Master Degree in Economics and Finance

Gender Differentials in Returns to Education in Developing Countries

Authors:
Ignatius Barnardt, Golschan Khun Jush,
Thies Wollesen, Samuel Hayden,
Eva Sotosek

Directors:
Manuel García-Santana
and
Libertad González

June 2017

ABSTRACT IN ENGLISH:

We investigate a possible gender gap in returns to education using data from the World Bank' STEP program for seven developing and emerging countries. We control for cognitive skills, non-cognitive skills and parental education - previously unobserved due to unavailability of data - to investigate how this heterogeneity is playing a role in estimating the gender differential in educational returns. We also model selection using the Heckman two-step estimation procedure to examine whether selection may be driving this phenomenon. Our findings suggest that gender gaps in returns to education are not as prominent in the countries in our sample as previously suggested. We also find that controlling for unobserved heterogeneity on the one hand, and selection on the other, has different effects in different countries, highlighting the importance of understanding individual countries' labour markets in detail before drawing conclusions regarding the existence of a gender gap in returns to education.

ABSTRACT IN CATALAN:

Investiguem una possible bretxa de gènere en els retorns a l'educació mitjançant dades del programa STEP del Banc Mundial per a set països en desenvolupament i emergents. Controlem les habilitats cognitives, les habilitats no cognitives i l'educació dels pares, que abans no es van observar a causa de la inexistència de dades, per investigar com aquesta heterogeneïtat juga un paper en l'estimació de la diferència de gènere en els rendiments educatius. També modelem la selecció amb el procediment d'estimació de dos passos de Heckman per examinar si la selecció pot conduir aquest fenomen. Les nostres dades suggereixen que les llacunes de gènere en els retorns a l'educació no són tan importants en els països de la nostra mostra com es va suggerir anteriorment. També veiem que controlar, per una banda, la heterogeneïtat no observada, i la selecció, d'altra banda, té diferents efectes en diferents països, destacant la importància d'entendre detalladament els mercats laborals dels països individuals abans de treure conclusions sobre l'existència d'una bretxa de gènere en els rendiments a l'educació.

Master Project

Gender Differentials in Returns to Education in Developing Countries

Master in Economics and Finance - 2016/2017

Authors

Ignatius Barnardt 138322

Golschan Khun Jush 138339

Thies Wollesen 138353

Samuel Hayden 138350

Eva Sotošek 138357

Supervisors

Prof. Libertad González

Prof. Manuel García-Santana

Submission date

1 June 2017

Barcelona GSE Graduate School of Economics

Contents

1 Introduction **1**

2 Literature review **2**

3 Description of the data **5**

4 Estimation strategy **6**

4.1 Returns gap across genders in a Mincerian framework 6

4.2 Controlling for previously unobserved heterogeneity 7

4.3 Modelling and controlling for selection 8

4.3.1 Discussion of excludable variables for Heckman procedure 9

5 Results and interpretation **10**

5.1 Existence and size of estimated returns gaps 10

5.1.1 Estimation results 10

5.1.2 Level estimates of the returns to education for females and males 10

5.1.3 Estimates of the gender gap in returns to education 11

5.2 Effect of adding STEP controls 13

5.3 Effect of controlling for selection 15

5.4 Limitations and potential extensions 19

6 Conclusion **20**

Abstract

We investigate a possible gender gap in returns to education using data from the World Bank' STEP program for seven developing and emerging countries. We control for cognitive skills, non-cognitive skills and parental education - previously unobserved due to unavailability of data - to investigate how this heterogeneity is playing a role in estimating the gender differential in educational returns. We also model selection using the Heckman two-step estimation procedure to examine whether selection may be driving this phenomenon. Our findings suggest that gender gaps in returns to education are not as prominent in the countries in our sample as previously suggested. We also find that controlling for unobserved heterogeneity on the one hand, and selection on the other, has different effects in different countries, highlighting the importance of understanding individual countries' labour markets in detail before drawing conclusions regarding the existence of a gender gap in returns to education.

1 Introduction

In recent decades, an increasing number of countries have been experiencing changing gender differentials in returns to education. Contrary to the trend prevailing in most of the 20th century and before, where male returns to education were higher compared to those of females, females seem to have been catching up in many developed and developing countries. Numerous studies find that the returns to education for females in developing countries now exceed those of males (Montenegro and Patrinos, 2014). This has led many to believe that wage gaps between women and men are slowly diminishing and that returns to education are now in fact higher for women than for men in some countries.

However, researchers such as Heckman and Li (2004) and Card (2001) have made the case that simple Mincerian OLS estimates without additional controls may yield misleading results. First, failing to include relevant yet unobserved individual characteristics, such as specific cognitive and non-cognitive skills or a more supportive family background, directly leads to omitted variable bias if the omitted variables are correlated with both wages and education. Second, for many countries there is a lot of evidence that women who work do not represent a random sample of the population. If heterogeneous unobserved characteristics both drive their entry into employment and are further correlated with returns to education, estimating the returns without accounting for selection leads to inconsistent estimates.

The first aim of this paper is to investigate the existence and size of the returns gap for a set of seven developing and emerging countries. Based on previous literature, we expect to find that returns to education are higher for females than for males, due to relevant unobserved factors being more highly correlated with labour market and educational outcomes for females compared to males.

