eastern & Southern Europe, which includes Ukraine, Macedonia, Georgia and Armenia, the effect od education on hourly earnings is smaller than Africa Sub-Saharan, but the difference between male and female is larger. The effects of age, tenure, training and certificate are also smaller than those in Africa Sub-Saharan. The effect of education goes larger as quantile increases. The educational effect on female is greater than male; at quantile 40, the difference between male and female is the largest, which is 1.69%. At quantile 80 and 90, the effect of training turns insignificant on hourly earnings. The effect of children on male's wage is positive and significant at most quantiles, but no effect on female's earnings. In the Southern and Eastern Asia, which includes China, Vietnam and Lao PDR, the difference of educational effect between male and female is the largest among these regions, where one-year increase in education increases male's hourly earnings by 6.4%, and 8.43% for female. The effects of tenure and training are insignificant, but certificate and urban is significant. The effect of urban on male's earnings is 36.4%, and 19.9% on female. The effect of education on earnings decreases as quantile increase. At quantile 10, the difference between male and female is the largest, where the educational return rate of female is greater than male by 4.74%. The difference at high quantiles is smaller than low quantiles. The effect of certificate on female is larger than male at middle and high quantiles. Urban has significant effect on earnings at low quantiles, but from quantile 70 it becomes insignificant. Number of children has negative effect on female's earnings at quantile 20, but positive at quantile 70 and 90. The effect on male is insignificant.

Table 1.4 goes about here.

Table 1.4 shows the result from traditional Blinder-Oaxaca decomposition. The row difference of log hourly earnings between male and female is 0.272 and significant, where 0.0361 could be explained by the difference in the characteristics and 0.241 is unexplained, which is from discrimination. In the explained part, 0.0192 comes from the difference in education, and 0.00136 comes from certification. The effect of children contributes -0.000796 of wage gap. In the unexplained part, wage gap significantly comes from constant term.

Table 1.5 goes about here

Table 1.5 shows the result that follows Machado and Mata (2005) approach, which provide an aggregate decomposition. The differences between male's and female's earning are significant at all distributional levels. The gender wage gap at low quantiles is larger than high quantiles. At middle quantiles 40 to 70, the contribution of characteristics differences is small. However, the contribution of discrimination is larger than characteristics differences

1.6. Conclusion

In this paper, I investigate the gender wage gap across 12 developing countries. I applied the classical Blinder-Oaxaca decomposition and combine it with quantile regression to find the distributional effects. I apply the similar approach to the Machado and Mata (2005) to obtain the distributional decomposition. This approach is popular in recent quantile decomposition research. The data that applied in the estimation are from STEP (Skills Toward Employment and

Productivity) Measurement data. This household survey provides information on household demographic characteristics, education and training, employment, behavior and preferences, family background, and some other characteristics from 12 developing countries.

From the estimated results, we find that gender wage gap exists at all wage levels. The wage gap is larger at high wage level than low level. Education training and certificate play the important roles in the wage equation. The effect of number of children varies across the quantiles: in the low level, it has negative on female's earning; at middle levels, it has positive effect on male's earnings; and at high wage level it has positive effect on both female's and male's earnings. Then gender discrimination exists at all wage levels. Specifically, the discrimination in number of children, sometimes we call it motherhood penalty, which exists at low wage level groups.

As for the policy makers, education is the best channel to reduce gender wage gap, or labor market inequality. Considering the motherhood penalty at low wage level groups, government should open more mother friendly positions and increase the welfare to the families with more children. Mothers at low wage levels should also be encouraged to take part in the training programs or get certificates, since these are the efficient ways to increase their earnings and reduce the gender wage gap.

Figure 1.1. Kernel Density Plot of Hourly Earnings

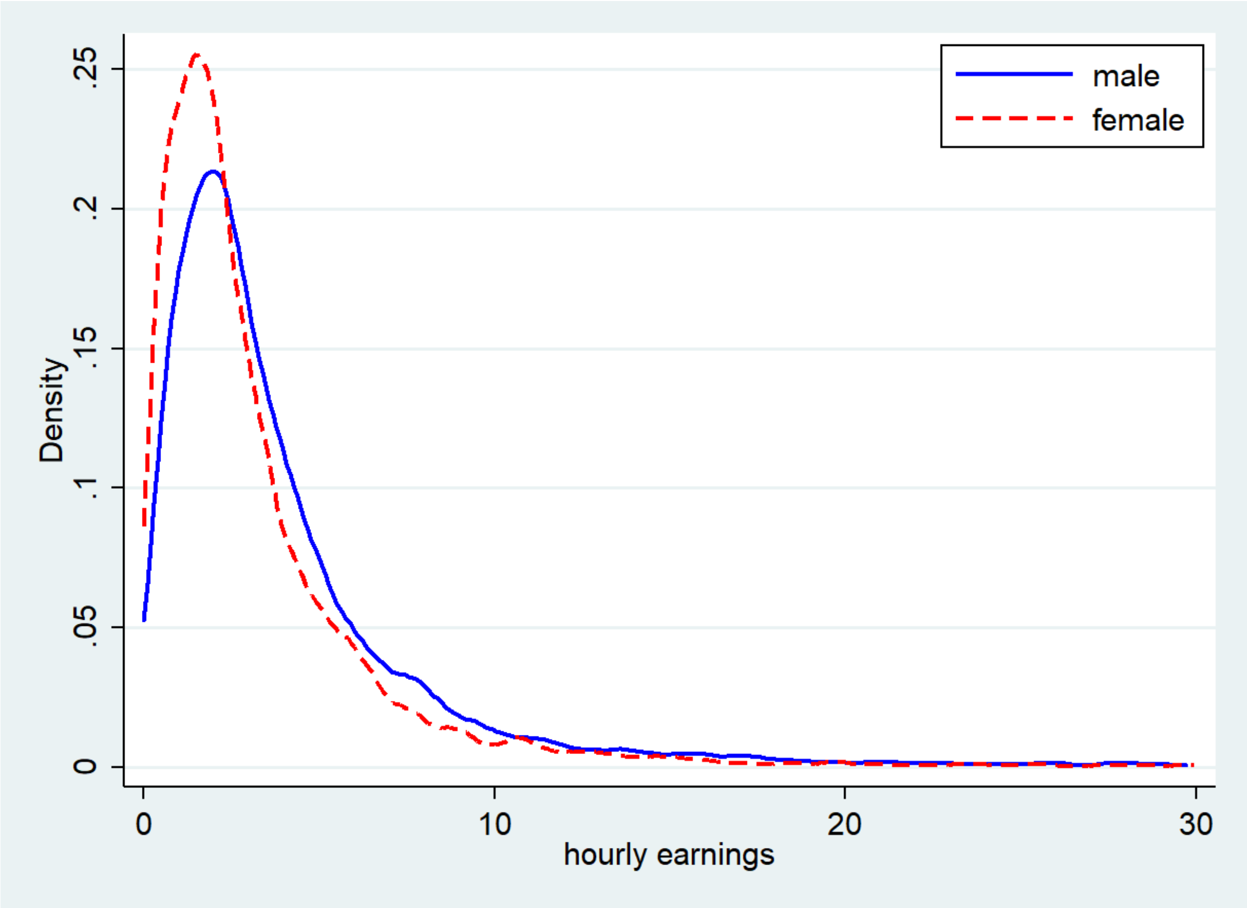

Figure 1.2. Quantile Regression Plot of Male's Log Hourly Earnings

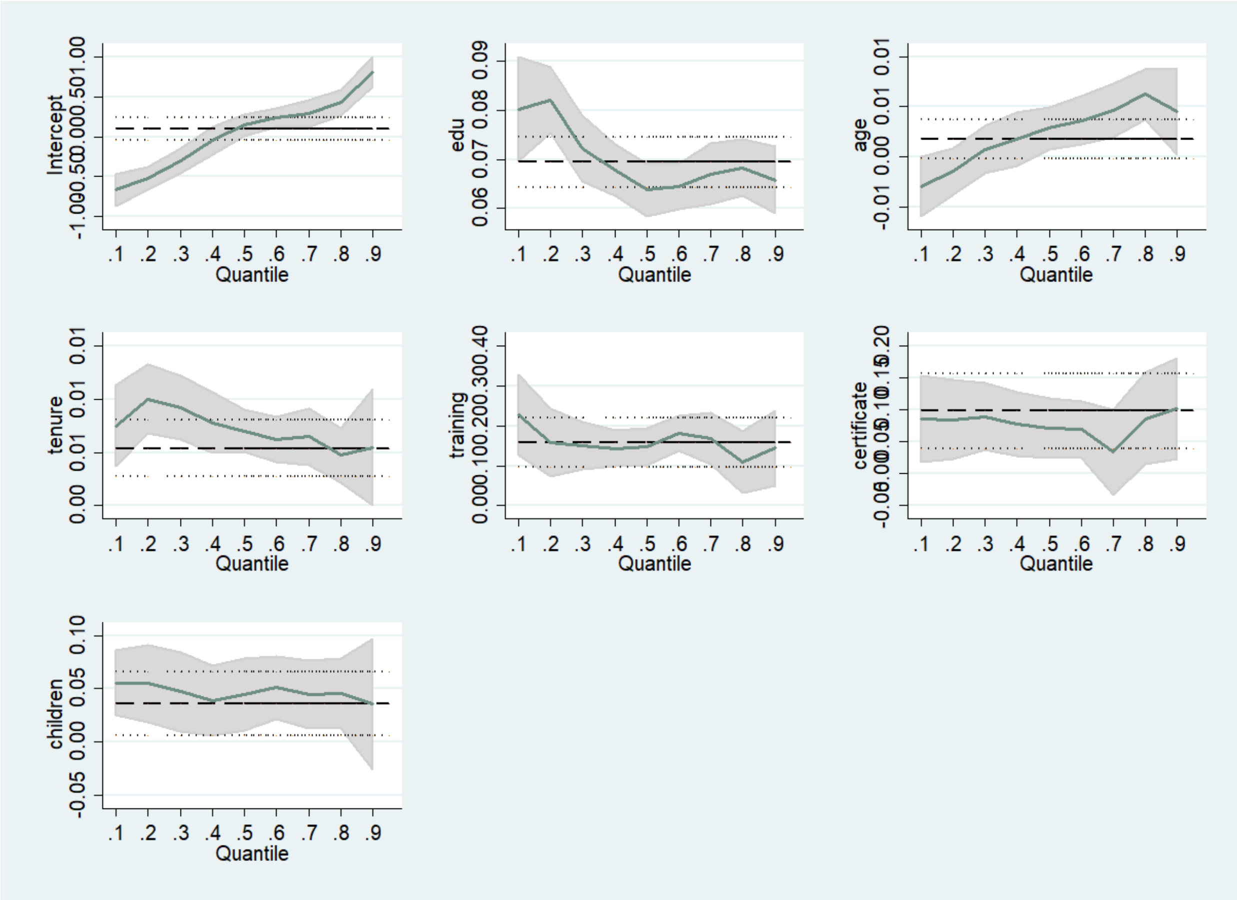

Figure 1.3. Quantile Regression Plot of Female's Log Hourly Earnings

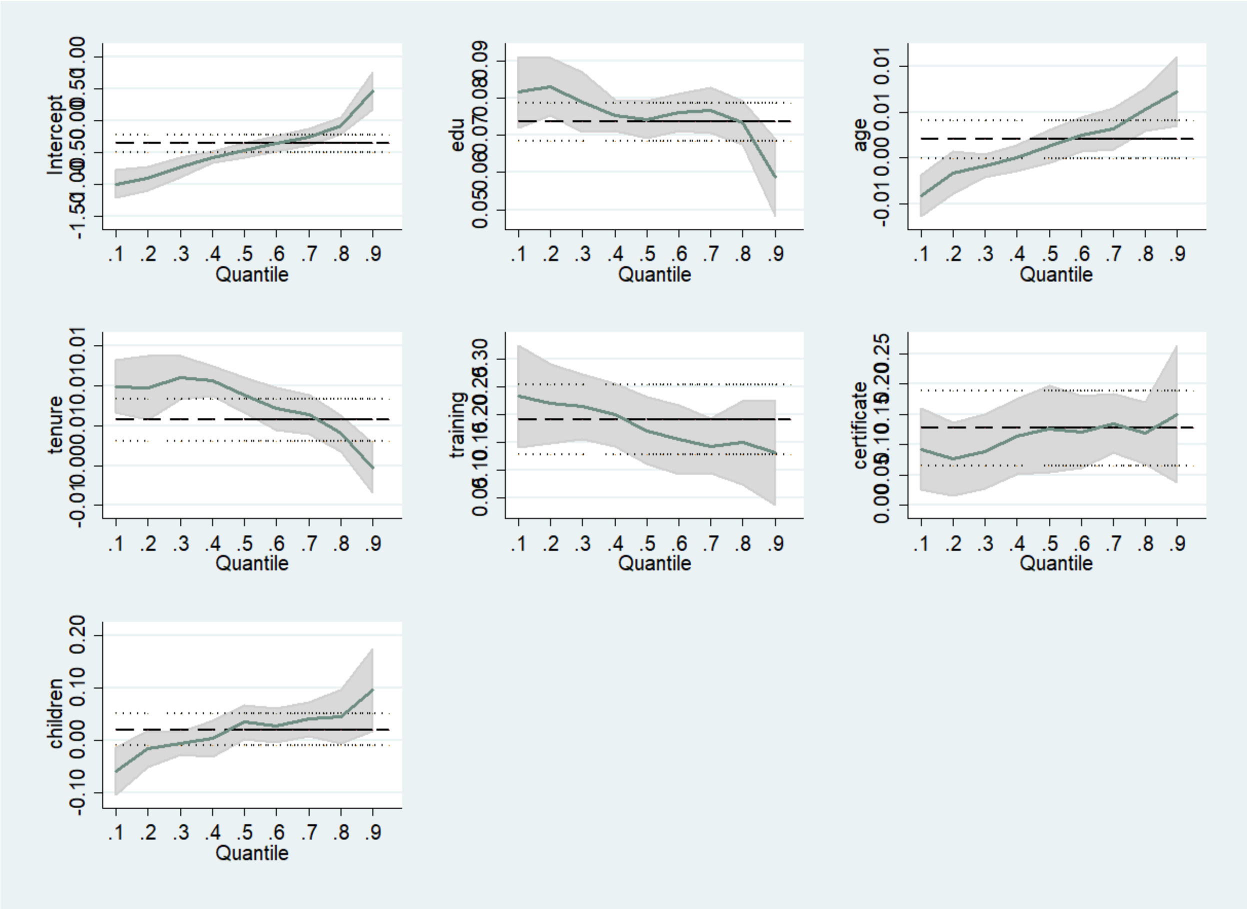

Table 1.1. Data Definition

Variable Name	Definition
ln_earnings_h_usd	USD log of Hourly earnings
edu	Number of years of education
age	Age in years
tenure	Number of years in current job
training	Participated in a training course such as work-related training or private skills training, that lasted at least 5 days/ 30 hours (not part of the formal educational system) in last 12 months. 1: yes, 0: no.
certificate	An industry-recognized or government certificate in a particular field (not from a formal ed. Institution)? 1: has, 0: no.
children	Number of children under 6 years old
urban	

Table 1.2. Data summary

Male					
Variable	Obs	Mean	Std. Dev.	Min	Max
ln_earning~d	8,788	1.0328	1.0175	-7.708	6.5461
years_educ~t	14,180	11.061	4.054	0	28
age	14,322	36.503	13.751	15	64
tenure_	9,890	8.0192	9.0851	0	70
training	14,314	0.0975	0.2967	0	1
certificate	14,322	0.1184	0.3231	0	1
children	14,322	0.3313	0.6341	0	7
urban	14,322	0.926	0.2618	0	1
Female					
Variable	Obs	Mean	Std. Dev.	Min	Max
ln_earning~d	9,580	0.7596	1.073	-7.1484	6.2947
years_educ~t	20,840	10.709	4.3482	0	28
age	21,247	37.457	13.632	15	64
tenure_	11,011	8.0255	9.1858	0	53
training	21,241	0.0752	0.2637	0	1
certificate	21,247	0.0977	0.2969	0	1
children	21,247	0.4251	0.6862	0	6
urban	21,247	0.9272	0.2598	0	1

Table 1.3. Results from OLS and Quantile Regression with Country Fixed Effect

Quantiles	OLS		10		20		30		40	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0690***	0.0745***	0.0770***	0.0808***	0.0795***	0.0795***	0.0704***	0.0794***	0.0665***	0.0773***
	(0.00270)	(0.00265)	(0.00510)	(0.00416)	(0.00397)	(0.00319)	(0.00406)	(0.00329)	(0.00379)	(0.00309)
Edu2	0.0120***	0.0164***	0.00204	0.00217	0.00665**	0.00963***	0.00968***	0.0105***	0.0119***	0.0158***
	(0.00324)	(0.00330)	(0.00389)	(0.00378)	(0.00268)	(0.00268)	(0.00241)	(0.00242)	(0.00385)	(0.00187)
age	0.0403***	0.0323***	0.0573***	0.0184***	0.0496***	0.0298***	0.0428***	0.0332***	0.0411***	0.0300***
	(0.00574)	(0.00583)	(0.00968)	(0.00626)	(0.00565)	(0.00641)	(0.00682)	(0.00638)	(0.00686)	(0.00622)
Age2	-0.000489***	-0.000384***	-0.000773***	-0.000299***	-0.000651***	-0.000399***	-0.000545***	-0.000425***	-0.000504***	-0.000376***
	(0.0000715)	(0.0000725)	(0.000126)	(0.0000874)	(0.0000785)	(0.0000841)	(0.0000941)	(0.0000815)	(0.0000975)	(0.0000812)
tenure	0.00576***	0.00647***	0.00765***	0.0114***	0.0105***	0.0105***	0.0105***	0.0114***	0.00862***	0.0107***
	(0.00135)	(0.00133)	(0.00234)	(0.00174)	(0.00176)	(0.00176)	(0.00146)	(0.00117)	(0.00136)	(0.000885)
training	0.155***	0.190***	0.207***	0.213***	0.160***	0.237***	0.150***	0.207***	0.148***	0.188***
	(0.0317)	(0.0321)	(0.0392)	(0.0352)	(0.0285)	(0.0224)	(0.0286)	(0.0250)	(0.0335)	(0.0274)
certificate	0.0918***	0.121***	0.0919**	0.0802*	0.0693**	0.0605*	0.0864***	0.0956***	0.0702***	0.0971***
	(0.0303)	(0.0314)	(0.0408)	(0.0415)	(0.0327)	(0.0331)	(0.0317)	(0.0291)	(0.0271)	(0.0278)
children	0.0207	0.0162	0.0259	-0.0586**	0.0298*	-0.0163	0.0262	-0.0116	0.0188	0.00290
	(0.0152)	(0.0154)	(0.0237)	(0.0267)	(0.0161)	(0.0195)	(0.0182)	(0.0219)	(0.0184)	(0.0188)
Urban	0.0836	0.110*	0.0674	0.328***	0.148**	0.306***	0.0704***	0.0794***	0.0546	0.122*
	(0.0511)	(0.0566)	(0.0871)	(0.0966)	(0.0707)	(0.0864)	(0.00406)	(0.00329)	(0.0459)	(0.0647)
_cons	-0.751***	-1.156***	-1.749***	-1.704***	-1.541***	-1.794***	-1.152***	-1.589***	-0.862***	-1.364***
	(0.133)	(0.138)	(0.244)	(0.138)	(0.161)	(0.146)	(0.156)	(0.152)	(0.166)	(0.131)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quantiles	50		60		70		80		90	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0635***	0.0749***	0.0660***	0.0757***	0.0671***	0.0776***	0.0673***	0.0745***	0.0676***	0.0610***
	(0.00323)	(0.00240)	(0.00348)	(0.00203)	(0.00358)	(0.00190)	(0.00349)	(0.00277)	(0.00549)	(0.00416)

Edu2	0.0137***	0.0169***	0.0142***	0.0199***	0.0164***	0.0229***	0.0144***	0.0226***	0.0177***	0.0296***
	(0.00347)	(0.00209)	(0.00222)	(0.00269)	(0.00267)	(0.00353)	(0.00388)	(0.00444)	(0.00539)	(0.00489)
age	0.0421***	0.0318***	0.0392***	0.0339***	0.0352***	0.0311***	0.0204***	0.0252***	0.0212***	0.0273**
	(0.00557)	(0.00736)	(0.00548)	(0.00588)	(0.00701)	(0.00683)	(0.00511)	(0.00945)	(0.00664)	(0.0115)
Age2	-0.000495***	-0.000391***	-0.000450***	-0.000390***	-0.000389***	-0.000345***	-0.000180***	-0.000259**	-0.000206**	-0.000258**
	(0.0000791)	(0.0000889)	(0.0000731)	(0.0000721)	(0.0000874)	(0.0000843)	(0.0000663)	(0.000118)	(0.0000882)	(0.000125)
tenure	0.00693***	0.0105***	0.00742***	0.00788***	0.00669***	0.00761***	0.00574***	0.00495***	0.00641**	0.000630
	(0.000975)	(0.000785)	(0.00101)	(0.000945)	(0.00136)	(0.00114)	(0.00137)	(0.00129)	(0.00261)	(0.00195)
training	0.146***	0.187***	0.168***	0.164***	0.149***	0.134***	0.132***	0.132***	0.113**	0.100**
	(0.0298)	(0.0282)	(0.0257)	(0.0244)	(0.0351)	(0.0233)	(0.0390)	(0.0350)	(0.0526)	(0.0460)
certificate	0.0607**	0.121***	0.0616*	0.120***	0.0328	0.119***	0.101***	0.126***	0.120**	0.131***
	(0.0273)	(0.0207)	(0.0352)	(0.0196)	(0.0307)	(0.0244)	(0.0379)	(0.0376)	(0.0559)	(0.0427)
children	0.0252*	0.0307	0.0336***	0.0287*	0.0368**	0.0401**	0.0262*	0.0512*	0.0173	0.0892***
	(0.0148)	(0.0201)	(0.0125)	(0.0147)	(0.0155)	(0.0203)	(0.0158)	(0.0278)	(0.0263)	(0.0297)
Urban	0.000625	0.0758	-0.00926	0.0361	-0.00239	-0.0661	0.0132	-0.0525	0.0501	0.0701
	(0.0495)	(0.0897)	(0.0651)	(0.0818)	(0.103)	(0.0804)	(0.112)	(0.125)	(0.104)	(0.192)
_cons	-0.618***	-1.229***	-0.515***	-1.090***	-0.361***	-0.866***	0.0636	-0.560**	0.309	-0.218
	(0.133)	(0.184)	(0.122)	(0.142)	(0.139)	(0.154)	(0.167)	(0.241)	(0.203)	(0.356)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N									8656	9369

Standard errors in parentheses

[*] $p < 0.1$, [**] $p < 0.05$, [***] $p < 0.01$

Table 1.4. Result from Blinder-Oaxaca Decomposition with Country Fixed Effect

	ln_earnings_h_usd
overall	
group_1	1.032***
	(0.0109)
group_2	0.760***
	(0.0111)
difference	0.272***
	(0.0156)
explained	0.0316***
	(0.00730)
unexplained	0.241***
	(0.0141)
explained	
edu	0.0192***
	(0.00476)
age	-0.000508
	(0.000371)
tenure	-0.000364
	(0.000790)
training	0.00123
	(0.000858)
certificate	0.00139**
	(0.000620)
children	-0.000796*
	(0.000427)
urban	-0.00209*
	(0.00108)
unexplained	
edu	-0.0450
	(0.0450)
age	-0.00619
	(0.0590)
tenure	-0.00527
	(0.0153)
training	-0.00393
	(0.00505)
certificate	-0.00371
	(0.00538)
children	0.00654
	(0.00911)
urban	-0.0266
	(0.0848)
_cons	0.497***
	(0.130)
Country FE	Yes
N	18104

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5. Results from Machado and Mata (2005) Approach with Country Fixed Effect

Quantile	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Raw	0 .408***	0. 320***	0.277***	0. 255***	0. 240***	0. 229***	0. 217***	0. 211***	0. 224***
	(0. 0151)	(0. 00978)	(0. 00917)	(0. 00927)	(0. 0101)	(0. 0105)	(0. 0112)	(0. 0131)	(0. 0190)
Characteristics	0. 0873***	0. 0486***	0. 0295**	0. 0196*	0. 0159*	0. 0178*	0. 0195*	0. 0203*	0. 0338**
	(0. 0216)	(0. 0150)	(0. 0127)	(0. 0116)	(0. 0105)	(0. 0108)	(0. 0111)	(0. 0123)	(0. 0169)
Coefficients	0. 321***	0. 271***	0. 248***	0. 235***	0. 224***	0. 212***	0. 197***	0. 190***	0. 190***
	(0. 0209)	(0. 0147)	(0. 0122)	(0. 0107)	(0. 00957)	(0. 00997)	(0. 0103)	(0. 0116)	(0. 0150)

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2 The effects of parental education on children's income

2.1. Introduction

Generally, personal income is affected by the educational level, experience, age, gender, race and some personal characteristics. A large majority of researches on wage investigate the effect of education on personal income. These research studies contributed significantly to policy making and help to reduce social inequality problem.

However, the characteristics of parents also have influence on children's income. Biologically, children who were born in families with high educational level or great wealth may enable the development of talents in learning skills and investing. Wealthy families may also provide more social resources and human capital to their children. Becker and Tomes (1979) explain the fortunes of children are linked to their parents through investments and endowments from their parents. The income of children would be higher if they receive more human capital; also, children with more endowments such as ability, family reputation and other family characteristics have high return in labor market. Those intergenerational effects cause problems such as intergenerational inequality and mobility. Checchi et al. (1999) exploit the relationship between education financing and intergenerational mobility in Italy and the U.S. Based on the fact that family background is important for labor market success, they find that a centralized and egalitarian education can't help poor children in the upward mobility. Titma et al. (2003)

investigate intergenerational mobility in Soviet Society and found it is strongly affected by education.

Based on the phenomenon that family characteristics have large impact on intergenerational mobility, the studies of intergenerational effects are important. The purpose of this paper is to investigate the effects of parental education on children's income in developing countries. Education is likely a proxy for networks of parent's and strength of their connections. Then the return to parental education could be considered as the return to parental networks and social resources.

With high quality of family social network and human capital, those children may have higher return in labor market than others. The hypothesis of this paper is that parental educational level has positive effect on children's earnings in developing countries. Besides parental education, there are also many other factors that affect income, such as age, gender, training experience and educational level of themselves. I applied Heckman Selection model to fix selection bias. To solve this endogeneity problem of the omitted ability bias in the wage equation, I use two stage least square (2SLS) in the empirical estimation. The instrumental variable for children's education is the school location. In the estimation, parental educational levels are categorical variables.

There are five sections in this paper. In the first section I present some related literature about wage study and intergenerational effects. Next, I describe the basic information of the dataset. In the third sections I talk about the empirical model that will be applied in this paper. Then I discuss the estimated results with different models. The last section is the conclusion.

2.2. Literature review

There are plenty of papers written on the intergenerational effects of education and income. Those intergenerational effects emphasize social inequality and intergenerational mobility. Altonji and Dunn (1996) examine the role of parental education on wage in the U.S. The empirical result with fixed effects suggests a one-year increase in mother's education raises the return to education by one percent. This result indicates family background characteristics have large effects on labor market payoff to schooling. Borisov and Pissarides (2016) focus on the intergenerational transmission of human capital. Using longitudinal data for Russian over 1994-2013, they find that educational attainment has high intergenerational correlation with earning capacity. They attribute this relationship to the informal networks. Belzil and Hansen (2003) use a structural dynamic programming model to estimate the relative importance of family background and individual specific abilities in explaining the differences in schooling and wages. They find that individual specific abilities account for 73% of the variations in wages. Family background accounts for 19%, while ability corrected for family background, account for 8.0%. Heckman and Hotz (1986) find that after adding the parental education into the wage equation, the return to the own education decrease by 25%.

However, parental educational level may not affect children's earnings directly, but through children's education. Lee et al. (2004) claim that parents can play a key role in children's college enrollment, since students may receive more information and support from higher educated parents to attend college. They explore how the students' experiences differ by generational status, especially parental education levels. They conduct five logistic regression equations, using parental education level as dependent variables. The findings indicate students' experience do not linearly related to parental educational level. Restuccia and Urrutia (2004) focus on the intergenerational persistence of earnings in the U.S. They find that almost 50% of the intergenerational correlation on earnings is accounted by the investment in children's education, especially early education. Acemoglu and Pischke (2001) exploit the effect of parental resources on inequality using data from U.S. National Center for Education Statistics (NCES). From the fact that families at the bottom of income distribution were much poorer from 1970s to 1990s, they find that a 10 percent increase in family income increases the probability of attending college by 1.4 percent.

The studies on wage inequality are extensive. Juhn et al. (1993) examine the increasing wage inequality for males over 1963 to 1989. They find the increase in wage inequality is due to the rise in return to skill other than year of schooling and experience. The explanation might be the increasing demand for skill in the United States. Mincer (1997) claims that persons with more schooling tend to invest more in job training. Therefore, the correlation between wage and years of education is positive. Lemieux (2006) applied a quantile regression model and explain the increasing wage inequality. He suggests that high return to postsecondary education plays a crucial role in the concentrated wage distribution. Balestra and Beckes-Gellner (2007) study returns to education over wage distribution. They use quantile regression with IV and find evidence of high level of heterogeneity. The results show that returns to education are higher at lower quantile than high quantile, which means at higher quantile individuals with different educational levels earn almost the same. Martins and Pereira (2004) investigate the relationship between education and wage inequality. They use quantile regression to estimate returns to education across the wage distribution. The empirical evidence from 15 countries shows that individuals who are more skilled have higher returns to schooling. Abdullah at al. (2015) find that education is an effective way to reduce inequality in Africa. The evidence is that education shows negative effect on the income share of the top earners but positive effect in the bottom earners. Besides, the effect of secondary schooling is larger than primary schooling. The effect of education may be different in different gender group. Kadir & Sukma (2019) find that for female, returns to education is higher than male. But education is still an efficient way to reduce social inequality by promoting same level of education for both male and female.

In recent years, there are also many studies on wage inequality in developing countries. Yirmiyahu et al. (2017) investigate how the labor market reflects the Israeli Academic Colleges Law, which is issued to improve the higher education accessibility to the Israeli Arab minority. They find evidence showing that the improvement of access to higher education increases the earnings of Israeli Arabs. Jackman and Bynoe (2014) study the "Free Education for All" policy in Barbados. Under this policy, government provides free education from primary to tertiary level. They find this policy help improve income in lower group and reduce wage inequality. In developing countries, openness to trade is also an important factor that narrow wage gap between workers with different skills (Wood, 1997; Hanson &Harrison, 1999; Beyer et al., 1999; Marjit et al., 2004). One of the reasons is growing international trade may increase the demand of low-educated/skilled workers in developing countries (Sachs & Shatz, 1996). Gonzalez and Miles (2001) explore the cause of increasing wage inequality in Uruguay. They find that not like the case in developed countries that the fall in the real minimum wage cause the wage inequality, the wage inequality in Uruguay is explained by the increase in return to education. Panagides and Patrinos (1994) first study the union-nonunion wage differential in developing country. They use a household survey in 1989 of Mexico, and find that overall gap is 10.4%, where union has positive effect on women and indigenous people. Psacharopoulos (1994) finds that at global scale, the returns of education decline as schooling level increase. The returns of investment in women's education are higher than man.

The estimation of wage equation often comes with sample selection problem, since the data of wage are missing if an individual choose "not work". To fix the selection bias, Heckman (1976; 1977) present a sample selection model which contains two stages estimation. In the first stage, he uses a probit model to estimate an individual is working or not. The second stage equation is the wage equation with the estimated Inverse Mill's Ratio from the first stage. Heckman selection model has been widely applied in wage studies (Neuman & Oaxaca, 2004; Dubin & River, 1989; Mulligan & Rubinstein, 2008). Arellano & Bonhomme (2017) combine Heckman selection model with quantile regression. They find sample selection has strong effect on male wage at the bottom of the distribution, but smaller effect for female. This result indicates the gender wage gap at the bottom could be decreased by solving selection problem. Biltagy (2014) applied Blinder-Oaxaca decomposition with Heckman Selection model to estimate gender wage difference in Egypt. The result shows that the 25% wage difference between males and females is attributed to education, experience and discrimination against women.

The contribution of this paper is to analyze the intergenerational effect in the developing world and correct both selection bias and omitted ability bias problems. Considering the impact on children's schooling, I use the location of school as the IVs. I also compare the results form OLS model and 2SLS models with IVs and Heckman selection model.

2.3. Data

In this paper, I use the STEP (Skills Toward Employment and Productivity) Measurement data. The dataset is household survey data from the World Bank STEP

Measurement Program. This survey provides information on the supply distribution of skills and the demand for skills in the labor market of developing countries. The scope of the survey includes household demographic characteristics, education and training, employment, job skill requirements, personality, behavior and preferences, family background, and some other characteristics. In this paper, I use data from 12 developing countries, including Colombia, Ghana, Kenya, Sri Lanka, Ukraine, Armenia, Lao PDR, Macedonia, Vietnam, China, Bolivia, and Georgia. Those 12 countries household survey data were collected in either 2012 or 2013, where we have Armenia, Georgia, Ghana, Kenya, Macedonia, Ukraine data in 2013, and others in 2012.

Table 2.1 goes about here

Table 2.2 goes about here

There are 32,517 observations in the dataset. However, only 16,515 hourly labor earnings are available. The average log hourly labor earning is 0.94. Years education act provides the years of education that actually completed. The average year of education is 11.211 years and the maximum year is 28. Max parent edu is a categorical variable providing the maximum level of parents' education. This variable is classified by the International Standard Classification of Education (ISCED), which is a statistical framework organizing educational information maintained by the United Nations Educational, Scientific and Cultural Organization (UNESCO). The range of the values is from 0 to 3. According to ISCED 1997 levels of education, 0 represents pre-primary education; 1 represents Primary education or first stage of basic education; 2 represents Lower secondary education or second stage of basic education and Upper secondary education; 3 represents Post-secondary non-tertiary education and higher level. The ages of the interviewees are from 15 to 60, while the average age is 36.875. Tenure denotes the number of mouths in the current job, which indicates the experience of an individual. Gender, training, public sector, certificate and has_spouse are dummy variables. 1 in gender indicates female and 0 indicates male. If an individual didn't participate in any training course in the last 12 months, the value of training is 0, otherwise 1. 0 in the pub_emp indicates the private sector of the employee, otherwise public sector. In the question of "does the individual have certificate and/or a spouse?" if the answer is yes we mark it as 1, otherwise 0. The average value of the number of children is 0.4, with the minimum number of children 0, and maximum 7.

Table 2.3 goes about here

Table 2.3 provides the information from different countries. Comparing the mean values among those countries, we can find Macedonia has the highest log hourly earnings, which is 1.51, while Ghana has the lowest 0.41. People in Georgia have the highest year of education, 14.197, while the year of education in Laos is 7.87. The average age of the workers ranges from 29 to 42. Ghana has the largest average number of children below 6 years old, which is 0.523. Due to the one child policy, China has the lowest number of children, which is 0.126.

2.4. Identification/ model specification

2.4.1. Wage Equation

According to Mincer (1974), wage is determined by educational background, experience, job training, gender, age, etc. Blau and Kahn (2017) claim that marital status and industries are also important explanatory factors. In this paper, children's hourly earnings are considered as a function of maximum level of parents' education, children's years of education completed, age, number of months in current job (tenure), and some dummy variables that indicate gender, marital status, public or private sector, industry-recognized certificate and training status. The regression equation with OLS estimation could be written as

$$(1) \qquad \ln(hourly\ earning)_i = \beta_0 + \beta_1 \,\mathrm{Max}\, parents\ edu_i + \beta_2 Year\ edu\ act_i + \beta_3 age_i + \beta_4 {age_i}^2 + \beta_5 tenure_i + \beta_6 gender_i + \beta_7 training_i + \beta_8 public_i + \beta_9 certificate_i + \beta_{10} marriage_i + \beta_{11} urban_i + \varepsilon_i$$

Where i denotes individual interviewee of the survey. $Max\ parents\ edu$ is a categorical variable, gender, training, public sector, certificate and marriage are 0,1 variables.