Next, we explore whether or not the returns gap, if present, is driven by heterogeneity in the form of individual-specific cognitive abilities, non-cognitive abilities and family background. Together, we refer to these three groups of variables as the "STEP controls." Comparing gender-specific returns to education with and without these controls could provide evidence on whether previously unobserved heterogeneity is partially driving any observed returns differential. Since our hypothesis is that unobserved characteristics influence females' labour market and educational outcomes to a greater extent than those of males, we expect to observe a narrowed returns gap after adding the STEP controls.

Finally, we examine whether selection into employment may be fuelling the observed returns gaps (or lack of observed returns gaps) using the Heckman two-step procedure to model selection. We expect selection to be more relevant for females than for males because female labour force participation is generally relatively low, making the importance of selection more likely. As we explain below, controlling for selection could result in a higher estimate of the returns gap if females are positively selected and education is positively related to the probability of working.

We use World Bank data from the Skills Towards Employability and Productivity (STEP) program conducted in 2012 and 2013 for seven different countries: Bolivia, Colombia, Georgia, Ghana, Kenya, Ukraine, and Vietnam. Compared to data used in previous studies, this household survey includes more detailed qualitative and quantitative information regarding individual abilities such as non-cognitive and cognitive skills, and characteristics such as parental education.

In view of the above, we make the following contributions to the existing literature: 1) we use a unique data source with variables unavailable to previous studies to control for individual heterogeneity; 2) we undertake an analysis over multiple countries thanks to the survey instruments used in each country being highly comparable; 3) we investigate whether gender gaps in returns to education found in developed and other developing countries are also present in our set of developing and emerging economies; 4) we use quantile regressions to investigate the existence and size of any returns gap at different parts of the wage distribution; and 5) we use a novel combination of excludable variables for the first stage of the Heckman procedure to model selection in estimating the returns gap, again thanks to the STEP data.

The rest of the paper is structured as follows. Section 2 introduces the existing literature on the role of different potential control variables and selection bias in estimating returns to education, and Section 3 elaborates on the underlying data used. Section 4 discusses in more detail the methodology and econometric approaches adopted for our research. Section 5 evaluates and interprets the results, before highlighting the limitations of our approach and potential directions for future research. Finally, Section 6 concludes.

2 Literature review

The predominant share of empirical work in labour economics is motivated by human capital theory (Becker, 1964), which suggests that an individual's set of knowledge, skills, abilities, and other characteristics is a key determinant of individual as well as aggregate economic success. Therefore, most of the studies examining returns to education rely on the Mincerian framework (Mincer, 1974), capturing the idea that schooling develops general skills and is thus a good proxy for human capital.

Furthermore, large parts of the economic literature examine the returns to schooling across developed and developing economies. Some scholars believe that returns to education are higher in developing countries than in industrialised countries (Card, 2001), and there exist some initial empirical suggestions that this may hold in practice (Psacharopoulos and Patrinos, 2004; Montenegro and Patrinos, 2014). However, strong evidence is surprisingly scarce and, due to the lack of comparable data, often difficult to interpret.

As far as gender differentials are concerned, a pattern of returns to education for females exceeding those of males in developing countries does emerge (Peet et al., 2015; Montenegro and Patrinos, 2014). Dougherty (2005) investigates potential channels through which returns for females could be larger than those of males. He argues that discrimination may cause women to accept wage offers that undervalue their characteristics, and that this effect weakens with increasing education levels. Part of the differential could also be attributable to gender differences in the quality of educational attainment, and women's choice to work in sectors where education is valued relatively highly.[1]

These apparent differentials in female and male returns should, however, be interpreted with caution as there are many potential econometric issues involved in their estimation. Card (1999, 2001) outlines potential sources of bias when returns are estimated by OLS using a standard Mincerian model. First, omitted variable bias may arise from the correlation of unobserved ability with both earnings and the level of education. Schooling is not the only systematic source of skill differences; instead, there are other factors, such as an individual's inherent ability, family inputs, quality of education and labour market characteristics that also determine skill acquisition and are likely correlated with education and earnings. In this respect, Heckman et al. (2006) show that a variety of labour market and behavioural outcomes may be attributed to cognitive and non-cognitive skills. Non-cognitive skills and personality traits such as perseverance, motivation, sociability, and emotional stability strongly influence schooling decisions and further affect wages after controlling for schooling decisions, thereby potentially introducing bias in the standard Mincerian wage equation.

Hanushek and Woessmann (2008) investigate the role played by cognitive skills on economic outcomes such as individual earnings, the distribution of income and economic development. They argue that failing to consider cognitive skills and ability strongly biases simple OLS estimates of returns to education based on a Mincerian wage equation. They further provide evidence in favour of relatively higher returns in developing countries compared to developed countries.

Acosta et al. (2015) use STEP data for Colombia to examine the relationship between ability and economic outcomes in labour markets. They show that cognitive and non-cognitive skills are important for these outcomes, and address different channels. Cognitive abilities are important for increased earnings and obtaining more qualified jobs, whereas personality appears to matter more for labour market participation. They also use the Big Five model (discussed below) as a proxy for personality characteristics.