2.4.2. Identification

To investigate the effect of parental education on children's earning, we need to identify two potential problems. The first one is sample selection bias, and the second is the omitted ability bias in the wage equation.

The data of hourly earnings are truncated as those of unemployed individuals are unobservable. The estimation from OLS may cause a biased conditional mean. Heckman Selection Model (Heckman, 1977) is a popular approach to avoid this problem. Heckman selection model contains two stages estimation. The first stage is the selection equation, which use probit model to estimate the probability of being selected. The selection equation could be written as:

$$(1) \qquad y_{1i} = x_{1i}\beta_1 + u_{1i}$$

Where y_{1i} is a binary variable with properties

$$(2) \qquad y_{1i} = 1 \ \ if\ employ = 1, \qquad y_{1i} = 0 \ \ otherwise$$

X are variables that determine the employment status. Here I use year of education, age and number of children in the first stage estimation. After getting the estimated $\hat{\beta}_1$ from the first stage, we can obtain the Inverse Mill's Ratio, and apply the estimated IMR in the final equation. The estimated IMR could be calculated by the formula:

$$(3) \qquad \widehat{IMR} = \frac{\phi(x'_{1i}\widehat{\beta}_1)}{\Phi(x'_{1i}\widehat{\beta}_1)}$$

In the estimation of wage equation, there is a common endogeneity problem, which is the omitted ability. The estimated coefficients of OLS model will be unbiased only when the schooling and labor market ability are uncorrelated (Blackburn & Neumark, 1991; Belzil & Hansen, 2002). However, the ability affects both schooling and wage. Since it's unobserved the effect of ability is involved in the error term, causing

$$(4) \qquad \text{cov}(\text{edu}_i, \varepsilon_i) \neq 0$$

To identify the "true" effect of schooling, I use interruption of schooling as to instrument education status. This is a binary variable, which is defined as "Did you ever interrupt (have a gap in) your studies for one academic year or more?". On the one hand, an individual who interrupt the schooling may has lower desire for schooling and tends to stop the education earlier (Belzil & Hansen, 2002). On the other hand, the individuals who have interrupted schooling have the same rate of return in the labor market as those who never interrupted schooling (Marcus, 1984). Therefore, the interruption of schooling has no correlation with earnings, but is related to the year of education. The equation of $Year\ edu\ act_i$ identification could be written as:

$$(1) \qquad Year\ \widehat{edu}\ act_i = f(z_i)$$

Where z_i denotes the indicator of interruption. Follow the idea from Semykina & Wooldridge (2005, 2010), which allows the presence of both endogeneity and selection, I use the estimated value of $Year\ edu\ act$ and Invers Mill's Ratio (IMR) as instruments in the final equation:

$$(2) \qquad \ln(hourly\ earning)_i = \beta_0 + \beta_1 \,\text{Max}\, parents\ edu_i + \beta_2 Year\ \widehat{edu}\ act_\iota + \beta_3 age_i + \beta_4 age_i{}^2 \beta_5 tenure_i + \beta_6 gender_i + \beta_7 training_i + \beta_8 public_i + \beta_9 certificate_i + \beta_{10} marriage_i + \beta_{11} urban_i + \sigma \widehat{IMR} + \varepsilon_i$$

where $Year\ \widehat{edu}\ act_\iota$ and $\widehat{IMR}$ is estimated from the previous stages.

2.5. Results

Table 2.4 goes about here

The result in the first column is the estimated coefficients with OLS and country fixed effect. Over all, parental education has positive effect on children's hourly earnings. Compare with pre-primary education, the children's earnings of parents with primary education or first stage of basic education is not significantly different. However, parents with lower or upper secondary education increase children's hourly earnings by 7.36%. The effect of post-secondary or higher educational parents on children's earning is greater than the lower level, leading 20.5% increase in children's earning. Effect of the individual's educational year is positive, where one more year of education, increases hourly earnings by 7.16%.

Column 2 in Table 2.4 provides the estimated results with the correction of omitted ability bias. After instrumenting the year of education with school location, the effect of parental education becomes larger than the estimation with OLS. Compare with pre-primary education, parents with primary education or first stage of basic education increase children's hourly insignificantly; parents with lower or upper secondary education increase children's hourly earnings by 42.2%. The effect of parents with post-secondary or higher education is the largest, which increase children's earnings by 75.7%. However, after correcting the ability bias, the effect of children's education insignificant. I use the Hausman Test to check for the endogeneity of educational year. The p-value is 0.0000 which indicates the endogeneity does exist.

Column 3 is the estimated results with the correction of omitted ability bias and selection bias, where the standard errors are obtained by using bootstrap. The coefficient of the inverse mills' ratio is significant, which indicates there is self-selection problem in the survey data. The effect of parental education on children's hourly earnings is positive. However, the coefficients are slightly larger than column 2. Compare with pre-primary education, parents with primary education or first stage of basic education increase children's hourly insignificantly; parents with lower or upper secondary education increase children's hourly earnings by 55.4%; parents with post-secondary or higher education increase children's hourly earnings by 94.0%. The effect of educational year is not significant.

Overall, the high educational level of parents increases children's hourly earnings. Holding other conditions constant, female have lower earnings than male. Experience, training, certificate and living in urban have positive effect on earnings. Individuals in public sectors earn more than private sectors. And marriage has positive effect on the personal earnings.

2.6. Conclusion

In this paper, I investigate the impact of parents' education on children's income. Many studies claim that education level is a key factor that influences income, while parental education also have spillover effects on their children's earning. Parental education is likely a proxy for networks of parent's and strength of their connections These phenomenons reflect the problem of intergenerational mobility and social inequality. In this paper, I use maximum parents' educational level as independent variable to find the relationship with children's hourly earnings. The data that applied in the estimation are from STEP (Skills Toward Employment and Productivity) Measurement data. This household survey provides information on household demographic characteristics, education and training, employment, behavior and preferences, family background, and some other characteristics from 12 developing countries. Considering the omitted ability bias in the wage equation, I use instrumental variable to identify the educational year with two-stage least square estimation. Hackman selection model was also applied in the estimation to fix selection bias. The results with Heckman and IVs show that parents' educational level has negative effects on children's earnings. The result shows the parental educational level has positive effects on children's hourly earnings, which is consistent with previous researches. This result also shows that the intergenerational inequality exists in those developing countries. Parental networks and social resources do have positive effect on children's earnings.

For policy makers, these effects should be considered in educational policies. To help the upward mobility of poor families, government should make education more accessible to the children at the bottom of the society. Besides educational equity, labor market equity is also important. For example, government sectors should provide more training programs to the individuals who have low educational level. This may help them to be more skilled and satisfy the requirement of the employees. The employees should also be encouraged to concentrate more on the personal characteristics but not family background.

Table 2.1. Data Definition

Variable Name	Definition
ln_earnings_h_usd	USD log of Hourly earnings
years_educ_act	Number of years of education
max_parent_educ	Maximum of parents' education
age	Age in years
tenure	Number of years in current job
training	Participated in a training course in last 12 months. 1: yes, 0: no.
pub_emp	Public or private sector employee. 1: Pubic, 0: private.
certificate	An industry-recognized or government certificate. 1: has, 0: no.
has_spouse	Has a spouse. 1: yes, 0: no.
Children	Number of children under 6 years old
Urban	1: Urban, 0: rural.

Table 2.2. Data Summary

Variable	Obs	Mean	Std.	Min	Max
ln_earning	16,515	0.9392	1.0224	-7.708	6.5461
max_parent_edu	32,517	1.689	1.015	0	3
years_educ	31,975	11.211	4.0218	0	28
age	32,517	36.857	13.737	15	64
tenure	18,650	7.8256	9.0248	0	70
gender	32,517	0.5984	0.4902	0	1
training	32,516	0.0882	0.2835	0	1
pub_emp	11,854	0.3771	0.4847	0	1
certificate	32,517	0.1103	0.3133	0	1
has_spouse	32,511	0.6023	0.4894	0	1
children	32,517	0.3739	0.6544	0	7
urban	32,517	0.932	0.2517	0	1
interrupt	32,232	0.3874	0.4872	0	1

Table 2.3. Average Values from Different Countries

Log hourly earnings	Obs	Mean	Std. Dev.	Min	Max
All	16,515	0.9392	1.0224	-7.708	6.5461
Armenia	947	0.9624	0.7002	-1.0334	6.294729
Bolivia	1,474	1.1296	1.0578	-3.925	6.546106
Colombia	1,558	1.1153	0.9607	-6.8266	5.2539
Georgia	890	1.0453	0.8603	-1.8094	5.1902
Ghana	1,180	0.4218	1.253	-4.8936	5.2818
Kenya	2,148	0.5947	1.1025	-3.4731	5.5629
Laos	1,296	0.4125	1.2753	-3.8329	5.287
Macedonia	1,608	1.5086	0.6408	-2.4658	4.8962
Sri Lanka	1,212	0.9153	1.0001	-2.8898	5.4649
Ukraine	1,016	1.1129	0.6547	-2.9573	4.1431
Vietnam	1,979	1.0422	0.9856	-7.708	6.2402
Yunnan (China)	1,207	1.018	0.7271	-2.1203	5.1815
Max parents educational	Obs	Mean	Std. Dev.	Min	Max
All	32,517	1.689	1.015	0	3
Armenia	2,896	2.4541	0.5889	1	3
Bolivia	2,231	1.4939	1.2197	0	3
Colombia	2,451	1.4651	0.7938	0	3
Georgia	2,971	2.6314	0.5588	0	3
Ghana	1,893	2.0032	0.7241	0	3
Kenya	3,735	1.5015	1.0741	0	3
Laos	2,217	0.7208	1.0157	0	3
Macedonia	3,987	1.4678	0.9811	0	3
Sri Lanka	2,688	1.6741	0.6318	0	3
Ukraine	2,346	2.2899	0.7174	0	3
Vietnam	3,135	1.2801	0.9237	0	3
Yunnan (China)	1,967	1.1886	0.824	0	3
Years of esucation	Obs	Mean	Std. Dev.	Min	Max
All	31,975	11.211	4.0218	0	28
Armenia	2,896	12.747	2.7899	0	21
Bolivia	2,231	11.473	4.2683	0	23
Colombia	2,451	10.193	3.7333	0	20
Georgia	2,454	14.197	3.3164	0	28
Ghana	1,893	9.4031	3.8049	0	19
Kenya	3,716	10.41	4.2319	0	22
Laos	2,217	7.8705	5.0161	0	23

Macedonia	3,984	12.037	3.4614	0	25
Sri Lanka	2,686	9.4702	3.1343	0	20
Ukraine	2,345	12.912	2.2553	0	22
Vietnam	3,135	11.168	3.9202	0	20
Yunnan (China)	1,967	11.963	3.6054	0	20
Age	Obs	Mean	Std. Dev.	Min	Max
All	32,517	36.857	13.737	15	64
Armenia	2,896	39.324	14.157	15	64
Bolivia	2,231	32.522	12.963	15	64
Colombia	2,451	35.74	13.774	15	64
Georgia	2,971	39.545	14.034	15	64
Ghana	1,893	30.506	10.94	15	64
Kenya	3,735	29.452	9.8587	15	64
Laos	2,217	35.35	13.382	15	64
Macedonia	3,987	40.419	14.079	15	64
Sri Lanka	2,688	37.565	13.362	15	64
Ukraine	2,346	42.037	14.476	15	64
Vietnam	3,135	38.027	13.765	15	64
Yunnan (China)	1,967	41.114	11.616	15	64
Tenure	Obs	Mean	Std. Dev.	Min	Max
All	18,650	7.8256	9.0248	0	70
Armenia	1,013	9.5114	9.9631	0	45
Bolivia	1,637	6.0554	7.9344	0	48
Colombia	1,638	0.1357	0.0403	0.0833	0.1667
Georgia	952	8.5664	9.6232	0	50
Ghana	1,302	6.0735	6.884	0	40
Kenya	2,330	4.5583	5.0928	0	44
Laos	1,861	11.103	10.043	0	53
Macedonia	1,845	11.561	10.435	0	43
Sri Lanka	1,479	10.097	10.093	0	50
Ukraine	1,148	9.6755	9.1867	0	44.167
Vietnam	2,182	9.3878	8.8807	0	70
Yunnan (China)	1,263	8.6914	8.8921	0	45
Gender	Obs	Mean	Std. Dev.	Min	Max
All	32,517	0.5984	0.4902	0	1
Armenia	2,896	0.7169	0.4506	0	1
Bolivia	2,231	0.5746	0.4945	0	1
Colombia	2,451	0.5814	0.4934	0	1
Georgia	2,971	0.6728	0.4693	0	1
Ghana	1,893	0.5647	0.4959	0	1

Kenya	3,735	0.5224	0.4996	0	1
Laos	2,217	0.6085	0.4882	0	1
Macedonia	3,987	0.537	0.4987	0	1
Sri Lanka	2,688	0.6217	0.4851	0	1
Ukraine	2,346	0.6645	0.4723	0	1
Vietnam	3,135	0.5946	0.4911	0	1
Yunnan (China)	1,967	0.5445	0.4981	0	1
Training	Obs	Mean	Std. Dev.	Min	Max
All	32,516	0.0882	0.2835	0	1
Armenia	2,896	0.0625	0.2421	0	1
Bolivia	2,231	0.195	0.3963	0	1
Colombia	2,451	0.1624	0.3689	0	1
Georgia	2,971	0.0778	0.2678	0	1
Ghana	1,893	0.0824	0.2751	0	1
Kenya	3,735	0.1114	0.3146	0	1
Laos	2,217	0.0514	0.2209	0	1
Macedonia	3,987	0.0815	0.2737	0	1
Sri Lanka	2,688	0.0629	0.2428	0	1
Ukraine	2,345	0.0188	0.1357	0	1
Vietnam	3,135	0.0558	0.2296	0	1
Yunnan (China)	1,967	0.1134	0.3171	0	1
Public sectors	Obs	Mean	Std. Dev.	Min	Max
All	11,854	0.3771	0.4847	0	1
Armenia	891	0.6689	0.4709	0	1
Bolivia	876	0.2089	0.4068	0	1
Colombia	943	0.0923	0.2895	0	1
Georgia	801	0.5443	0.4983	0	1
Ghana	545	0.3321	0.4714	0	1
Kenya	1,305	0.1226	0.3281	0	1
Laos	524	0.5821	0.4937	0	1
Macedonia	1,504	0.4335	0.4957	0	1
Sri Lanka	843	0.3879	0.4876	0	1
Ukraine	1,285	0.4016	0.4904	0	1
Vietnam	1,253	0.4677	0.4992	0	1
Yunnan (China)	1,084	0.4068	0.4915	0	1
Certificate	Obs	Mean	Std. Dev.	Min	Max
All	32,517	0.1103	0.3133	0	1
Armenia	2,896	0.0735	0.2611	0	1
Bolivia	2,231	0.2039	0.403	0	1
Colombia	2,451	0	0	0	0

Georgia	2,971	0.1097	0.3126	0	1
Ghana	1,893	0.084	0.2775	0	1
Kenya	3,735	0.0768	0.2664	0	1
Laos	2,217	0.0424	0.2015	0	1
Macedonia	3,987	0.1242	0.3298	0	1
Sri Lanka	2,688	0.0707	0.2563	0	1
Ukraine	2,346	0.0678	0.2514	0	1
Vietnam	3,135	0.251	0.4337	0	1
Yunnan (China)	1,967	0.215	0.411	0	1
Spouse	Obs	Mean	Std. Dev.	Min	Max
All	32,511	0.6023	0.4894	0	1
Armenia	2,896	0.6136	0.487	0	1
Bolivia	2,231	0.4818	0.4998	0	1
Colombia	2,451	0.4537	0.498	0	1
Georgia	2,970	0.5892	0.4921	0	1
Ghana	1,892	0.3879	0.4874	0	1
Kenya	3,732	0.5206	0.4996	0	1
Laos	2,217	0.6721	0.4696	0	1
Macedonia	3,987	0.6353	0.4814	0	1
Sri Lanka	2,688	0.7128	0.4525	0	1
Ukraine	2,346	0.7084	0.4546	0	1
Vietnam	3,134	0.6394	0.4802	0	1
Yunnan (China)	1,967	0.8053	0.3961	0	1
Number of children	Obs	Mean	Std. Dev.	Min	Max
All	32,517	0.3739	0.6544	0	7
Armenia	2,896	0.3339	0.6292	0	4
Bolivia	2,231	0.4868	0.7533	0	7
Colombia	2,451	0.3333	0.6097	0	6
Georgia	2,971	0.3329	0.6276	0	6
Ghana	1,893	0.523	0.7985	0	5
Kenya	3,735	0.5116	0.7244	0	4
Laos	2,217	0.5011	0.7116	0	4
Macedonia	3,987	0.296	0.6289	0	6
Sri Lanka	2,688	0.4185	0.6162	0	3
Ukraine	2,346	0.2272	0.4999	0	3
Vietnam	3,135	0.3834	0.6551	0	4
Yunnan (China)	1,967	0.1256	0.339	0	2
Urban	Obs	Mean	Std. Dev.	Min	Max
All	32,517	0.932	0.2517	0	1

Armenia	2,896	1	0	1	1
Bolivia	2,231	1	0	1	1
Colombia	2,451	1	0	1	1
Georgia	2,971	1	0	1	1
Ghana	1,893	1	0	1	1
Kenya	3,735	1	0	1	1
Laos	2,217	0.7325	0.4427	0	1
Macedonia	3,987	1	0	1	1
Sri Lanka	2,688	0.3984	0.4897	0	1
Ukraine	2,346	1	0	1	1
Vietnam	3,135	1	0	1	1
Yunnan (China)	1,967	1	0	1	1
Interrupt	Obs	Mean	Std. Dev.	Min	Max
all	32,232	0.3874	0.4872	0	1
Armenia	2,896	0.0401	0.1961	0	1
Bolivia	2,213	0.23	0.4209	0	1
Colombia	2,436	0.2931	0.4553	0	1
Georgia	2,971	0.171	0.3766	0	1
Ghana	1,893	1	0	1	1
Kenya	3,735	0.9989	0.0327	0	1
Laos	2,034	0.088	0.2834	0	1
Macedonia	3,987	0.9992	0.0274	0	1
Sri Lanka	2,665	0.1114	0.3147	0	1
Ukraine	2,344	0.0299	0.1702	0	1
Vietnam	3,111	0.1353	0.3421	0	1
Yunnan (China)	1,947	0.0329	0.1783	0	1