Wichert and Pohlmeier (2010) investigate the specific influence of non-cognitive skills, i.e. personality traits, on female labour force participation. They analyse this relation-

[1] Dougherty does not differentiate between statistical and taste-based discrimination in his analysis.

ship using the Big Five model, a theory proposing five broad dimensions (openness, conscientiousness, extroversion, agreeableness and neuroticism) that may be used to describe the human personality (Costa and McCrae, 1999). They argue that personality traits are influential for the labour participation decision and further show that not taking them into account will overestimate the impact of education on wages. They find a strong impact of extraversion and agreeableness, whereby the first has a positive and the latter a negative effect on wages. Conscientiousness is the Big Five trait with the strongest direct positive effect on wages, while neuroticism and openness have a negative effect of about the same size on the probability of female labour market participation. However, they do not find a direct wage effect of these variables in their study.

Another potential source of omitted variable bias may arise from family background (Card, 2001). Economists have long been interested in the effects of family environment on the subsequent labour market success of individuals. Some suggest that the bias might be larger in less developed economies, since liquidity constraints and family background are relatively more important determinants of both education and earnings for low-income countries. Behrman and Wolfe (1984) identify a large family background bias in returns estimated for females in Nicaragua; Heckman and Hotz (1986) find the same for Panamanian males, while Lam and Schoeni (1993) estimate the bias to be only modest in Brazil.

Further, since earnings are only observed for employed individuals, arguably a non-random sample of the population, sample selection can lead to biased OLS estimates. Individuals with higher (potential) returns to education have an incentive to acquire more schooling, and thus a cross-sectional regression of earnings on schooling is likely to yield an upward biased estimate of the average marginal return to schooling (Card, 2001). Mulligan and Rubinstein (2008) make the case for the U.S. that selection is an important determinant of the gender wage gap and, moreover, observe a changing selection pattern over time. Using Heckman's two-step estimator, they find that selection into the female full-time, full-year workforce shifted from negative in the 1970s to positive in the 1990s, and that much of the supposed narrowing of the gender wage gap reflects changes in female workforce composition.

Finally, in their meta-study, Harmon et al. (2001) show that selection bias in female returns to education in European countries is small, but significant. This means they find significantly different returns for females in European countries when modelling selection compared to estimates that do not rely on a selection model. Additionally, countries with the lowest rates of female participation show the highest differences in educational returns.

3 Description of the data

We use data from the World Bank's Skills Towards Employability and Productivity (STEP) program, captured in 2012 or 2013 depending on the country.[2] The data include survey questions pertinent to this paper's research question, which measure cognitive skills, non-cognitive skills and contain information on parental education for people in the working age population.

The program includes 12 developing and emerging countries. Due to unavailability of certain variables relevant for our research, we are forced to drop some countries from the analysis. We exclude Laos, Sri Lanka, Macedonia and Yunnan since these countries lack data on literacy test scores. Armenia is excluded due to a large proportion of missing values for the parental education variables. Hence, the seven countries remaining are Bolivia (2012), Colombia (2012), Georgia (2013), Ghana (2013), Kenya (2013), Ukraine (2012), and Vietnam (2012). We analyse all seven countries using OLS and quantile regression estimation. However, we exclude Kenya from the analysis of selection bias, due to implausibilities in the labour dependency variable which serves as exclusion restriction in our Heckman two-step procedure.[3]

The variables we use as controls for previously unobserved heterogeneity are reading proficiency as a proxy for cognitive ability; the Big Five personality dimensions, measures of grit and decision making to proxy non-cognitive skills; and parental education (both maternal and paternal) as a proxy for parental background. Variables capturing cognitive and non-cognitive skills are standardised to have mean 0 and standard deviation equal to 1.

The survey measures non-cognitive abilities as follows. For each variable, three questions are asked with answers on a 1 to 4 integer Likert Scale.[4] Subsequently, the scores for the three questions are averaged and then standardised. Literacy rates are assessed through a short reading and comprehension test designed by the Educational Testing Service (ETS), and reported on a scale from 0 to 500 and then standardised. Parental education is measured using the International Standard Classification of Education (ISCED). For each parental gender, the measure has four categories: category 1 (ISCED 0) refers to early childhood education or less; category 2 (ISCED 1) to primary education; category 3 (ISCED 2 & 3) to secondary education; and category 4 (ISCED 4

[2]Although the data have been gathered in either 2012 or 2013 for each of the countries, we assume that returns to education have not changed significantly over this short period of two years. Furthermore, it is not our aim to compare coefficients across countries, but rather to look at the education returns gap for each individual country and how it changes once we control for previously unobserved heterogeneity and selection.

[3]For Kenya, the labour dependency variable only takes values of 1 and 0, which implies that either all or no members of the household work. Since this seems highly implausible, we decided to exclude Kenya from our Heckman analysis.

[4]The Likert Scale is a linear rating scale which transforms individual survey responses into a comparable scale.

and higher) to post-secondary non-tertiary education or higher. For extended summary statistics on variables of interest see Table 5 in the Appendix.

Following Acosta et al. (2015), we include only wage workers in the sample of employed persons. We omit self-employed individuals and family workers because for this group we cannot separate wage earnings from returns to capital, while our hypotheses are focused on returns to education in the labour market. Furthermore, we exclude the top 1% of wage earners from the sample, in which we follow the literature in guarding against outliers skewing our results. We further restrict our sample to only urban residents, since the surveys were aimed overwhelmingly at urban populations. When discussing the outputs, it is therefore important to bear in mind that the findings are specific to the kind of workers addressed in our sample. Although there might be external validity for other types of workers, caution should be exerted in drawing general conclusions since selection patterns might be different for rural as compared to urban areas.