Table 2.4. Estimated Results of Parental Education on Children's Hourly Earnings

	OLS ln_earnings_h_usd	IV ln_earnings_h_usd	Heckman with IV ln_earnings_h_usd
1.max_parent_educ	-0.00936	0.187	0.235
	(0.0249)	(0.132)	(0.166)
2.max_parent_educ	0.0736***	0.422*	0.554*
	(0.0250)	(0.234)	(0.316)
3.max_parent_educ	0.205***	0.757**	0.940*
	(0.0285)	(0.372)	(0.487)
4.max_parent_educ	0.0716***	-0.0535	-0.0589
	(0.00222)	(0.0851)	(0.101)
years_educ_act	0.0176***	0.0103*	0.0101*
	(0.00211)	(0.00539)	(0.00594)
age	0.0381***	0.0509***	-0.0361
	(0.00424)	(0.0102)	(0.0317)
age2	-0.000464***	-0.000642***	0.000459
	(0.0000528)	(0.000139)	(0.000398)
tenure	0.00878***	0.0105***	0.0109***
	(0.000961)	(0.00164)	(0.00181)
gender	-0.220***	-0.211***	-0.00256
	(0.0140)	(0.0172)	(0.0862)
training	0.167***	0.352***	0.359**
	(0.0198)	(0.127)	(0.152)
pub_emp	0.0636***	0.279*	0.288*
	(0.0164)	(0.146)	(0.174)
certificate	0.0618***	0.177**	0.0631
	(0.0198)	(0.0805)	(0.0758)
has_spouse	0.0565***	0.0601***	0.0629***
	(0.0162)	(0.0186)	(0.0206)
urban	0.0288	0.175	0.182
	(0.0469)	(0.116)	(0.134)
Bolivia	0.390***	0.458***	0.462***

	(0.0375)	(0.0639)	(0.0646)
Colombia	0.606***	0.560***	0.558***
	(0.0374)	(0.0530)	(0.0572)
Georgia	0.0260	0.210	0.217
	(0.0376)	(0.132)	(0.166)
Ghana	-0.172***	-0.317***	-0.326**
	(0.0411)	(0.110)	(0.159)
Kenya	0.0176	0.000555	-0.00314
	(0.0348)	(0.0417)	(0.0519)
Laos	0.119***	0.195***	0.202**
	(0.0431)	(0.0735)	(0.0898)
Macedonia	0.666***	0.776***	0.776***
	(0.0329)	(0.0828)	(0.0967)
Sri Lanka	0.306***	0.111	0.101
	(0.0471)	(0.143)	(0.176)
Ukraine	0.261***	0.310***	0.310***
	(0.0347)	(0.0514)	(0.0582)
Vietnam	0.279***	0.332***	0.328***
	(0.0344)	(0.0526)	(0.0567)
Yunnan (China)	0.141***	0.279***	0.286**
	(0.0360)	(0.104)	(0.119)
mills			-0.758**
			(0.304)
_cons	-1.119***	-0.352	1.513
	(0.0985)	(0.528)	(1.267)
N	10862	10827	10827

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3 Effect of Insurance Subsidies on Agricultural Land Use

3.1. Introduction

Land is an important resource for agricultural production and urban development. Generally, land use change is driven by climate, geographic characteristics, market factors and government policy. Land use has a significant environmental effect on local regions. Land use change affects biodiversity, water quality, Green House Gas (GHG) emission and so on. Searching et al. (2008) found the growth of cropland for biofuels increase GHG emission in the U.S. for 167 years. Agricultural land use also impacts soil dust emission (Tegen et al., 2004) and climate change (Pielke, 2005). Besides the environmental effects, agricultural land use also influences supply of agricultural products.

Government policies play an important role in land use allocation, especially agricultural land use. For example, an agricultural subsidy on crop production encourages farmers to plant more crops. With subsidy, crop production generates more net return to farmers, and farmers are encouraged to extend the acreage of cropland to increase total production. Besides direct subsides, policies on other agricultural products, such as biofuels, will also influence agricultural land allocation. The extension of biofuel stimulates the demand for ethanol, which makes corn more profitable. As a result, corn land increases. Motamed et al. (2016) found in the United States, corn area had a significant and large response to local ethanol markets.

The objective of this paper is to investigate the effect of crop insurance subsidies on agricultural land use. In the U.S., crop insurance is an important instrument for farmers to control the risk from natural disaster and market volatility. The Federal Crop Insurance Corporation (FCIC) was created in 1938, which focused on some major crops in a few regions. In 1980, the government expanded the species and regions covered by Federal Crop Insurance. In order to encourage crop insurance participation, government improved the crop insurance program with greater subsidy levels in 1994 and 2000. There are two basic types of crop insurance—yield based and revenue based, that guarantee crop production based on historical yields and prices. The Federal Crop Insurance Program helps farmers in risk management. Under the assumption of risk averse, lower risk increases the expected utility from crop production profit. The higher expected utility of crop production profit encourages farmers to extend farmland to plant more crops. Schatzki (2003) focuses on the effects of uncertainty and sunk costs on land use change. He finds that higher uncertainty in returns to potential use will decrease the likelihood of conversion from agriculture to forest. Moreover, crop insurance subsidies reduce the cost of insurance premium. Lower insurance cost will also increase expected total profit, encouraging farmers to plant more crops. Therefore, investigating the effect of crop insurance subsidies on agricultural land use is important: for policy makers, crop insurance subsidies could be considered a tool to adjust agricultural land use; the change of land use allocation will influence agricultural product supply and environment quality in local areas.

To find the effects of insurance subsidies on agricultural land use, I apply the Multinomial Fractional Logit Model with county level data from all continental states in the U.S. In this paper, I classify land use type in two ways. The first way is following the classification in Hardie and Parks (1997), which classified farmland into irrigated and unirrigated farmland. The second way is to classify farmland more specifically, such as cropland, woodland, and pastureland. The contribution of this paper is using the Multinomial Fractional Logit Model to investigate the effect of insurance subsidies. Based on the single outcome model, which focused on crop area in Yu et al. (2016), I classified land use into different types. By considering the effects of insurance subsidies on multiple land uses within the same Fractional Response Model, we are able to indirectly control for the substitution between various land use alternatives.

The paper has five sections. The next section is a literature review, which provides previous studies in land use and effects of government policies. The following part is data description. Then I discuss the theoretical framework and econometric model that would be used. The forth section contains the results of estimation with agricultural census data and discussion. The final section is the conclusion.

3.2. Literature review

Generally, the usage of land is depended on the spatial characteristics, demographic characteristics, and land quality. Besides, land use is also influenced by market demand and government policy (Lambin et al., 2001; Veldkamp & Lambin, 2001).

Lubowski et al. (2008) focus on the factors that drive land-use change. They analyze change in the U.S. land-use between 1982 and 1997 and consider the net returns as the drivers of land-use change. In their model, the factors from both supply and demand sides are included. The results show that the private land-use decisions were dependent on land quality, economic returns, and public policies. Veldkamp and Fresco (1996) use a theoretical model to study land use. They claim that land use change depended on both biophysical and human demands. Typical biophysical drivers are biophysical suitability, land use history, and spatial distribution of infrastructure. Important human land use drivers are population density, regional technology level, economical conditions, attitude and values. Newburn et al. (2004) claim that the site-selection of land-use is influenced by three important factors: biological benefits, land cost, and likelihood of land-use change. After comparing different targeting strategies, they find that the relationship between economics and land use change is important. Wang et al. (2014) develop a spatial autoregressive multinomial probit model to analyze land development decisions, including spatial clustering and cross-alternative correlation. The explanatory variables include parcel area, parcel perimeter-to-area ratio, network distances and soil slope. The results show that distance to CBD has positive effect on the likelihood of residential development.

The research that related to the effect of insurance subsidies on land use is limited. Yu et al. (2016) investigate the effects of crop insurance premium subsidies on crop acreage. In their study, crop insurance premium subsidies affect crop acreage in two ways. The first way is by increasing expected returns, the second is by reducing riskiness of crop. These two ways will encourage farmers to increase crop acreage. The results show that a 10% increase in the crop insurance premium subsidy increase crop acreage by 0.43%. However, Yu et al. (2016) only considered the effect of land use on crop area, but not other types on farmland. To address the question "why do we subsidize crop insurance?" Coble and Barnett (2012) consider the contributions made by subsidies to the policy objectives. They explain the mechanism behind premium subsidies and outlined some potential problems caused by the insurance subsidies. In the end, the authors suggest four research topics. One of which is to study the resource allocation decisions affected by insurance subsidies. Wu (1999) estimates the effect of crop insurance in the Central Nebraska Basin. He finds that crop insurance would increase chemical use by shifting land from hay and pasture to corn. Goodwin et al. (2004) use multi equation structural models to analyze the acreage effect of Federal Crop Insurance Program. The response of crop acreage change was significant. In some cases, 30% decrease in premiums, which implied the increase in subsidies, would lead crop acreage increase from 0.2% to 1.1%. Young (2001) investigate d the effect of crop insurance on farmers' planting choices. The simulation results show that crop insurance subsidies has positive effect on aggregate planting, especially for wheat and cotton. Over all, previous studies on insurance subsidies and land use focused on the acreage effect on cropland, but the proportion of allocation and effects on other farm land have not been mentioned.

There are also many studies about other factors that impact agricultural land use. Miao (2013) applies a logit land-share model by using panel data from 1997 to 2009. He finds that corn-based ethanol plants have a significant effect on the proportion of land planted with corn. Plantinga et al. (2002) develop a spatial city model to estimate agricultural land price. In their study, land prices reflect not only current uses of land, but also potential uses. Therefore, the current land value reported is influenced by agricultural production rents and rents from future land development. Mann et al. (2010) focus on the effect of agricultural rent on cropland conversion in Amazonia. The results show that besides transportation cost, the expected returns from the venture also affect agricultural expansion. Lichtenberg (1989) focus on the effect of land quality on land use, crop choice, and technological change. He finds that quality of land had a significant impact on cropland allocation. Technologies tend to be applied primarily on land with low quality. Otherwise, the irrigation was sensitive to tax policies.

Hardie and Parks (1997) conduct research on southeastern land use. They apply an Area Base Model and analyze the effect of land quality on the probability of different types of land use. The variables used in their paper are: crop revenue, crop cost, saw-timber price, pulpwood price, timber cost, land class which reflects land quality, population density and per capita income. They also use discrete explanatory variables to evaluate the land use in different regions. The analysis of that paper is based on rent-maximization hypothesis, and use variables such as costs and prices as economic characteristics. However, they do not consider the influence of government action, which could also be an important factor affecting land use allocation.

3.3. Data

The data about land use acres and agricultural market are collected from the USDA Census of Agriculture. The Census of Agriculture provides county level data of farm land acreage, land cash rent, farmers' net income, and agricultural production expenditures. According to the definition and explanation from the Census of Agriculture, the category of irrigated land includes all land watered by any artificial or controlled means. For approximate land area, the proportion of farmland or cropland area may exceed 100%, since some of the operations have land in two or more counties, while all acres are reported in the principal county of operation. As the Multinomial Fractional Logit Model that was applied in this paper requires the outcomes to be fractions in [0, 1] interval, I dropped land shares that exceeded 100%. The land share data that applied in the estimation sum up to unity for each county-year by construction. The category of total cropland includes harvested cropland, abandoned cropland, other pasture and grazing land that could be used for crops without any improvement, and land used for cover crops. Total woodland includes natural or planted woodlots, cutover, deforested land, and pastured woodland. Net cash income is derived by subtracting total farm expenses from total sales, government payments, and other farm-related income. The census data are reported at an interval of every five years. In this paper I use the data from the following years 2002, 2007 and 2012, as they are the three most recent reported years. The data contain counties in all continental states.

The data of crop insurance information are collected from Risk Management Agency (RMA) and Summary of Business (SOB). These data sources provide county level data of crop insurance premium, subsidies from government, coverage level and other information. The data are reported by different species, delivery type, coverage levels and other details. Therefore, I construct county level data by summing up subsidy amount from those subgroups. The average values of crop market value and subsidies for crop insurance are constructed by total value and cropland acreage, while the average values of farmer expenses and net income come from dividing the total value by farmland acreage.

Table 3.1 goes about here.

Table 3.1 shows the summary of data that were used in the estimation. Other 1 is calculated by subtracting irrigated and unirrigated farmland share from one, while other 2 is calculated by subtracting cropland, woodland and pastureland share from one. From the table, we can find that the average irrigated farmland share is 1.36% in 2007, while average unirrigated farmland share is 10.83%. Both of them increased from 2002 to 2012. For more detailed land use type, cropland has the largest average land share, which is 28.08%, while pastureland share is 18.14% in 2007. However, the cropland share decreased during these 10 years. The acreage of woodland has the smallest average share, which is 6.37%. The average cash rent is $76.808 per acre in 2007 and increased from 58.07 to 110.97. The average crop insurance subsidy increased a lot, which from 3.43 to 15.5. In the estimation, all dollar-valued regressors are logged to get the effects with percentage change.

3.4. Model

In this paper, the use of land is classified in two ways. First is following the classification from Hardie and Parks (1997). Hardie and Parks (1997) claim that irrigation is important. Therefore, I will classify farmland use into irrigated farm land and unirrigated farm land. To investigate the effect of insurance subsidies on more specific land use, the second classifications of land use type is cropland, woodland, pasture land, and other land. The profit of a particular land use depends on its cost and revenue. For example, as the cost of crop production decreases or revenue of crop increases, the profit from crop production increases. With higher profit, land will be more likely to be used for crop production. As the total acreage of each county is constant, the purpose of land allocation is maximizing total profit from all types of land. Therefore, the share of land in use j could be represented as a function of independent variables:

$$(1) \qquad p_j = P(sub_j, z_{jk})$$

where p_j is the proportion of land use j. The independent variables should be related to revenue and cost, including crop insurance subsidies, farmers net income, labor expenses, fertilizer expenses, chemical expense, cash rent of land, and some dummy variables.

I focus on the effect of insurance subsidies, with the hypothesis: insurance subsidies would increase the proportion of farmland. The objective of Federal Crop Insurance Program is to help farmers reduce risk from natural disasters and market volatility. Farms that enrolled in crop insurance programs are supposed to have lower risk. Crop insurance subsidies could improve the participation rate; moreover, crop insurance subsidies also reduce farmers cost of insurance premium, and increase expected returns. With lower risk and higher expected returns, farmers are encouraged to increase crop production. On the other hand, insurance can affect farmers' behavior due to moral hazard and adverse selection (Quiggin et al., 1993). Insurance subsidies may have the similar effects. Therefore, with insurance subsidies, marginal land and land with less profitable usage are likely to be transferred into farmland to achieve high crop production. Marginal land is the land that is sensitive to the demand of different land uses. Then hypothesis of this paper could be raised: the insurance subsidies increase farmland proportion, especially farmland in crop production. Land proportion with less profitable usage will be decreased.

According to the current and previous research works, logit/probit model is widely popular to investigate land use conversion. Based on the theory of maximizing net benefit, Carrion-Flores and Irwin (2004) estimated residential land conversion using probit model. They used parcel-level data and spatial statistics and found that urban development was affected by preferences for lower density area.

In this paper, I use the Multinomial Fractional Logit (MFLOGIT) Model. MFLOGIT Model is used when the outcomes are fractional variables, such as rates and proportions. For example, Mullahy and Robert (2010) applied this model to the time budget problem. They explored how people with different education levels allocated time to physical activities, where the allocation of time is a sort of fractional outcome. Papke and Wooldridge (1996) introduced the quasi-maximum likelihood estimator (QMLE) to avoid the wrong distribution assumption, leading to a relatively efficient estimator. Based on Papke and Wooldridge (1996), Mullahy (2010) discussed the application of this model on economic share data outcomes. He extended the univariate fractional regression from Papke and Wooldridge (1996) to the Multivariate Fractional Logit Model. For the application in land use, Molowny-Horas et al. (2015) applied this model to investigate the effect of nature forces and landscape on land use. Based on the fractional regression study by Papke and Wooldridge (1996), they used multivariate data of Barcelona province, Spain. The results showed that the land use was not only influenced by geographical and environmental variables, but by the neighboring landscape.

Following Mullahy (2015), the Multinormial Logit Functional form is

$$(2) \qquad \mathrm{E}[y_{im}|x] = \frac{\exp(x_i\beta_m)}{\sum_{k=1}^{M}\exp(x_i\beta_k)} = \frac{\exp(x_i\beta_m)}{1+\sum_{k=2}^{M}\exp(x_i\beta_k)}, m = 1, \dots, M$$

With normalization $\beta_1 = 0$. Where x_i represents the independent variables that affect land use, m represents land use type, y_{im} represents share of the mth land use.

The average partial effects (APE) for continuous variables is

$$(3) \qquad \widehat{APE}_{mk} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial E[y_{im}|x_i]}{\partial x_{ik}}$$

Where

$$(4) \qquad \frac{\partial E[y_{im}|x_i]}{\partial x_{ik}} = \exp(x_i\beta_m) * \frac{(1+\sum_{k=2}^{M}\exp(x_i\beta_j))*\beta_{mk}-\sum_{k=2}^{M}\exp(x_i\beta_j)*\beta_{jk}}{(1+\sum_{k=2}^{M}\exp(x_i\beta_k))^2}$$

Consider the potential endogeneity problem that the increase of farmland may cause the increase in total insurance subsidies, I use the average subsidies per acre instead of total subsidies. I also use average values per acre for other independent variables to avoid endogeneity problem. However, the endogeneity problem still exists, since increasing farmland share may decrease average cash rent. As all the explanatory variables are predetermined, I estimate the model with all regressors replaced with their lags.

In the estimation, I apply the district fixed effect. The agricultural district is classified into nine categories, which are northwest and mountain, north central, northeast, west central, central, east central, southwest, south central, and southeast.