In order to ensure consistency, we use the same sample across all of our regressions. Full regression results including all coefficients are available upon request, as are Stata do files.

4 Estimation strategy

As discussed above, our empirical strategy is threefold. We first test whether there is a significant gap in returns to education across genders separately by country, also investigating possible gaps at different quantiles of the wage distribution. Next, we examine to what extent such gaps, if they exist, can be explained by previously unobserved heterogeneity and, if so explore whether there are certain parts of the wage distribution where this is more likely the case. Finally, we assess whether selection might be driving the differentials (or lack of differentials) in different countries.

4.1 Returns gap across genders in a Mincerian framework

We first investigate whether there is in fact a gender gap in returns to education for wage workers in the countries in our sample, adopting a standard Mincerian framework. We do this by running OLS regressions separately for each country, where we regress the natural logarithm of hourly wages on years of education and potential experience ($age - years\_educ - 6$) and its square.

We also include a dummy for part-time work (less than 30 hours per week) in the basic model, since part-time employment could affect female and male wages differently, as set out by Goldin (2014). Finally, we include a gender dummy, which equals one if the individual is female, and interactions of the dummy with each of the covariates. This allows us both to obtain an estimate for the returns gap by looking at a single coefficient, and to estimate different marginal effects by gender for all the variables.

$$\begin{aligned} \log wage_i = \beta_0 + \beta_1 years\_educ_i + \boldsymbol{\beta_2 mincer}_i + \beta_3 parttime_i + \\ + \beta_4 d_i + \beta_5 years\_educ_i * d_i + \boldsymbol{\beta_6 mincer}_i * d_i + \beta_7 parttime_i * d_i + u_i \end{aligned} \tag{1}$$

In Equation 1, $\boldsymbol{mincer}_i$ is a vector of Mincerian controls ($experience$ and $experience^2$), $parttime_i$ is a dummy variable that takes a value of 1 if the worker works part-time and d_i is a gender dummy of 1 for females and 0 for males.

In addition to reporting the estimated returns gap $\hat{\beta}_5$, we also report the level estimates separately for males and females. This informs our discussion of the estimated returns gap for each country, and allows us to compare the level estimates of returns to education with the existing literature.

Following this, we examine whether the returns gap varies across the wage distribution by running a quantile regression using the same covariates as set out above. This permits us, first, to check whether any significant returns differentials we do find are driven by gaps at certain parts of the wage distribution in the relevant countries. Second, where we find no significant returns gap, we can also check whether in fact there is a gap, but one which is only present for specific parts of the wage distribution.

4.2 Controlling for previously unobserved heterogeneity

After this, we explore how much of the gender gap (if present) in returns to education can be explained by the STEP controls (non-cognitive skills, cognitive skills, and parental background). To control for non-cognitive skills, we add first and second degree polynomials for non-cognitive abilities measured by the Big Five personality traits, plus grit and decision-making ability, following Heineck and Anger (2010). To control for cognitive ability, we add literacy scores, and control for parental background using dummies for the various ISCED categories of parental education, where we set early childhood education or less as reference category.

$$\begin{aligned} \log wage_i = \beta_0 + \beta_1 years\_educ_i + \boldsymbol{\beta_2 mincer}_i + \boldsymbol{\beta_3 STEP}_i + \\ + \beta_4 parttime_i + \\ + \beta_5 d_i + \beta_6 years\_educ_i * d_i + \boldsymbol{\beta_7 mincer}_i * d_i + \boldsymbol{\beta_8 STEP}_i * d_i \\ + \beta_9 parttime_i * d_i + u_i \end{aligned} \tag{2}$$

In regression equation 2, $\boldsymbol{STEP_i}$ is a vector that includes all the above mentioned variables.

Comparing the results from the two sets of regressions with and without the STEP controls allows us to investigate whether accounting for this previously unobserved heterogeneity affects the gap in returns to education. In terms of level estimates, we expect that adding the STEP controls should attenuate returns to education compared to the baseline regression. Moreover, if these controls are more strongly correlated with educational outcomes for females than for males, we should see a larger decrease in the level estimates for females than for males, which would in turn lower the estimated returns differential.

Finally, we also run quantile regressions (Koenker and Hallock, 2001) using the expanded specification with STEP controls, in order to examine whether adding these controls has different effects on returns to education for different parts of the wage distribution.

4.3 Modelling and controlling for selection

Last, we examine the potential impact of controlling for selection on the gender differentials that we have estimated. For this we perform a Heckman two-step procedure. In the first stage we run a probit model with the following specification:

$$\Pr(Wage.worker_i = 1) = \boldsymbol{\Phi(x_i\gamma)} \tag{3}$$

The outcome variable takes on a value of 1 if the individual is a wage worker, and 0 otherwise. There are three groups of independent variables, namely a vector of four variables excluded from the second stage (discussed below), a vector of Mincerian controls, and a vector of STEP controls. We omit the part-time dummy since it is a function of the working status.

We run this regression separately by country and gender, and also compute the inverse Mills ratio (IMR) separately by gender. For the purposes of the second stage regression, we compile the IMR by combining the IMR variables for males and females into a single variable.