3.5. Results

Table 3.2 goes about here.

Table 3.2 and Table 3.3 provide the estimated coefficients in the Multinomial Fractional Logit Model with either year fixed effect or agricultural district fixed effect. The standard errors are clustered at state level. Table 3.2 shows that the effect of insurance subsidies on both irrigated land and unirrigated land is significant and positive. The positive coefficients indicate as premium subsidies increase, the probability of changing other land into irrigated/unirrigated farmland increases. The effects of average net income on irrigated and unirrigated allocation are not significant. The effects of agricultural production expenses such as labor expense, fertilizer expense and chemical expense on other land use are positive and significant, which indicate that as those expenses increase, farmers' net income will decrease, and land are less likely to be transferred into irrigated and unirrigated farmland.

Table 3.3 goes about here.

For classification of cropland, woodland and pasture land (Table 3.3), the influence of insurance subsidies on cropland and pasture land are significant with year fixed effect. As the amount of insurance subsidies increase, the proportion of cropland and pastureland increases. This result is consistent to the hypothesis, which is the insurance subsidies from government encourage crop production, and increase cropland share. However, the effect of insurance subsidies on woodland land share is not significant. Average cash rent per acre has a negative effect on woodland share but a positive effect on cropland, indicating as average cash rent increase, land is more likely to be transferred from woodland land into cropland. This result might be due to high profit from cropland. The effect of labor expense on cropland share is negative, but woodland share is positive. Since as farm labor expense increase, land owners may intend to abandon some of agricultural production and transfer cropland into woodland and other land. Labor and fertilizer expenses have positive effects on other land with both year and district fixed effects. The increasing agricultural expenses lead other land share increases.

Table 3.4 and Table 3.5 goes about here.

Table 3.4 and Table 3.5 show the estimated results with lagged variables. The results are similar with the estimation with un-lagged variables, but the significant levels are increased. Crop insurance subsidies have positive effects on irrigated/unirrigated farmland, cropland and pastureland, while the effects on other land are negative.

Table 3.6 goes about here.

Table 3.6 and

Table **3.7** show the average marginal effect estimated in the Multinomial Fractional Logit Model with both un-lagged and lagged variables. Signs of coefficients coincide with sighs of the corresponding marginal effects. For the effect of insurance subsidies on irrigated farmland and unirrigated farmland, a 10% increase in average insurance subsidies per acre increase irrigated farmland share by 3.127 percentage points, and unirrigated farmland share by 0.229 percentage points. The marginal effect of insurance subsidies on unirrigated farmland is smaller than that on irrigated farmland. The marginal effects of average cash rent and agricultural expenses are also significant. As cash rent per acre increase 10%, the unirrigated farmland share increase by 0.245 percentage points, and the other land share decreases by 0.239 percentage points. Average cash rent has positive marginal effect on crop land share and negative marginal effects on woodland share. For the effect of labor expense, a 10% increase in expense decreased cropland share by 1.046 percentage points.

Table **3.7** goes about here.

Table **3.7** shows the estimated average marginal effects with lagged variables. The marginal effects that estimated with lags are more reasonable than that without lags. As the crop insurance subsidies increase 10%, shares of irrigated and unirrigated farmland increase by 0.06 and 0.206 percentage points, respectively. With classification 2, lagged average subsidies increase 10%, cropland and pastureland share increase 0.265 and 0.292 percentage points respectively, while woodland share decrease 0.05 percentage points. Agricultural expenses such as labor expense, fertilizer expense and chemical expense also have positive average marginal effects on other land share under both classifications. The result is consistent to previous studies in crop acreage effects, which showed insurance subsidies had positive effect on crop acreage. Since the effect of insurance subsidies on both cropland and pastureland is positive, the increased cropland share may not be transferred from pastureland but other types of land. Average cash rent has positive marginal effect on cropland shares, but a negative effect on woodland shares. A 10% increase in average cash rent leads cropland shares to increase by 1.135 percentage points, and woodland share to decrease by 0.251 percentage points. It indicates that cropland is likely to have higher cash rent than woodland, since cropland is more profitable. Production expenditures have significantly positive marginal effect on other land share. As the operation cost increase, farmers are likely to abandon some of crop production and transfer cropland into woodland and pasture land.

3.6. Conclusion

This paper investigated the effect of insurance subsidies on agricultural land use. As agricultural land use allocation has great influence on the agricultural market and local environmental quality, policy makers should know how the policy tool such as subsidies affect land use allocation. Based on the Multinomial Fractional Logit Model, which is used with fractional outcomes, I investigate the effect of crop insurance on the allocation of irrigated farmland, unirrigated farmland, cropland, woodland, and pasture land. The estimated results provide coefficients with both year fixed effect and agricultural district fixed effect. The results with lagged variables show that insurance subsidies, as a kind of government tool, have a significant effect on agricultural land use allocation. A 10% increase in insurance subsidies increases the share of irrigated farmland, unirrigated farmland, cropland and pastureland by 0.06, 0.206, 0.265 and 0.292 percentage points, respectively. We can conclude that crop insurance subsidies have positive effect on both irrigated farmlands share and unirrigated farmland share, and the effect on unirrigated farmland is larger. For more specific classification, as insurance subsidies increase, lands other than pastureland are likely to be transferred into cropland. The results are consistent to the hypothesis, which claims crop insurance subsidies have positive effect on farmland allocation by encouraging farmers to plant more crops.

With regard to policy makers, the objective of crop insurance subsidies is to help farmers with risk management, which could reduce risk with the guarantee of crop yield or revenue. However, the effect of insurance is not only on production risk, but also on agricultural land use allocation. As the allocation of agricultural land has influence in the agricultural market and environment quality, policy makers should consider the effects on land allocation when they subsidize crop insurance. In addition, crop insurance subsidies could also be considered as an efficient way for policy makers to adjust agricultural land allocation.

Table 3.1. Data Summary

Variables of 2002	Obs	Mean	Std. Dev.	Min	Max
Irrigated land share	2,196	0.0123	0.0263	0.000045	0.2517
unirrigated land share	2,196	0.1031	0.0805	0.000097	0.4939
other1	2,196	0.8846	0.0843	0.4897	0.9996
Cropland share	2,196	0.2877	0.2392	0.002	0.9523
Woodland share	2,196	0.0669	0.0544	0.0000174	0.3118
Pastureland share	2,196	0.1753	0.1884	0.0011	0.9336
other2	2,196	0.47	0.2886	0.0012	0.9946
Average subsidy per acre	1,964	3.4341	3.7814	0	64.167
Net income per acre	2,194	372.85	955.29	-2570.2	20800
Cash rent per acre	2,169	58.07	172.1	0.1808	7371.4
Labor expense per acre	2,102	18.613	44.183	0.0471	595.13
Fertilizer expense per acre	2,192	68.087	62.2	0.0104	739.64
Chemical expense per acre	2,177	47.279	60.62	0.0104	860.44
Variables of 2007	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Irrigated land share	2,291	0.0136	0.029	0.00000732	0.2732
unirrigated land share	2,291	0.1083	0.0853	0.0000455	0.5147
other1	2,291	0.878	0.0894	0.4714	0.9997
Cropland share	2,291	0.2801	0.2478	0.0011	0.9559
Woodland share	2,291	0.0637	0.0515	0.0000465	0.2554
Pastureland share	2,291	0.1814	0.1977	0.0012	0.9693
other2	2,291	0.4749	0.2932	0.0008	0.9944
Average subsidy per acre	2,028	7.6583	6.7049	0	93.546
Net income per acre	2,287	462.93	820.77	-4655.8	10813
Cash rent per acre	2,274	76.808	101.93	1.4959	2037.6
Labor expense per acre	2,187	22.383	58.417	0.1367	1019
Fertilizer expense per acre	2,290	119.84	110.62	0.029	1532.5
Chemical expense per acre	2,265	58.53	72.623	0.0667	1161.3
Variables of 2012	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**

Irrigated land share	2,385	0.0146	0.0319	0.0000028	0.2943
unirrigated land share	2,385	0.114	0.0907	0.0000171	0.5195
other1	2,385	0.8713	0.0946	0.4552	0.9997
Cropland share	2,385	0.275	0.2566	0.0007	0.9589
Woodland share	2,385	0.0632	0.0516	0.000049	0.344
Pastureland share	2,385	0.1779	0.2081	0.0008	0.9767
other2	2,385	0.4839	0.2962	0.0001	0.9968
Average subsidy per acre	2,139	15.519	16.576	0	487.2
Net income per acre	2,382	510	1006	-6443	22116
Cash rent per acre	2,363	110.97	138.61	1.5881	1826.9
Labor expense per acre	2,309	30.643	80.35	0.1019	1829.2
Fertilizer expense per acre	2,374	172.24	175.78	0.0059	2657.8
Chemical expense per acre	2,353	90.622	135.87	0.0354	3489.4

Table 3.2. Estimated Coefficients on Land Use of Irrigated Farmland and Unirrigated Farmland in Multinomial Fractional Logit Model

	Irrigated farmland	**Unirrigated farmland**	**Other1**	**Irrigated farmland**	**Unirrigated farmland**	**other1**
Average net income	-0.00612 (0.0558)	0.0238 (0.0210)	-0.0200 (0.0181)	0.00334 (0.0527)	0.0198 (0.0206)	-0.0177 (0.0175)
Average subsidy	0.351* (0.145)	0.224*** (0.0352)	-0.243*** (0.0326)	0.292** (0.103)	0.218*** (0.0260)	-0.233*** (0.0260)
Average cash rent	-0.0315 (0.169)	0.231** (0.0766)	-0.203** (0.0668)	-0.0253 (0.168)	0.236** (0.0759)	-0.209** (0.0654)
Average labor expense	-0.110 (0.0808)	-0.323*** (0.0323)	0.299*** (0.0259)	-0.120 (0.0833)	-0.324*** (0.0305)	0.302*** (0.0244)
Average fertilizer expense	-0.524** (0.175)	-0.0466 (0.0613)	0.121* (0.0546)	-0.519** (0.159)	-0.0772 (0.0609)	0.146** (0.0549)
Average chemical expense	0.528* (0.242)	-0.242*** (0.0614)	0.144** (0.0483)	0.532** (0.197)	-0.212*** (0.0578)	0.119** (0.0452)
_cons	-4.010*** (0.285)	-1.679*** (0.179)	1.508*** (0.159)	-4.081*** (0.340)	-1.791*** (0.193)	1.621*** (0.186)
Year fe	Yes	Yes	Yes	No	No	No
Dist fe	No	No	No	Yes	Yes	Yes
N	5163	5163	5163	5163	5163	5163

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3. Estimated Coefficients on Land Use of Cropland, Woodland and Pastureland in Multinomial Fractional Logit Model

	Cropland	**Woodland**	**Pasture land**	**other2**	**Cropland**	**Woodland**	**Pasture land**	**other2**
Average net income	0.0539	-0.0120	-0.152***	0.0460	0.0380	-0.00703	-0.125***	0.0350
	(0.0320)	(0.0281)	(0.0429)	(0.0452)	(0.0275)	(0.0288)	(0.0337)	(0.0371)
Average subsidy	0.216***	-0.0502	0.274***	-0.299***	0.0293	-0.0507	0.302***	-0.195***
	(0.0416)	(0.0423)	(0.0703)	(0.0613)	(0.0371)	(0.0363)	(0.0599)	(0.0500)
Average cash rent	0.554***	-0.428***	-0.135	-0.446***	0.568***	-0.425***	-0.100	-0.451***
	(0.132)	(0.0552)	(0.0809)	(0.111)	(0.137)	(0.0525)	(0.0791)	(0.112)
Average labor expense	-0.510***	0.137**	0.0518	0.373***	-0.559***	0.139**	0.0404	0.413***
	(0.0482)	(0.0451)	(0.0772)	(0.0718)	(0.0434)	(0.0442)	(0.0674)	(0.0632)
Average fertilizer expense	0.318**	0.514***	-0.396**	0.0361	0.114	0.426***	-0.278*	0.0977
	(0.105)	(0.109)	(0.124)	(0.163)	(0.103)	(0.103)	(0.112)	(0.136)
Average chemical expense	-0.150	-0.140	-0.308**	0.343*	0.107	-0.0819	-0.441***	0.228*
	(0.116)	(0.117)	(0.118)	(0.137)	(0.0965)	(0.107)	(0.102)	(0.109)
_cons	-3.016***	-2.903***	1.726***	-0.738	-3.152***	-3.044***	1.247***	-0.243
	(0.282)	(0.277)	(0.310)	(0.428)	(0.287)	(0.271)	(0.287)	(0.360)
Year fe	Yes	Yes	Yes	Yes	No	No	No	No
Dist fe	No	No	No	No	Yes	Yes	Yes	Yes
N	5163	5163	5163	5163	5163	5163	5163	5163

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4. Estimated Coefficients on Land Use of Irrigated Farmland and Unirrigated Farmland in Multinomial Fractional Logit Model with Lagged Variables

	Irrigated farmland	**Unirrigated farmland**	**Other1**	**Irrigated farmland**	**Unirrigated farmland**	**other1**
Average net income	-0.0672	0.0174	-0.00592	-0.0717	0.0156	-0.00418
	(0.0507)	(0.0230)	(0.0206)	(0.0492)	(0.0222)	(0.0193)
Average subsidy	0.408^{*}	0.210^{***}	-0.235^{***}	0.377^{**}	0.199^{***}	-0.223^{***}
	(0.163)	(0.0412)	(0.0390)	(0.124)	(0.0300)	(0.0293)
Average cash rent	0.0375	0.272^{***}	-0.250^{***}	0.0348	0.277^{***}	-0.254^{***}
	(0.146)	(0.0756)	(0.0631)	(0.144)	(0.0761)	(0.0631)
Average labor expense	-0.143^{*}	-0.289^{***}	0.275^{***}	-0.146^{*}	-0.290^{***}	0.276^{***}
	(0.0705)	(0.0312)	(0.0257)	(0.0695)	(0.0307)	(0.0251)
Average fertilizer expense	-0.466^{*}	-0.0274	0.0904	-0.501^{**}	-0.0643	0.126^{*}
	(0.189)	(0.0652)	(0.0601)	(0.155)	(0.0588)	(0.0531)
Average chemical expense	0.446	-0.288^{***}	0.201^{***}	0.483^{**}	-0.257^{***}	0.171^{***}
	(0.244)	(0.0666)	(0.0529)	(0.178)	(0.0593)	(0.0450)
_cons	-3.834^{***}	-1.731^{***}	1.536^{***}	-3.780^{***}	-1.852^{***}	1.641^{***}
	(0.271)	(0.161)	(0.144)	(0.324)	(0.176)	(0.174)
Year fe	Yes	Yes	Yes	No	No	No
Dist fe	No	No	No	Yes	Yes	Yes
N	3106	3106	3106	3106	3106	3106

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5. Estimated Coefficients on Land Use of Cropland, Woodland and Pastureland in Multinomial Fractional Logit Model with Lagged Variables

	Cropland	Woodland	Pasture land	other1	Cropland	Woodland	Pasture land	other1
Average net income	0.0336	-0.00681	-0.126**	-0.00592	0.00538	-0.000495	-0.103**	-0.00418
	(0.0367)	(0.0305)	(0.0451)	(0.0206)	(0.0329)	(0.0310)	(0.0365)	(0.0193)
Average subsidy	0.261***	-0.0765	0.227**	-0.235***	0.139***	-0.0819*	0.236***	-0.223***
	(0.0522)	(0.0406)	(0.0747)	(0.0390)	(0.0401)	(0.0397)	(0.0669)	(0.0293)
Average cash rent	0.606***	-0.424***	-0.125	-0.250***	0.629***	-0.417***	-0.101	-0.254***
	(0.130)	(0.0500)	(0.0776)	(0.0631)	(0.135)	(0.0457)	(0.0752)	(0.0631)
Average labor expense	-0.505***	0.103*	0.0542	0.275***	-0.519***	0.106*	0.0378	0.276***
	(0.0462)	(0.0416)	(0.0710)	(0.0257)	(0.0441)	(0.0417)	(0.0649)	(0.0251)
Average fertilizer expense	0.342**	0.515***	-0.319**	0.0904	0.126	0.452***	-0.224*	0.126*
	(0.115)	(0.103)	(0.120)	(0.0601)	(0.102)	(0.0977)	(0.105)	(0.0531)
Average chemical expense	-0.219*	-0.0826	-0.403**	0.201***	0.0117	-0.0428	-0.503***	0.171***
	(0.110)	(0.105)	(0.127)	(0.0529)	(0.0920)	(0.0929)	(0.109)	(0.0450)
_cons	-3.110***	-3.091***	1.542***	1.536***	-3.110***	-3.264***	1.019***	1.641***
	(0.252)	(0.267)	(0.318)	(0.144)	(0.273)	(0.272)	(0.287)	(0.174)
Year fe	Yes	Yes	Yes	Yes	No	No	No	No
Dist fe	No	No	No	No	Yes	Yes	Yes	Yes
N	3106	3106	3106	3106	3106	3106	3106	3106

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6. Estimated Marginal Effects on Land Use in Multinomial Fractional Logit Model

	Classification 1			Classification 2			
	Irrigated farmland	**Unirrigated farmland**	**Other1**	**Cropland**	**Woodland**	**Pasture land**	**other2**
Average net income	-0.0017	0.0019	-0.0019	0.0064	-0.0009	-0.0174***	0.0114
	(0.0559)	(0.0021)	(0.0021)	(0.0055)	(0.0017)	(0.0054)	(0.0097)
Average subsidy	0.3127***	0.0229***	-0.0275***	0.0086	-0.0034*	0.0397***	-0.0490***
	(0.0944)	(0.0029)	(0.0031)	(0.0069)	(0.0020)	(0.0082)	(0.0115)
Average cash rent	-0.0343	0.0245***	-0.0239***	0.1054***	-0.0251***	-0.0136	-0.0943***
	(0.1691)	(0.0079)	(0.0075)	(0.0247)	(0.0038)	(0.0094)	(0.0250)
Average labor expense	-0.1196	-0.0340***	0.0351***	-0.1046***	0.0080***	0.0081	0.0865***
	(0.0860)	(0.0031)	(0.0029)	(0.0072)	(0.0024)	(0.0092)	(0.0138)
Average fertilizer expense	-0.5434***	-0.0066	0.0159**	0.0187	0.0285***	-0.0400***	0.0293
	(0.1513)	(0.0065)	(0.0065)	(0.0204)	(0.0061)	(0.0135)	(0.0322)
Average chemical expense	0.5553***	-0.0234***	0.0147***	0.0207	-0.0066	-0.0490***	0.0477*
	(0.1939)	(0.0060)	(0.0052)	(0.0191)	(0.0066)	(0.0129)	(0.0254)

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7. Estimated Marginal Effects on Land Use in Multinomial Fractional Logit Model with Lagged Variables

	Classification 1			Classification 2			
	Irrigated farmland	**Unirrigated farmland**	**Other1**	**Cropland**	**Woodland**	**Pasture land**	**other1**
Average net income	-0.0011	0.0014	-0.0003	0.0009	-0.0005	-0.0136***	-0.0003
	(0.0008)	(0.0023)	(0.0023)	(0.0062)	(0.0018)	(0.0053)	(0.0023)
Average subsidy	0.0060**	0.0206***	-0.0257***	0.0265***	-0.0050**	0.0292***	-0.0257***
	(0.0026)	(0.0033)	(0.0036)	(0.0073)	(0.0022)	(0.0081)	(0.0036)
Average cash rent	0.0006	0.0284***	-0.0289***	0.1135***	-0.0251***	-0.0131	-0.0289***
	(0.0023)	(0.0078)	(0.007)	(0.0231)	(0.0033)	(0.0088)	(0.007)
Average labor expense	-0.0023*	-0.0305***	0.0321***	-0.0953***	0.0060***	0.0066	0.0321***
	(0.0014)	(0.0031)	(0.0029)	(0.007)	(0.0023)	(0.0082)	(0.0029)
Average fertilizer expense	-0.0079***	-0.005	0.0129**	0.0237	0.0291***	-0.0310***	0.0129**
	(0.0023)	(0.0062)	(0.0061)	(0.0201)	(0.0056)	(0.012)	(0.0061)
Average chemical expense	0.0077***	-0.0275***	0.0203***	0.0018	-0.0037	-0.0532***	0.0203***
	(0.0024)	(0.0061)	(0.0052)	(0.0174)	(0.0056)	(0.013)	(0.0052)

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

Acemoglu, D., & Pischke, J. S. (2001). Changes in the wage structure, family income, and children's education. *European Economic Review*, *45*(4-6), 890-904.