The second stage regressions are the same as those in Equations 1 and 2, apart from the fact that the IMR is added as a control variable of selection in each case, yielding the following specification:

$$\begin{aligned}\log wage_i = {} & \beta_0 + \beta_1 years\_educ_i + \boldsymbol{\beta_2 mincer}_i + \\ & + \beta_3 parttime_i + \beta_4 \lambda_i(\boldsymbol{x_i\gamma}) \\ & + \beta_5 d_i + \beta_6 years\_educ_i * d_i + \boldsymbol{\beta_7 mincer}_i * d_i + \\ & + \beta_8 parttime_i * d_i + \beta_9 \lambda_i(\boldsymbol{x_i\gamma}) * d_i + u_i\end{aligned} \tag{4}$$

$$\begin{aligned}\log wage_i = {} & \beta_0 + \beta_1 years\_educ_i + \boldsymbol{\beta_2 mincer}_i + \boldsymbol{\beta_3 STEP}_i + \\ & + \beta_4 parttime_i + \beta_5 \lambda_i(\boldsymbol{x_i\gamma}) \\ & + \beta_6 d_i + \beta_7 years\_educ_i * d_i + \boldsymbol{\beta_8 mincer}_i * d_i + \boldsymbol{\beta_9 STEP}_i * d_i \\ & + \beta_{10} parttime_i * d_i + \beta_{11} \lambda_i(\boldsymbol{x_i\gamma}) * d_i + u_i\end{aligned} \tag{5}$$

4.3.1 Discussion of excludable variables for Heckman procedure

The Heckman procedure requires that one or more independent variables in the first stage probit regression can be excluded from the second stage wage regression ("excludable variables"). This is necessary in order to avoid identification only by the functional form of the inverse Mills ratio. The criteria for the excludable variable(s) are therefore twofold: they need to affect employment status, but must not be correlated with wages given all the other regressors in the wage equation.

A variable previously used in the first stage regression for women in developed countries is a woman's number of children interacted with marital status. However, it appears likely, especially in developing countries, that women that earn lower wages have more children due to a lack of family planning resources at the lower end of the wage distribution. The excludable variables we use in this paper are therefore the labour dependency ratio, household asset index, number of economic shocks before age 15, and socio-economic status at age 15, which we discuss in turn below. These are the same excludable variables used by Valerio et al. (2015), who make use of the same dataset.

The labour dependency ratio is the number of working individuals in a respondent's household divided by the total household size, not taking the respondent into account. The motivation is that a lower ratio implies higher pressure to work on household members. The use of the household asset index is motivated by an expected negative relationship between household assets and the probability of working, resulting from lower financial pressure. As far as the number of economic shocks before age 15 is concerned, we expect that experiencing more economic shocks at a young age should increase the probability of joining the labour force due to increased financial pressure (Verick, 2014). Economic shocks include, for example, the death of a household member or destruction of harvest.

Finally we expect socio-economic status at age 15 to be a positive determinant of labour force participation for males, but a negative one for females. We expect males with a higher socio-economic status at a young age to have a better network and other background characteristics, which increase the probability of finding employment. For females, we expect a negative relationship due to the dynamics of the marriage market, in particular positive assortative mating, as explained in Anukriti and Shatanjaya (2017). Given a wealthier background, women are more likely to marry a wealthy husband, which, in turn, would make them less likely to enter employment.

Summary statistics of the excludable variables are provided in Table 6 in the Appendix. We also discuss the limitations of our exclusion assumptions in more detail in subsection 5.4.

5 Results and interpretation

5.1 Existence and size of estimated returns gaps

5.1.1 Estimation results

Table 1 shows OLS results without and with STEP controls included. The last two columns contain estimates of the difference in returns to education for females and males, defined as returns of females minus returns of males ("the returns gap").

Table 1: OLS estimates of returns to education

	Females		Males		Returns gap	
STEP controls	No	Yes	No	Yes	No	Yes
Bolivia	0.108***	0.103***	0.110***	0.098***	−0.002	0.000
	(0.012)	(0.018)	(0.016)	(0.016)	(0.019)	(0.022)
Colombia	0.110***	0.104***	0.102***	0.086***	0.009	0.036
	(0.016)	(0.016)	(0.015)	(0.019)	(0.022)	(0.024)
Georgia	0.077***	0.061***	0.095***	0.075***	−0.017	−0.015
	(0.015)	(0.016)	(0.018)	(0.019)	(0.022)	(0.022)
Ghana	0.249***	0.166***	0.115***	0.091**	0.134**	0.074
	(0.041)	(0.058)	(0.037)	(0.043)	(0.056)	(0.073)
Kenya	0.134***	0.112***	0.129***	0.117***	0.006	0.000
	(0.014)	(0.017)	(0.012)	(0.013)	(0.016)	(0.018)
Ukraine	0.080***	0.066***	0.019	0.009	0.061**	0.056**
	(0.012)	(0.012)	(0.025)	(0.023)	(0.027)	(0.026)
Vietnam	0.101***	0.088***	0.087***	0.087***	0.014	0.001
	(0.008)	(0.011)	(0.010)	(0.012)	(0.013)	(0.016)

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Returns gap defined as return for females less return for males = $\hat{\beta}_5$ for OLS without STEP controls (see Equation 1) and $\hat{\beta}_6$ for OLS with STEP controls (see Equation 2). See Table 5 for summary statistics and the sample sizes for each regression.

5.1.2 Level estimates of the returns to education for females and males

Looking at level returns in columns 1 and 3 of Table 1, we see that most countries show returns to education of around 10%.[5] This is consistent with Montenegro and Patrinos

[5]Note that our dependent variable is measured in natural logs. In order to interpret the marginal effect of one additional year of education on wages, as opposed to $\log wage$, we need to transform our $\hat{\beta}$ estimates to marginal effect $= e^{\hat{\beta}}$. Therefore, the estimates displayed in the tables will not match the numbers mentioned in the text.

(2014), who estimate the world average returns to education at 9.7%. Moreover, our pattern of findings across geographical regions is similar to what they find, suggesting returns to be on average higher in less developed countries for both females and males. For example, returns in Kenya and Ghana, classified as lower-middle-income countries by the World Bank, are 14.0% and 17.6% - considerably above the 8.7% returns in Georgia and 11.1% in Colombia, which the World Bank classifies as upper-middle-income countries. In our dataset, Ghana and Kenya are the countries with the lowest hourly wages for both genders, while Colombia and Georgia are among countries with the highest wages (see Table 5 in the Appendix).

Despite this overall result, some countries demand particular attention. The high point estimate of 28.3% for female returns to education in Ghana is striking, and cannot be attributed solely to noise in the data, since the standard error is only approximately 0.04. This result may be due to a variety of factors. One of these could be the fact that females in Ghana work mostly in the service and sales sector (Ghana Statistical Service, 2014). This is confirmed in our data source, which shows that 62% of working females in Ghana engage in the service and sales sectors, compared to 20% of working males. We do not observe such a high employment difference between females and males in the service sector of any of the other countries in our sample. Since higher levels of education are likely to be more valuable in these sectors, this could serve to explain, at least partly, why returns to education for females in Ghana are so high.

The returns of Ukrainian males also stand out, with a point estimate of only 1.9%, a relatively low figure and insignificantly different from zero. The upper bound of this estimate's 95% confidence interval is 7.0%, and it therefore appears to be significantly lower than the returns for males in any of the other countries in our sample. Hence, this represents strong evidence of exceptionally low returns to education for males in the Ukraine. This finding may be explained by skills mismatch and overeducation in transition economies (Kupets, 2015). Misalignment of the education with labour market needs in transitions economies may cause large imbalances between the supply and demand for skills. In Ukraine, males seem to be more prone to being overeducated for their occupation than females, with 10.1% of males being overeducated for their jobs compared to 7.9% of females. Meanwhile, undereducation is more common among females (Kupets, 2010). In addition, a further explanation refers to an institutional setting specific to the Ukrainian labour market, where wages of both low and high skilled workers are determined through a political process as opposed to market dynamics (Svejnar, 2010). This system may prevent employers from rewarding skills according to their market value, which would suppress the wages of highly skilled workers.

5.1.3 Estimates of the gender gap in returns to education

Regarding our estimates of the gaps in returns to education without STEP controls, we see that they are either small or not statistically different from zero for five of the countries in our sample. However, we do find a statistically significant returns differential in favour of females in Ghana and Ukraine, where the estimated returns gaps are 14.3% and 6.3%, respectively.

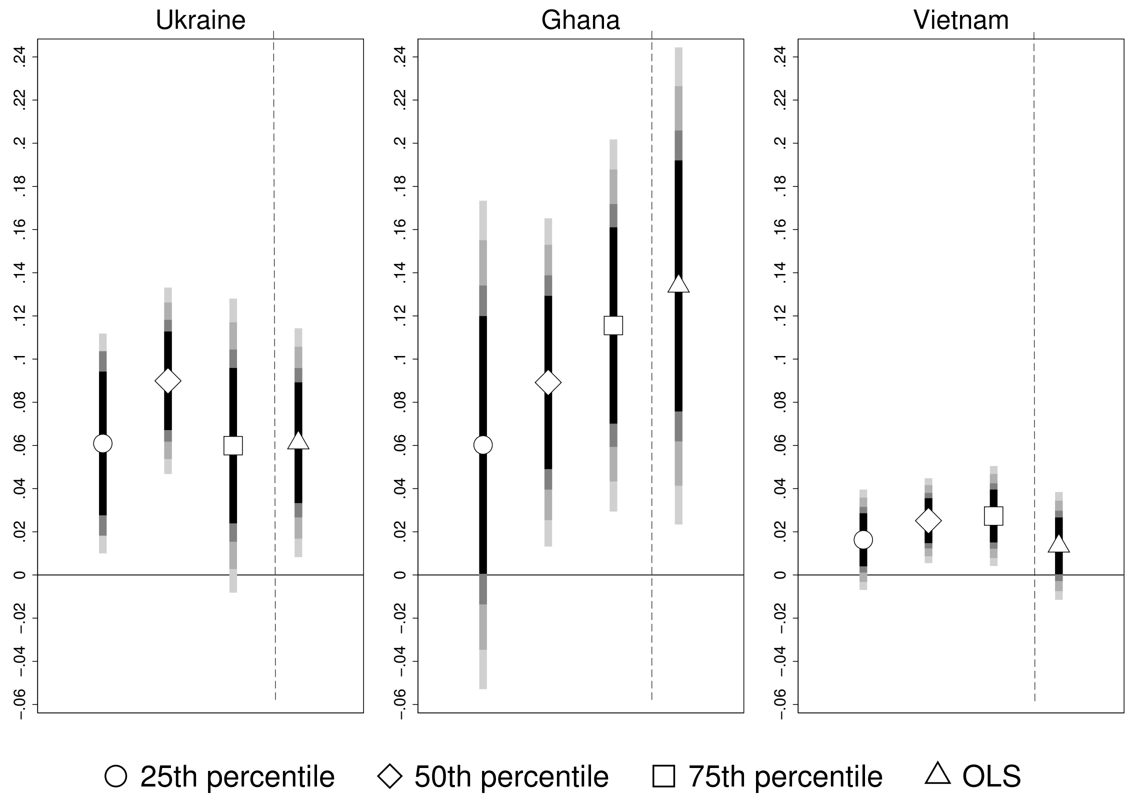
In this respect, it is important to differentiate between insignificant estimates of the returns gap with large standard errors on the one hand, and those with small standard errors on the other. Large standard errors imply large confidence intervals, which are a result of imprecisely estimated coefficients. On the other hand, a confidence interval tightly centred around zero allows one to draw a conclusion about the non-existence of a gender differential in returns to education. In Table 1 we observe such tight confidence intervals for Bolivia, Colombia, Georgia, Kenya and Vietnam. Hence, it seems that, under the baseline specification, there are no additional returns to education for female wage workers in the urban labour market in these countries, or that any returns differential is very small.