Abdullah, A., Doucouliagos, H., & Manning, E. (2015). Does education reduce income inequality? A meta-regression analysis. *Journal of Economic Surveys*, *29*(2), 301-316.

Ahmed, S., & McGillivray, M. (2015). Human capital, discrimination, and the gender wage gap in Bangladesh. *World Development*, *67*, 506-524.

Albrecht, J., Björklund, A., & Vroman, S. (2003). Is there a glass ceiling in Sweden?. *Journal of Labor economics*, *21*(1), 145-177.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of labor economics*, *3*, 3143-3259.

Altonji, J. G., & Dunn, T. A. (1996). The effects of family characteristics on the return to education. *The Review of Economics and Statistics*, 692-704.

Amado, C. A., Santos, S. P., & São José, J. M. (2018). Measuring and decomposing the gender pay gap: A new frontier approach. *European Journal of Operational Research*, *271*(1), 357-373.

Amaya, E., & Mougenot, B. (2019). The gender differences in highly paid wage: a case study of Peruvian physicians. *Cadernos de saude publica*, *35*, e00043018.

Arellano, M., & Bonhomme, S. (2017). Quantile selection models with an application to understanding changes in wage inequality. *Econometrica*, *85*(1), 1-28.

Balestra, S., & Backes-Gellner, U. (2013). *Heterogeneous returns to education over wage distribution: Who profits the most* (No. 0091). University of Zurich, Department of Business Administration (IBW).

Bartlett, R. L., & Callahan, C. (1984). Wage determination and marital status: Another look. *Industrial Relations: A Journal of Economy and Society*, *23*(1), 90-96.

Becker, G. S., & Becker, G. S. (2009). *A Treatise on the Family*. Harvard university press.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy*, *87*(6), 1153-1189.

Belzil, C., & Hansen, J. (2002). Unobserved ability and the return to schooling. *Econometrica*, *70*(5), 2075-2091.

Belzil, C., & Hansen, J. (2003). Structural estimates of the intergenerational education correlation. *Journal of Applied Econometrics*, *18*(6), 679-696.

Beyer, H., Rojas, P., & Vergara, R. (1999). Trade liberalization and wage inequality. *Journal of Development Economics*, *59*(1), 103-123.

Biltagy, M. (2014). Estimation of Gender Wage Differentials using Oaxaca Decomposition Technique. *Topics in Middle Eastern and North African Economies*, *16*.

Blackburn, M., & Neumark, D. (1991). *Unobserved ability, efficiency wages, and interindustry wage differentials* (No. w3857). National Bureau of Economic Research.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789-865.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Borisov, G., & Pissarides, C. (2016). The intergenerational transmission of human capital and earnings in contemporary Russia.

Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American sociological review*, 204-225.

Carrion-Flores, C., & Irwin, E. G. (2004). Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics*, *86*(4), 889-904.

Checchi, D., Ichino, A., & Rustichini, A. (1999). More equal but less mobile?: Education financing and intergenerational mobility in Italy and in the US. *Journal of public economics*, *74*(3), 351-393.

Dearden, L., Machin, S., & Reed, H. (1997). Intergenerational mobility in Britain. *The Economic Journal*, 47-66.

DiNardo, J., Fortin, N. M., & Lemieux, T. (1995). *Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach* (No. w5093). National bureau of economic research.

Dubin, J. A., & Rivers, D. (1989). Selection bias in linear regression, logit and probit models. *Sociological Methods & Research*, *18*(2-3), 360-390.

England, P., Bearak, J., Budig, M. J., & Hodges, M. J. (2016). Do highly paid, highly skilled women experience the largest motherhood penalty?. *American Sociological Review*, *81*(6), 1161-1189.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953-973.

Firpo, S., Fortin, N., & Lemieux, T. (2007). Decomposing wage distributions using recentered influence function regressions. *University of British Columbia (June)*.

Gardeazabal, J., & Ugidos, A. (2005). Gender wage discrimination at quantiles. *Journal of population economics*, *18*(1), 165-179.

Glauber, R. (2007). Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. *Journal of Marriage and Family*, *69*(4), 951-961.

Glauber, R. (2018). Trends in the motherhood wage penalty and fatherhood wage premium for low, middle, and high earners. *Demography*, *55*(5), 1663-1680.

González, X., & Miles, D. (2001). Wage inequality in a developing country: decrease in minimum wage or increase in education returns. *Empirical Economics*, *26*(1), 135-148.

Goodwin, B. K., Vandeveer, M. L., & Deal, J. L. (2004). An empirical analysis of acreage effects of participation in the federal crop insurance program. American Journal of Agricultural Economics, 86(4), 1058-1077.

Goraus, K., Tyrowicz, J., & van der Velde, L. (2017). Which gender wage gap estimates to trust? A comparative analysis. *Review of Income and Wealth*, *63*(1), 118-146.

Hanson, G. H., & Harrison, A. (1999). Trade liberalization and wage inequality in Mexico. *ILR Review*, *52*(2), 271-288.

Hardie, I. W., & Parks, P. J. (1997). Land use with heterogeneous land quality: an application of an area base model. *American Journal of Agricultural Economics*, *79*(2), 299-310.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement, Volume 5, number 4* (pp. 475-492). NBER.

Heckman, J. J. (1977). Sample selection bias as a specification error (with an application to the estimation of labor supply functions).

Heckman, J. J., & Hotz, V. J. (1986). An investigation of the labor market earnings of panamanian males evaluating the sources of inequality. *Journal of Human resources*, 507-542.

Heinze, A. (2006). Decomposing the Gender Wage Differences using Quantile Regressions. *Centre for European Economic Research, ZEW Mannheim*.

Irwin, E. G., & Geoghegan, J. (2001). Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment*, *85*(1), 7-24.

Jackman, M., & Bynoe, R. (2014). Wages, Wage Inequality, and Education: The Case of a Microstate with a "Free Education for All" Policy. *International Journal of Public Administration*, *37*(12), 866-874.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of political Economy*, *101*(3), 410-442.

Kadir, K., & Sukma, W. L. (2019). RETURNS TO EDUCATION AND WAGE DISTRIBUTION IN INDONESIA: A COMPARISON ACROSS GENDER GROUPS. *Jurnal Litbang Sukowati: Media Penelitian dan Pengembangan*, *2*(2), 19-19.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33-50.

Korenman, S., & Neumark, D. (1990). *Marriage, motherhood, and wages* (No. w3473). National Bureau of Economic Research.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., ... & George, P. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global environmental change*, *11*(4), 261-269.

Lee, J. J., Sax, L. J., Kim, K. A., & Hagedorn, L. S. (2004). Understanding students' parental education beyond first-generation status. *Community College Review*, *32*(1), 1-20.

Lemieux, T. (2006). Postsecondary education and increasing wage inequality. *American Economic Review*, *96*(2), 195-199.

Lichtenberg, E. (1989). Land quality, irrigation development, and cropping patterns in the northern high plains. *American Journal of Agricultural Economics*, *71*(1), 187-194.

Lubowski, R. N., Plantinga, A. J., & Stavins, R. N. (2008). What drives land-use change in the United States? A national analysis of landowner decisions. *Land Economics*, *84*(4), 529-550.

Machado, J. A., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, *20*(4), 445-465.

Mann, M. L., Kaufmann, R. K., Bauer, D., Gopal, S., Vera-Diaz, M. D. C., Nepstad, D., ... & Amacher, G. S. (2010). The economics of cropland conversion in Amazonia: The importance of agricultural rent. *Ecological Economics*, *69*(7), 1503-1509.

Marcus, R. D. (1984). Measuring the rate of return to interrupted schooling. *Journal of Educational Statistics*, *9*(4), 295-310.

Marjit, S., Beladi, H., & Chakrabarti, A. (2004). Trade and wage inequality in developing countries. *Economic Inquiry*, *42*(2), 295-303.

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

Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour economics*, *12*(4), 577-590.

Miao, R. (2013). Impact of ethanol plants on local land use change. *Agricultural and Resource Economics Review*, *42*(2), 291-309.

Mihăilă, R. (2016). Female labor force participation and gender wage discrimination. *Journal of Research in Gender Studies*, *6*(1), 262-268.

Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.

Mincer, J. (1997). The production of human capital and the life cycle of earnings: Variations on a theme. *Journal of labor economics*, *15*(1, Part 2), S26-S47.

Molowny-Horas, R., Basnou, C., & Pino, J. (2015). A multivariate fractional regression approach to modeling land use and cover dynamics in a Mediterranean landscape. *Computers, Environment and Urban Systems*, *54*, 47-55.

Montes-Rojas, G., Siga, L., & Mainali, R. (2017). Mean and quantile regression Oaxaca-Blinder decompositions with an application to caste discrimination. *The Journal of Economic Inequality*, *15*(3), 245-255.

Motamed, M., McPhail, L., & Williams, R. (2016). Corn area response to local ethanol markets in the United States: a grid cell level analysis. *American Journal of Agricultural Economics*, *98*(3), 726-743.

Mullahy, J. (2015). Multivariate fractional regression estimation of econometric share models. *Journal of Econometric Methods*, *4*(1), 71-100.

Mullahy, J., & Robert, S. A. (2010). No time to lose: Time constraints and physical activity in the production of health. *Review of Economics of the Household*, *8*(4), 409-432.

Mulligan, C. B., & Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, *123*(3), 1061-1110.

Neuman, S., & Oaxaca, R. L. (2004). Wage decompositions with selectivity-corrected wage equations: A methodological note. *The Journal of Economic Inequality*, *2*(1), 3-10.

Newburn, D., Reed, S., Berck, P., & Merenlender, A. (2005). Economics and Land-Use Change in Prioritizing Private Land Conservation. *Conservation Biology*, *19*(5), 1411-1420.

Nguyen, B. T., Albrecht, J. W., Vroman, S. B., & Westbrook, M. D. (2007). A quantile regression decomposition of urban–rural inequality in Vietnam. *Journal of Development Economics*, *83*(2), 466-490.

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The review of economics and statistics*, *90*(2), 290-299.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693-709.

Ogawa, H., & Fujita, M. (1980). Equilibrium land use patterns in a nonmonocentric city. *Journal of regional science*, *20*(4), 455-475.

Panagides, A., & Patrinos, H. A. (1994). *Union-nonunion wage differentials in the developing world: A case study of Mexico* (No. 1269). The World Bank.

Papke, L. E., & Wooldridge, J. (1993). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates.

Pielke, R. A. (2005). Land use and climate change. *Science*, *310*(5754), 1625-1626.

Plantinga, A. J., Lubowski, R. N., & Stavins, R. N. (2002). The effects of potential land development on agricultural land prices. *Journal of Urban Economics*, *52*(3), 561-581.

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

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

Quiggin, J. C., Karagiannis, G., & Stanton, J. (1993). Crop insurance and crop production: an empirical study of moral hazard and adverse selection. *Australian Journal of Agricultural Economics*, *37*(2), 95-113.

Schatzki, T. (2003). Options, uncertainty and sunk costs:: an empirical analysis of land use change. *Journal of environmental economics and management*, *46*(1), 86-105.

Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., ... & Yu, T. H. (2008). Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, *319*(5867), 1238-1240.

Tegen, I., Werner, M., Harrison, S. P., & Kohfeld, K. E. (2004). Relative importance of climate and land use in determining present and future global soil dust emission. *Geophysical Research Letters*, *31*(5).

Veldkamp, A., & Fresco, L. O. (1996). CLUE: a conceptual model to study the conversion of land use and its effects. *Ecological modelling*, *85*(2-3), 253-270.

Veldkamp, A., & Lambin, E. F. (2001). Predicting land-use change.

Wang, X., & Kockelman, K. M. (2009). Application of the dynamic spatial ordered probit model: Patterns of land development change in Austin, Texas. *Papers in Regional Science*, *88*(2), 345-365.

Wood, A. (1997). Openness and wage inequality in developing countries: the Latin American challenge to East Asian conventional wisdom. *The World Bank Economic Review*, *11*(1), 33-57.

Wu, J. (1999). Crop insurance, acreage decisions, and nonpoint-source pollution. *American Journal of Agricultural Economics*, *81*(2), 305-320.

Young, C. E., Vandeveer, M. L., & Schnepf, R. D. (2001). Production and price impacts of US crop

Appendix A

Rifreg following Firpo et al. (2007, 2009)

As to the contribution of individual covariates, Machado and Mata (2005) use the similar approach as DiNardo et al. (1996). The idea of DiNardo et al. (1996) is using reweighting procedure to construct counterfactual distribution of wage. However, it works only when the covariates are dummy variables. Based on DiNardo et al. (1996)'s counterfactual distribution of wage, Firpo et al. (2007, 2009) introduce the RIF (Recentered Influence Function) method. Compare with DiNardo et al. (1996), RIF is simpler and performs better than DiNardo et al. (1996). The first step in RFI, we need to construct the counterfactual distribution of wage. Then for the τth quantile of the distribution q_τ, the RIF regression is:

$$\mathrm{RIF}(\mathrm{y};\ q_\tau, F_Y) = q_\tau \frac{\tau - 1\{y \leq q_\tau\}}{f_Y(q_\tau)}$$

Where $\frac{\tau-1\{y\leq q_\tau\}}{f_Y(q_\tau)}$ is the influence function; $f_Y(q_\tau)$ is the density of Y evaluated at q_τ. As the expected value of influence function is 0, the expectation of RIF is

$$\mathrm{E}[\mathrm{RIF}(\mathrm{y};\ q_\tau, F_Y|X] = X_i^i \beta$$

Combine the RIF regression model with the counterfactual wage decomposition equation, the detailed decomposition could be written as:

$$\mathrm{v}\left(F_{Y_m}\right) - \mathrm{v}\left(F_{Y_f}\right) = (\bar{X}_m\beta_m^v - \bar{X}_m\beta_c^v) + \left(\bar{X}_m\beta_c^v - \bar{X}_f\beta_f^v\right) = \bar{X}_m(\beta_m^v - \beta_c^v) + \left(\bar{X}_m - \bar{X}_f\right)\beta_f^v + R_0$$

$$= \sum_{k=1}^{K}\left(\bar{X}_{m,k} - \bar{X}_{f,k}\right)\beta_{f,k}^v + \sum_{k=1}^{K}\bar{X}_{m,k}\left(\beta_{m,k}^v - \beta_{c,k}^v\right) + R_0$$

Where $\mathrm{v}(F_Y)$ is the distributional statistic of interest; $R_0 = \bar{X}_m(\beta_c^v - \beta_f^v)$ which represents the approximate error; K=6

In the Firpo et al. (2007, 2009) decomposition, the explained part of the traditional Blinder-Oaxaca decomposition will be called the composition effect, while the unexplained part is called wage structure effect. The first part reflects the difference in the characteristics, and the second part shows how the characteristics are “priced” in the labor market.

Table 1.6. Decomposition from RIF (Recentered Influence Function) Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	rif_y10	rif_y20	rif_y30	rif_y40	rif_y50	rif_y60	rif_y70	rif_y80	rif_y90
overall									
group_1	0.0278	0.414***	0.665***	0.845***	1.003***	1.192***	1.388***	1.614***	1.954***
	(0.0225)	(0.0168)	(0.0137)	(0.0126)	(0.0131)	(0.0136)	(0.0144)	(0.0166)	(0.0198)
group_2	-0.0835***	0.344***	0.567***	0.721***	0.891***	1.046***	1.234***	1.492***	1.782***
	(0.0204)	(0.0163)	(0.0141)	(0.0130)	(0.0134)	(0.0146)	(0.0156)	(0.0178)	(0.0207)
difference	0.111***	0.0703***	0.0984***	0.124***	0.112***	0.147***	0.154***	0.122***	0.172***
	(0.0304)	(0.0234)	(0.0197)	(0.0181)	(0.0187)	(0.0199)	(0.0213)	(0.0243)	(0.0287)
explained	-0.142***	-0.130***	-0.110***	-0.101***	-0.0955***	-0.0838***	-0.0692***	-0.0396***	-0.0122
	(0.0136)	(0.0113)	(0.00959)	(0.00880)	(0.00908)	(0.00942)	(0.00960)	(0.00991)	(0.0104)
unexplained	0.253***	0.200***	0.208***	0.225***	0.208***	0.231***	0.224***	0.161***	0.184***
	(0.0299)	(0.0224)	(0.0186)	(0.0172)	(0.0178)	(0.0189)	(0.0204)	(0.0237)	(0.0288)
explained									
years_educ_act	-0.0482***	-0.0452***	-0.0424***	-0.0430***	-0.0469***	-0.0515***	-0.0514***	-0.0556***	-0.0514***
	(0.00656)	(0.00558)	(0.00501)	(0.00498)	(0.00537)	(0.00583)	(0.00591)	(0.00652)	(0.00655)
age	0.00211	0.00282*	0.00168	0.000845	0.0000533	-0.000538	-0.00261**	-0.00529***	-0.00633***
	(0.00188)	(0.00149)	(0.00114)	(0.00101)	(0.00102)	(0.00107)	(0.00132)	(0.00189)	(0.00229)
tenure	-0.00537**	-0.00709***	-0.00731***	-0.00629***	-0.00678***	-0.00713***	-0.00591***	-0.00491***	-0.00318*
	(0.00230)	(0.00238)	(0.00232)	(0.00203)	(0.00217)	(0.00228)	(0.00201)	(0.00190)	(0.00181)
training	-0.00176	-0.00177*	-0.00218*	-0.00152*	-0.00207*	-0.00220*	-0.00291*	-0.00264*	-0.00285*
	(0.00121)	(0.00106)	(0.00117)	(0.000875)	(0.00112)	(0.00118)	(0.00152)	(0.00143)	(0.00159)
pub_emp	-0.00574	-0.0119***	-0.00797***	-0.00927***	-0.0130***	-0.0121***	-0.0119***	-0.00616*	0.00419
	(0.00427)	(0.00331)	(0.00262)	(0.00250)	(0.00275)	(0.00279)	(0.00295)	(0.00321)	(0.00387)
certificate	-0.000408	-0.000491	-0.000349	-0.000397	-0.000126	-0.000120	-0.0000243	-0.000134	-0.000325
	(0.000791)	(0.000927)	(0.000663)	(0.000748)	(0.000264)	(0.000257)	(0.000145)	(0.000298)	(0.000637)

children	0.00137	0.00166*	0.00131*	0.00170*	0.00159*	0.00135*	0.00156*	0.00168*	0.00172
	(0.00103)	(0.000987)	(0.000786)	(0.000930)	(0.000891)	(0.000802)	(0.000905)	(0.00100)	(0.00110)
Bolivia	0.0000507	-0.00000393	-0.00000562	-0.0000206	-0.0000504	-0.000108	-0.000137	-0.000286	-0.000328
	(0.000510)	(0.0000579)	(0.0000659)	(0.000208)	(0.000505)	(0.00108)	(0.00137)	(0.00286)	(0.00328)
Colombia	0.000381	0.000697	0.000604	0.000380	0.000321	0.000303	0.000378	0.000677	0.000863
	(0.00162)	(0.00295)	(0.00256)	(0.00161)	(0.00136)	(0.00129)	(0.00161)	(0.00287)	(0.00366)
Georgia	0.00669	0.000322	-0.00266	-0.000371	-0.000146	-0.00169	-0.00334	-0.00882***	-0.0117***
	(0.00442)	(0.00313)	(0.00257)	(0.00236)	(0.00244)	(0.00255)	(0.00277)	(0.00342)	(0.00423)
Ghana	-0.0494***	-0.0300***	-0.0198***	-0.0140***	-0.00809***	-0.00256	0.00102	0.00971***	0.0116***
	(0.00749)	(0.00484)	(0.00349)	(0.00287)	(0.00255)	(0.00246)	(0.00262)	(0.00331)	(0.00404)
Kenya	-0.0484***	-0.0374***	-0.0252***	-0.0197***	-0.00943***	-0.00237	0.00712*	0.0216***	0.0236***
	(0.00782)	(0.00582)	(0.00439)	(0.00387)	(0.00354)	(0.00356)	(0.00389)	(0.00500)	(0.00600)
Laos	-0.00124	-0.000300	-0.00168	-0.00417**	-0.00451**	-0.00345**	-0.00203*	0.00235*	0.00420**
	(0.00153)	(0.00107)	(0.00105)	(0.00168)	(0.00180)	(0.00149)	(0.00117)	(0.00137)	(0.00199)
Macedonia	0.000470	0.00122	0.00147	0.00128	0.00166	0.00208	0.00256	0.00277	0.00130
	(0.000791)	(0.00165)	(0.00196)	(0.00171)	(0.00221)	(0.00276)	(0.00340)	(0.00367)	(0.00177)
Sri Lanka	0.00254	0.00304	0.00152	-0.00147	-0.00117	0.00145	0.00466	0.0193***	0.0243***
	(0.00459)	(0.00333)	(0.00270)	(0.00250)	(0.00259)	(0.00270)	(0.00295)	(0.00414)	(0.00511)
Ukraine	0.00334	-0.00398	-0.00408	-0.00282	-0.00482	-0.00218	-0.00313	-0.00689	-0.000793
	(0.00607)	(0.00441)	(0.00359)	(0.00332)	(0.00346)	(0.00357)	(0.00385)	(0.00456)	(0.00551)
Vietnam	0.00186	-0.00109	-0.00279*	-0.00158	-0.00179	-0.00283*	-0.00292	-0.00709***	-0.00715**
	(0.00266)	(0.00191)	(0.00169)	(0.00148)	(0.00154)	(0.00169)	(0.00181)	(0.00262)	(0.00297)
Yunnan (China)	-0.0000960	-0.000213	-0.000240	-0.000215	-0.000194	-0.000219	-0.000158	0.0000919	0.0000879
	(0.000425)	(0.000893)	(0.00101)	(0.000898)	(0.000813)	(0.000918)	(0.000664)	(0.000398)	(0.000389)
unexplained									
years_educ_act	0.190*	-0.0196	-0.198***	-0.218***	-0.212***	-0.247***	-0.279***	-0.243***	-0.222**
	(0.105)	(0.0784)	(0.0654)	(0.0604)	(0.0625)	(0.0665)	(0.0716)	(0.0835)	(0.102)

age	-0.231**	-0.146*	-0.108	-0.0726	-0.0521	-0.0793	-0.0758	-0.0548	-0.154
	(0.115)	(0.0859)	(0.0717)	(0.0662)	(0.0685)	(0.0729)	(0.0786)	(0.0916)	(0.111)
tenure	0.0716**	0.0523**	0.0279	-0.00228	-0.00731	-0.00494	-0.00649	0.00211	0.0187
	(0.0322)	(0.0240)	(0.0200)	(0.0185)	(0.0191)	(0.0203)	(0.0219)	(0.0255)	(0.0311)
training	0.00331	-0.00336	0.000871	-0.00787	-0.00694	-0.00692	0.000821	-0.00301	-0.00826
	(0.0137)	(0.0102)	(0.00851)	(0.00787)	(0.00813)	(0.00864)	(0.00931)	(0.0109)	(0.0132)
pub_emp	-0.0141	-0.0104	-0.00720	0.000315	0.0182	-0.00503	-0.0172	-0.0520**	-0.0401
	(0.0282)	(0.0211)	(0.0175)	(0.0162)	(0.0168)	(0.0178)	(0.0192)	(0.0224)	(0.0273)
certificate	0.0170	0.0191*	0.0133	0.0137*	-0.00671	-0.0150*	-0.0241***	-0.0242**	-0.0189
	(0.0131)	(0.00981)	(0.00817)	(0.00755)	(0.00780)	(0.00830)	(0.00896)	(0.0104)	(0.0127)
children	0.0342**	0.0314***	0.0202**	0.0216**	0.0185*	0.0156	0.0142	0.00976	-0.00733
	(0.0161)	(0.0120)	(0.0100)	(0.00926)	(0.00957)	(0.0102)	(0.0110)	(0.0128)	(0.0156)
Bolivia	-0.0151	-0.00703	-0.0175**	-0.0208***	-0.0187***	-0.0269***	-0.0305***	-0.0170*	-0.00415
	(0.0119)	(0.00886)	(0.00739)	(0.00685)	(0.00707)	(0.00755)	(0.00814)	(0.00940)	(0.0114)
Colombia	0.0130	0.0135	-0.0102	-0.0294***	-0.0306***	-0.0312***	-0.0353***	-0.0209**	-0.00825
	(0.0125)	(0.00930)	(0.00772)	(0.00724)	(0.00749)	(0.00794)	(0.00856)	(0.00988)	(0.0120)
Georgia	-0.00199	0.0120	0.00351	-0.00917	-0.0103	-0.0137*	-0.0119	0.00784	0.0258**
	(0.0120)	(0.00892)	(0.00737)	(0.00683)	(0.00706)	(0.00748)	(0.00804)	(0.00939)	(0.0115)
Ghana	0.0155**	0.00735	-0.00204	-0.0121***	-0.0139***	-0.0206***	-0.0246***	-0.0195***	-0.0125*
	(0.00756)	(0.00562)	(0.00468)	(0.00438)	(0.00454)	(0.00490)	(0.00532)	(0.00608)	(0.00730)
Kenya	-0.00513	-0.00374	-0.0230***	-0.0354***	-0.0349***	-0.0437***	-0.0477***	-0.0344***	-0.0275**
	(0.0121)	(0.00899)	(0.00755)	(0.00708)	(0.00730)	(0.00783)	(0.00843)	(0.00966)	(0.0117)
Laos	-0.00767	-0.0106*	-0.0230***	-0.0202***	-0.0245***	-0.0219***	-0.0217***	-0.00814	-0.00516
	(0.00810)	(0.00605)	(0.00518)	(0.00477)	(0.00499)	(0.00524)	(0.00562)	(0.00641)	(0.00779)
Macedonia	-0.00232	-0.00548	-0.0262**	-0.0439***	-0.0429***	-0.0445***	-0.0447***	-0.0452***	-0.0497***
	(0.0166)	(0.0123)	(0.0102)	(0.00956)	(0.00987)	(0.0105)	(0.0112)	(0.0131)	(0.0159)

Sri Lanka	0.0122	0.0200***	0.00174	-0.0137***	-0.0214***	-0.0289***	-0.0298***	-0.00882	0.00457
	(0.00892)	(0.00671)	(0.00552)	(0.00515)	(0.00539)	(0.00581)	(0.00624)	(0.00706)	(0.00858)
Ukraine	-0.0195	-0.0141	-0.0370***	-0.0252**	-0.00892	-0.0167	0.000969	0.0165	0.0250
	(0.0178)	(0.0131)	(0.0109)	(0.0101)	(0.0104)	(0.0110)	(0.0118)	(0.0139)	(0.0169)
Vietnam	-0.0182	-0.0164	-0.0284***	-0.0366***	-0.0450***	-0.0484***	-0.0570***	-0.0260*	-0.0205
	(0.0174)	(0.0129)	(0.0107)	(0.00993)	(0.0103)	(0.0109)	(0.0118)	(0.0136)	(0.0166)
Yunnan (China)	-0.0164	-0.0402***	-0.0529***	-0.0437***	-0.0387***	-0.0453***	-0.0398***	-0.0177*	0.00149
	(0.0131)	(0.00985)	(0.00834)	(0.00766)	(0.00786)	(0.00838)	(0.00894)	(0.0103)	(0.0125)
_cons	0.227	0.321**	0.674***	0.780***	0.746***	0.915***	0.953***	0.700***	0.687***
	(0.198)	(0.148)	(0.123)	(0.114)	(0.118)	(0.125)	(0.135)	(0.157)	(0.192)
N	12405	12405	12405	12405	12405	12405	12405	12405	12405

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.6 shows the results using decomposition with RIF regression. Since this approach is different from MM(2005), the results are slightly different. Here we only focus on the detailed decomposition results. In the explained part, difference in educational years and experience contributes to the wage gap. The difference in number of children also causes the different wage. In the unexplained part, the discrimination is mainly from number of children, especially at low and middle wage levels. This kind of discrimination decreases as the increase in wage level. Discrimination in education also exists at middle wage groups.

Appendix B

Results from OLS and Quantile Regression by Regions

Table 5.1. Central America

Quantiles	OLS		10		20		30		40	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0673***	0.0723***	0.0639***	0.0672***	0.0578***	0.0714***	0.0629***	0.0621***	0.0603***	0.0682***
	(0.00623)	(0.00644)	(0.0152)	(0.0133)	(0.00817)	(0.00931)	(0.0100)	(0.00826)	(0.00705)	(0.00899)
Edu2	0.0207***	0.0184**	0.00999	-0.0000253	0.00991	0.0131	0.0162**	0.0193***	0.0181***	0.0200***
	(0.00683)	(0.00738)	(0.0110)	(0.0138)	(0.00912)	(0.00924)	(0.00683)	(0.00570)	(0.00604)	(0.00598)
age	0.0491***	0.00726	0.0915***	0.0247	0.0507***	0.0148	0.0461***	0.0204	0.0407***	0.0281**
	(0.0115)	(0.0129)	(0.0237)	(0.0202)	(0.0154)	(0.0145)	(0.0110)	(0.0130)	(0.0134)	(0.0114)
Age2	-0.000551***	-0.0000309	-0.00112***	-0.000348	-0.000604***	-0.000160	-0.000542***	-0.000199	-0.000427**	-0.000295**
	(0.000150)	(0.000167)	(0.000318)	(0.000288)	(0.000224)	(0.000202)	(0.000153)	(0.000169)	(0.000183)	(0.000147)
tenure	-0.000525	0.00253	-0.00561	-0.00988	-0.00327	-0.00601	-0.00289	0.00382	-0.00250	0.00658
	(0.00372)	(0.00427)	(0.00893)	(0.0124)	(0.00430)	(0.00781)	(0.00289)	(0.00504)	(0.00486)	(0.00477)
training	0.146**	0.132**	0.199**	0.205*	0.229***	0.227***	0.154***	0.264***	0.168***	0.213***
	(0.0569)	(0.0665)	(0.0882)	(0.109)	(0.0508)	(0.0676)	(0.0399)	(0.0666)	(0.0454)	(0.0701)
certificate	-0.0118	0.0463	-0.137	0.0387	-0.00447	-0.0224	-0.0512	0.0850	-0.00290	0.0320
	(0.0752)	(0.0829)	(0.168)	(0.126)	(0.102)	(0.0927)	(0.102)	(0.0888)	(0.0953)	(0.0925)
children	0.0347	0.0168	-0.00983	-0.120	0.0148	-0.102**	0.0303	-0.0167	0.0360	-0.0214
	(0.0327)	(0.0355)	(0.0507)	(0.0757)	(0.0609)	(0.0435)	(0.0400)	(0.0479)	(0.0379)	(0.0455)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)

_cons	-0.645***	-0.173	-2.151***	-1.149***	-1.030***	-0.836***	-0.856***	-0.743***	-0.640***	-0.753***
	(0.220)	(0.250)	(0.473)	(0.370)	(0.232)	(0.306)	(0.200)	(0.279)	(0.231)	(0.235)
Quantiles	50		60		70		80		90	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0629***	0.0672***	0.0739***	0.0757***	0.0782***	0.0810***	0.0751***	0.0736***	0.0688***	0.0586***
	(0.00720)	(0.00921)	(0.00750)	(0.00728)	(0.00846)	(0.00723)	(0.00983)	(0.00660)	(0.0138)	(0.0104)
Edu2	0.0198***	0.0199***	0.0299***	0.0192***	0.0419***	0.0213***	0.0310***	0.0134	0.0283**	0.0304*
	(0.00701)	(0.00581)	(0.00715)	(0.00699)	(0.00740)	(0.00797)	(0.00976)	(0.00938)	(0.0114)	(0.0173)
age	0.0461***	0.0129	0.0417***	0.0102	0.0422***	0.00247	0.0298*	-0.0110	0.0227	-0.0250
	(0.0135)	(0.00951)	(0.0128)	(0.0120)	(0.0163)	(0.0133)	(0.0180)	(0.0194)	(0.0248)	(0.0241)
Age2	-0.000483***	-0.000109	-0.000412**	-0.0000192	-0.000410*	0.0000674	-0.000216	0.000257	-0.000220	0.000402
	(0.000176)	(0.000149)	(0.000167)	(0.000169)	(0.000213)	(0.000182)	(0.000238)	(0.000285)	(0.000327)	(0.000294)
tenure	0.000495	0.00878	0.00532	0.00503	0.00425	0.00833	0.000914	0.00849*	0.00150	0.00352
	(0.00563)	(0.00540)	(0.00604)	(0.00438)	(0.00646)	(0.00555)	(0.00745)	(0.00497)	(0.00936)	(0.0115)
training	0.148***	0.175***	0.0939	0.117**	0.0712	0.0775	0.172*	-0.0552	0.0962	0.0247
	(0.0481)	(0.0614)	(0.0606)	(0.0597)	(0.0886)	(0.0597)	(0.0981)	(0.0954)	(0.141)	(0.0971)
certificate	0.00601	0.118	-0.0203	0.125	0.0164	0.0323	-0.0831	0.00325	0.179	-0.107
	(0.0951)	(0.0803)	(0.100)	(0.0794)	(0.114)	(0.0479)	(0.109)	(0.0913)	(0.162)	(0.119)
children	0.0198	0.0253	0.0454	0.0722**	0.0487	0.0819***	0.0417	0.0520	0.0866	0.0726
	(0.0310)	(0.0364)	(0.0360)	(0.0290)	(0.0413)	(0.0258)	(0.0476)	(0.0493)	(0.0952)	(0.0807)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-0.645***	-0.300*	-0.613***	-0.217	-0.532*	0.0656	0.00594	0.772**	0.787**	1.580***
	(0.202)	(0.164)	(0.217)	(0.230)	(0.286)	(0.282)	(0.296)	(0.339)	(0.350)	(0.601)
N									1566	1688