We note that for the two countries where we find significant returns gaps, the underlying reasons differ. Whereas in Ukraine this seems to be driven by exceedingly low male returns (8.3% for females and 1.9% for males), in Ghana it is driven by exceptionally high female returns (28.3% for females and 12.1% for males). This is in line with the reasoning set out above, according to which there appear to be characteristics specific to the labour markets in these countries which give rise to low returns for males in Ukraine and high returns for females in Ghana.

In order to gain some insights on whether OLS results on the return gaps are homogeneous over the entire wage distribution, we estimate quantile regressions for three quartiles of the wage distribution and report the significant findings of interest in Figure 1. We find that in the majority of countries where we could not reject the null hypothesis of no gap in the OLS framework, there are also no significant gaps arising at any of the three quartiles. More specifically, Bolivia, Colombia, Georgia, and Kenya show no significant gap at any of the three quartiles at the 10% significance level. This further validates our OLS findings for these countries.

In Ghana, where our Mincerian OLS results without STEP controls suggested a positive and significant returns gap, we observe that a significant gap (at the 5% level) is in fact only present at the second and the third quartile of the wage distribution (9.3% and 12.3% respectively). On the contrary, in Ukraine, significant gaps (at the 10% level) of similar magnitude are present across all three quartiles of the wage distribution (6.3.%, 9.4% and 6.3% respectively). Interestingly, in Vietnam, where we could not observe a significant gap using OLS, we find significant gaps (at the 5% level) at the upper two quartiles of the wage distribution (2.3% and 2.5% respectively).

Figure 1: Returns gap estimates (quantile regressions without STEP controls)



Shaded areas indicate 70% (darkest), 80%, 90% and 95% (lightest) confidence intervals. Returns gap defined as return for females less return for males = $\hat{\beta}_5$ for OLS without STEP controls (see Equation 1). See Table 7 for the quantile regressions for all countries, with and without STEP controls. See Table 5 for summary statistics and the sample sizes for each regression.

5.2 Effect of adding STEP controls

The level estimates of returns to education when the STEP controls are added are reported in columns 2 and 4 of Table 1. For all countries (except for Ghana, whose results we discuss below) the point estimates for both genders decrease slightly by around 0 to 2.5 percentage points. Looking at the coefficients of the STEP variables (unreported), we observe the expected positive relationship between wages and cognitive skills, non-cognitive characteristics and family background. Therefore controlling for them explains part of the job market returns otherwise attributed to education.

On the one hand, the changes in the level estimates are not substantial, and the new estimates fall comfortably within the confidence intervals of the previous estimates without the STEP controls. On the other hand, the small decrease in the estimated returns is consistent across six of the seven countries. Therefore our results suggest that previously unobserved heterogeneity with regard to the STEP controls could have a small (but not necessarily unimportant) effect for estimating the level estimates for returns to education in these six countries, and that excluding them might lead to a small overestimation of returns. However, since adding the STEP controls does not appear to have a different effect across genders, it appears that omitting them is unlikely to introduce serious bias for estimates of the gender gap in returns to education in these six countries.

Turning to Ghana, we observe that the estimated returns to education for females decrease by just more than 10 percentage points when adding the STEP controls. As a result, the estimated returns gap falls by 6.7 percentage points and becomes insignificant. This is consistent with the earlier discussion of the specific characteristics of the labour market in Ghana. Since females in Ghana work proportionally more in the sales and services sectors than males, and more than females in other countries in the sample, it is plausible that the STEP controls play a more important role in determining wages for females in Ghana than for males, or for females in other countries.