Standard errors in parentheses

[*] $p < 0.1$, [**] $p < 0.05$, [***] $p < 0.01$

Table 5.2. Africa Sub-Saharan

Quantiles	OLS		10		20		30		40	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0807***	0.0806***	0.0789***	0.0824***	0.0832***	0.0823***	0.0853***	0.0845***	0.0924***	0.0817***
	(0.00561)	(0.00616)	(0.0143)	(0.00947)	(0.00886)	(0.00899)	(0.00865)	(0.0122)	(0.00885)	(0.0101)
Edu2	0.00691	0.0102	-0.00206	-0.00280	-0.0117	0.00531	0.00521	-0.00210	0.0152	-0.00239
	(0.00883)	(0.0113)	(0.0207)	(0.0147)	(0.0129)	(0.0196)	(0.00964)	(0.0135)	(0.00977)	(0.0167)
age	0.0442***	0.0546***	0.0780**	0.0242	0.0683***	0.0588***	0.0629***	0.0785***	0.0591***	0.0587***
	(0.0140)	(0.0157)	(0.0305)	(0.0220)	(0.0162)	(0.0175)	(0.0211)	(0.0191)	(0.0193)	(0.0140)
Age2	-0.000549***	-0.000642***	-0.00102**	-0.000389	-0.000857***	-0.000797***	-0.000787***	-0.00106***	-0.000749***	-0.000714***
	(0.000178)	(0.000204)	(0.000422)	(0.000283)	(0.000205)	(0.000243)	(0.000265)	(0.000254)	(0.000251)	(0.000190)
tenure	0.0149***	0.0122***	-0.00296	0.0176**	0.00766*	0.0196**	0.0112**	0.0217**	0.0153***	0.0175***
	(0.00386)	(0.00447)	(0.00873)	(0.00823)	(0.00455)	(0.00780)	(0.00439)	(0.00860)	(0.00342)	(0.00594)
training	0.367***	0.517***	0.362***	0.171	0.362***	0.445***	0.441***	0.448***	0.422***	0.680***
	(0.0698)	(0.0983)	(0.0722)	(0.171)	(0.0784)	(0.154)	(0.0756)	(0.134)	(0.0663)	(0.181)
certificate	0.180**	0.0596	0.118	0.291	0.0774	0.263**	0.110	0.0935	0.167*	0.0909
	(0.0750)	(0.101)	(0.135)	(0.235)	(0.0721)	(0.112)	(0.0743)	(0.124)	(0.0931)	(0.116)
children	-0.0223	-0.00748	-0.0324	-0.121***	0.0324	-0.0719	0.0189	-0.0705	0.00181	-0.0236
	(0.0330)	(0.0315)	(0.0692)	(0.0432)	(0.0397)	(0.0538)	(0.0415)	(0.0595)	(0.0345)	(0.0419)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-1.200***	-1.685***	-2.794***	-2.220***	-2.260***	-2.549***	-2.050***	-2.479***	-1.893***	-1.936***
	(0.271)	(0.316)	(0.559)	(0.406)	(0.318)	(0.373)	(0.387)	(0.451)	(0.287)	(0.312)
Quantiles	50		60		70		80		90	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female

edu	0.0881***	0.0871***	0.0930***	0.0906***	0.0805***	0.0872***	0.0820***	0.0804***	0.0725***	0.0766***
	(0.00950)	(0.0104)	(0.0103)	(0.00937)	(0.00857)	(0.00673)	(0.00630)	(0.00807)	(0.00833)	(0.0108)
Edu2	0.0102	0.0119	0.0219***	0.0272**	0.0202**	0.0221	0.0196	0.0115	0.0256	0.0292*
	(0.0106)	(0.0113)	(0.00677)	(0.0107)	(0.00977)	(0.0150)	(0.0126)	(0.0196)	(0.0167)	(0.0172)
age	0.0567***	0.0700***	0.0539***	0.0680***	0.0270	0.0596***	0.00386	0.0533**	-0.0216	0.0577*
	(0.0207)	(0.0172)	(0.0194)	(0.0182)	(0.0245)	(0.0200)	(0.0199)	(0.0258)	(0.0336)	(0.0326)
Age2	-0.000695***	-0.000810***	-0.000661***	-0.000753***	-0.000335	-0.000634**	-0.0000312	-0.000488	0.000383	-0.000546
	(0.000250)	(0.000215)	(0.000244)	(0.000269)	(0.000310)	(0.000294)	(0.000262)	(0.000359)	(0.000451)	(0.000430)
tenure	0.0189***	0.0162***	0.0173***	0.0114*	0.0168***	0.0140**	0.0199***	0.00807	0.0274***	0.00372
	(0.00347)	(0.00588)	(0.00487)	(0.00586)	(0.00389)	(0.00646)	(0.00543)	(0.00633)	(0.00691)	(0.0106)
training	0.415***	0.645***	0.355***	0.564***	0.327***	0.530***	0.317***	0.639***	0.328***	0.736***
	(0.0702)	(0.110)	(0.0789)	(0.0751)	(0.0830)	(0.0965)	(0.0707)	(0.122)	(0.109)	(0.178)
certificate	0.105	0.0295	0.0984	-0.0102	0.122	0.0475	0.175	0.0229	0.364	-0.0577
	(0.0649)	(0.108)	(0.0730)	(0.0914)	(0.112)	(0.102)	(0.108)	(0.143)	(0.244)	(0.136)
children	-0.00671	0.0175	0.00729	0.00513	0.00609	0.0445	0.00786	0.120**	-0.0983**	0.144**
	(0.0357)	(0.0451)	(0.0401)	(0.0435)	(0.0442)	(0.0535)	(0.0309)	(0.0498)	(0.0425)	(0.0677)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-1.611***	-2.119***	-1.444***	-1.957***	-0.551	-1.538***	0.120	-1.030**	0.960	-0.698
	(0.330)	(0.421)	(0.304)	(0.339)	(0.432)	(0.399)	(0.389)	(0.516)	(0.587)	(0.616)
N									2155	2037

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.3. **Eastern & Southern Europe**

Quantiles	OLS		10		20		30		40	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0532***	0.0624***	0.0425***	0.0482***	0.0401***	0.0524***	0.0417***	0.0593***	0.0524***	0.0693***
	(0.00585)	(0.00515)	(0.00887)	(0.00632)	(0.0104)	(0.00800)	(0.00682)	(0.00674)	(0.00465)	(0.00780)
Edu2	0.00280	0.00553	-0.00956	-0.00941*	-0.00825*	-0.00165	-0.00312	-0.00207	-0.000714	0.00415
	(0.00444)	(0.00390)	(0.00792)	(0.00566)	(0.00499)	(0.00516)	(0.00487)	(0.00264)	(0.00318)	(0.00377)
age	0.0344***	0.0309***	0.0320**	0.0293**	0.0391***	0.0289**	0.0381***	0.0308***	0.0333***	0.0358***
	(0.0102)	(0.00874)	(0.0146)	(0.0114)	(0.0120)	(0.0129)	(0.0103)	(0.00914)	(0.00902)	(0.00628)
Age2	-0.000439***	-0.000411***	-0.000506***	-0.000416***	-0.000578***	-0.000410***	-0.000524***	-0.000430***	-0.000438***	-0.000483***
	(0.000123)	(0.000103)	(0.000178)	(0.000129)	(0.000151)	(0.000152)	(0.000131)	(0.000109)	(0.000118)	(0.0000850)
tenure	0.0141***	0.0107***	0.0147***	0.0100***	0.0157***	0.0112***	0.0180***	0.0101***	0.0171***	0.0112***
	(0.00203)	(0.00158)	(0.00292)	(0.00155)	(0.00196)	(0.00177)	(0.00202)	(0.00152)	(0.00228)	(0.00181)
training	0.126**	0.121***	0.327***	0.238***	0.245***	0.158***	0.230***	0.147***	0.229***	0.107***
	(0.0596)	(0.0401)	(0.0929)	(0.0620)	(0.0596)	(0.0562)	(0.0687)	(0.0376)	(0.0552)	(0.0293)
certificate	0.179***	0.189***	0.223***	0.0263	0.125***	0.117*	0.105*	0.146**	0.113*	0.187***
	(0.0516)	(0.0392)	(0.0644)	(0.0763)	(0.0468)	(0.0630)	(0.0638)	(0.0672)	(0.0642)	(0.0458)
children	0.0722***	0.00769	0.109***	-0.0521	0.0676***	0.00390	0.0692***	-0.000575	0.0561**	-0.00894
	(0.0263)	(0.0263)	(0.0391)	(0.0491)	(0.0207)	(0.0382)	(0.0249)	(0.0271)	(0.0260)	(0.0308)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-0.143	-0.495**	-0.532*	-0.876***	-0.372	-0.727**	-0.302	-0.680***	-0.226	-0.800***
	(0.212)	(0.194)	(0.282)	(0.277)	(0.231)	(0.300)	(0.197)	(0.238)	(0.153)	(0.143)
Quantiles	50		60		70		80		90	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female

edu	0.0618***	0.0688***	0.0626***	0.0687***	0.0640***	0.0719***	0.0682***	0.0699***	0.0618***	0.0646***
	(0.00593)	(0.00839)	(0.00457)	(0.00757)	(0.00795)	(0.00740)	(0.00534)	(0.00668)	(0.0119)	(0.00921)
Edu2	-0.00142	0.00829**	0.00416	0.0137***	0.00126	0.0117***	0.00480	0.0156***	0.00784	0.0237***
	(0.00482)	(0.00356)	(0.00344)	(0.00431)	(0.00316)	(0.00412)	(0.00430)	(0.00399)	(0.00828)	(0.00487)
age	0.0400***	0.0402***	0.0375***	0.0419***	0.0307***	0.0387***	0.0120	0.0102	0.0184*	0.00448
	(0.00587)	(0.00955)	(0.00690)	(0.0106)	(0.0110)	(0.0137)	(0.0109)	(0.0106)	(0.0108)	(0.0125)
Age2	-0.000482***	-0.000515***	-0.000440***	-0.000535***	-0.000359***	-0.000487***	-0.000130	-0.000134	-0.000191	-0.0000483
	(0.0000723)	(0.000121)	(0.0000802)	(0.000135)	(0.000131)	(0.000168)	(0.000131)	(0.000122)	(0.000134)	(0.000151)
tenure	0.0136***	0.0110***	0.0132***	0.0125***	0.0138***	0.0126***	0.0122***	0.0110***	0.0108***	0.00833**
	(0.00182)	(0.00190)	(0.00169)	(0.00178)	(0.00223)	(0.00175)	(0.00247)	(0.00213)	(0.00331)	(0.00358)
training	0.154***	0.0736	0.0995*	0.0417	0.0826**	0.0560	0.0237	0.0824*	-0.0100	0.117
	(0.0449)	(0.0482)	(0.0548)	(0.0452)	(0.0395)	(0.0468)	(0.0321)	(0.0496)	(0.0682)	(0.0821)
certificate	0.124**	0.198***	0.159***	0.251***	0.135***	0.222***	0.176***	0.211***	0.275***	0.269***
	(0.0585)	(0.0517)	(0.0516)	(0.0477)	(0.0472)	(0.0355)	(0.0353)	(0.0529)	(0.0887)	(0.0972)
children	0.0465*	0.0128	0.0357	0.0260	0.0528**	0.0286	0.0623**	0.0397	0.0318	0.0232
	(0.0259)	(0.0309)	(0.0263)	(0.0273)	(0.0245)	(0.0294)	(0.0244)	(0.0271)	(0.0608)	(0.0539)
Urban	0	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-0.339**	-0.800***	-0.216	-0.725***	0.0434	-0.563*	0.513**	0.180	0.668***	0.567*
	(0.145)	(0.248)	(0.165)	(0.259)	(0.253)	(0.312)	(0.233)	(0.249)	(0.249)	(0.333)
N					0.0640***	0.0719***			1887	2375

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.4. **Southern & Eastern Asia**

Quantiles	OLS		10		20		30		40	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0640***	0.0843***	0.0976***	0.145***	0.0808***	0.108***	0.0665***	0.0953***	0.0567***	0.0871***
	(0.00525)	(0.00484)	(0.0116)	(0.00901)	(0.00970)	(0.00640)	(0.00701)	(0.00577)	(0.00476)	(0.00550)
Edu2	0.0158**	0.0208***	0.0168*	0.0132	0.0175***	0.0167***	0.0206***	0.0205***	0.0136*	0.0228***
	(0.00695)	(0.00703)	(0.00940)	(0.0100)	(0.00572)	(0.00398)	(0.00513)	(0.00642)	(0.00803)	(0.00439)
age	0.0364***	0.0513***	0.0548***	0.0281	0.0438***	0.0378**	0.0306*	0.0438***	0.0262*	0.0428***
	(0.0125)	(0.0126)	(0.0173)	(0.0235)	(0.0161)	(0.0159)	(0.0173)	(0.0136)	(0.0150)	(0.0111)
Age2	-0.000468***	-0.000628***	-0.000787***	-0.000452	-0.000584***	-0.000529**	-0.000437**	-0.000582***	-0.000360*	-0.000571***
	(0.000153)	(0.000157)	(0.000237)	(0.000294)	(0.000187)	(0.000223)	(0.000216)	(0.000184)	(0.000190)	(0.000152)
tenure	0.00296	-0.000682	0.00917***	0.00821**	0.00297	0.00585***	0.00906***	0.00428**	0.00853***	0.00617***
	(0.00241)	(0.00237)	(0.00324)	(0.00359)	(0.00276)	(0.00202)	(0.00251)	(0.00206)	(0.00304)	(0.00183)
training	-0.0833	0.0800	-0.0352	0.174*	-0.112*	0.115**	-0.105	0.0521	-0.0375	0.0358
	(0.0716)	(0.0709)	(0.0705)	(0.0953)	(0.0665)	(0.0478)	(0.0733)	(0.0550)	(0.0708)	(0.0370)
certificate	0.138***	0.181***	0.117*	0.162**	0.163***	0.162***	0.211***	0.123***	0.176***	0.0781
	(0.0522)	(0.0547)	(0.0625)	(0.0704)	(0.0557)	(0.0421)	(0.0403)	(0.0331)	(0.0427)	(0.0476)
children	-0.00189	0.0181	-0.0534	-0.0724	0.00231	-0.0757**	-0.0133	-0.0651	0.00862	0.00168
	(0.0327)	(0.0322)	(0.0700)	(0.0739)	(0.0405)	(0.0325)	(0.0357)	(0.0406)	(0.0293)	(0.0249)
Urban	0.364***	0.199***	1.138***	0.577***	0.712***	0.663***	0.674***	0.426**	0.439***	0.428***
	(0.0773)	(0.0764)	(0.196)	(0.0912)	(0.135)	(0.113)	(0.174)	(0.208)	(0.148)	(0.153)
_cons	-0.904***	-1.516***	-3.257***	-2.963***	-2.143***	-2.489***	-1.526***	-2.020***	-0.908***	-1.757***
	(0.259)	(0.258)	(0.370)	(0.470)	(0.396)	(0.330)	(0.392)	(0.343)	(0.343)	(0.273)
Quantiles	50		60		70		80		90	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
edu	0.0542***	0.0771***	0.0497***	0.0763***	0.0474***	0.0727***	0.0477***	0.0672***	0.0385***	0.0525***

	(0.00458)	(0.00552)	(0.00534)	(0.00532)	(0.00733)	(0.00505)	(0.00762)	(0.00561)	(0.0135)	(0.00998)
Edu2	0.0137**	0.0216***	0.0144**	0.0245***	0.0163**	0.0213***	0.0225*	0.0233***	0.0321**	0.0356***
	(0.00643)	(0.00593)	(0.00595)	(0.00631)	(0.00727)	(0.00697)	(0.0121)	(0.00778)	(0.0151)	(0.0109)
age	0.0312***	0.0411***	0.0336***	0.0418***	0.0305**	0.0504***	0.0327***	0.0549***	0.0455*	0.0676***
	(0.0118)	(0.0119)	(0.0117)	(0.0114)	(0.0125)	(0.0134)	(0.0126)	(0.0190)	(0.0240)	(0.0253)
Age2	-0.000421***	-0.000554***	-0.000433***	-0.000508***	-0.000348**	-0.000604***	-0.000297*	-0.000639**	-0.000454	-0.000739**
	(0.000151)	(0.000163)	(0.000157)	(0.000162)	(0.000160)	(0.000177)	(0.000159)	(0.000250)	(0.000291)	(0.000348)
tenure	0.00881***	0.00594**	0.00781***	0.00312	0.00510**	0.00311	0.000788	0.00109	-0.00149	-0.00592
	(0.00177)	(0.00231)	(0.00185)	(0.00211)	(0.00224)	(0.00203)	(0.00320)	(0.00313)	(0.00433)	(0.00367)
training	-0.0637	0.0874**	-0.0753	0.0727*	-0.0375	0.0413	-0.0482	0.00851	-0.101	-0.0838
	(0.0561)	(0.0388)	(0.0632)	(0.0435)	(0.0800)	(0.0389)	(0.0969)	(0.0626)	(0.163)	(0.119)
certificate	0.123***	0.139***	0.0927**	0.160***	0.0919**	0.169***	0.113	0.173***	0.0546	0.183*
	(0.0330)	(0.0510)	(0.0360)	(0.0574)	(0.0462)	(0.0452)	(0.0840)	(0.0584)	(0.100)	(0.0970)
children	0.00898	0.0231	0.0488*	0.0195	0.0323	0.0595***	0.0619	0.0628	0.0726	0.173**
	(0.0302)	(0.0216)	(0.0274)	(0.0163)	(0.0327)	(0.0198)	(0.0508)	(0.0484)	(0.0973)	(0.0696)
Urban	0.352***	0.234	0.366***	0.0244	0.123	-0.124	-0.00915	-0.135	-0.164	-0.344
	(0.0976)	(0.194)	(0.125)	(0.185)	(0.152)	(0.184)	(0.271)	(0.181)	(0.222)	(0.210)
_cons	-0.716***	-1.266***	-0.595**	-0.980***	-0.149	-0.785**	0.0592	-0.586	0.489	-0.163
	(0.256)	(0.265)	(0.244)	(0.265)	(0.243)	(0.315)	(0.336)	(0.449)	(0.512)	(0.489)
N									2257	2776

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$