In support of this possibility, we observe that in the level regression with STEP controls for females in Ghana, the only individually significant STEP variables are agreeableness (one of the Big Five personality traits) and its square, both with positive coefficients and significance at the 10% level. Together with Vietnamese females, Ghanaian females are the only group for either of the two variables is significant. Given the high prevalence of female Ghanaian workers in the sales and services sector, it makes sense that a trait such as agreeableness would be positively related to the wages (and educational outcomes) of Ghanaian women, and that controlling for it reduces the estimated returns gap substantially. This stands in contrast with the findings of Costa and McCrae (1999) that agreeableness has a negative effect on wages.

When we investigate these results further using quantile regressions with the STEP controls, we see that for Bolivia, Colombia, Georgia, and Kenya our initial finding of a lack of significant returns gaps at all three quartiles of the distribution does not change once we add the STEP controls (see Table 7 in the Appendix).

For Ghana, we observed a returns gap at the second and the third quartile of the wage distribution under the baseline specification, and once we add the STEP controls a significant gap survives only for the highest quartile. Therefore adding the STEP controls only makes a difference for the estimated gap at the median. This indicates that, even in Ghana, adding the STEP controls does not account for the entire gap, showing that Ghana is not as much of an exception as the results discussed above appeared to suggest. Instead, the effects involving agreeableness and women in the sales and services sector are likely to be substantial only for the middle of the wage distribution.

In Ukraine, we found previously that returns gaps of similar magnitude are present across all three quartiles of the wage distribution under the baseline specification. After adding STEP controls, the gap persists only at the median. This shows that our earlier finding that adding the STEP controls does not account for the gender gap in Ukraine holds only for the middle of the wage distribution.

Finally, in Vietnam, where we found significant returns gaps at the upper two quartiles of the wage distribution under the baseline specification, we find no significant gaps at any of the quartiles once we add the STEP controls. This once again shows that the STEP controls may have an important impact on estimated returns for some quartiles of the distribution, although for estimates of the conditional mean gap, these controls do not appear to play an important role.

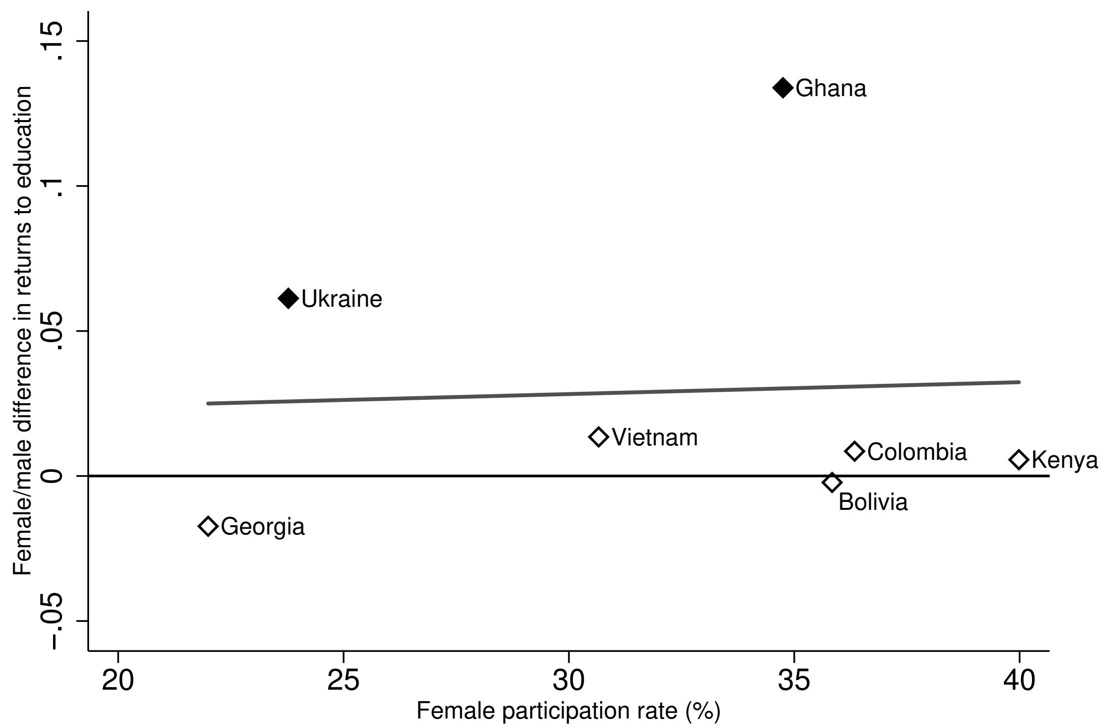
In summary, it appears from the data that estimates of returns gaps in returns to education are, in general, not driven strongly by heterogeneity in skills, cognitive abilities or parental background. Only for Ghana do differences in these characteristics seem to explain to some extent the female premium in returns to education, while in the case of Ukraine and Vietnam, any material effect of adding the controls is limited to specific quartiles of the wage distribution.

5.3 Effect of controlling for selection

Before undertaking the Heckman two-step estimation procedure, we investigate to what extent the returns gaps we have estimated may be related to female participation rates. This is in view of the fact that, as we have discussed above, the size of estimated gender returns gaps could be related to the female labour force participation rate, as low participation rates make endogenous selection more likely. This also allows for a comparison of our results with those of Harmon et al. (2001), who suggest that there may exist a negative relationship between female participation rates and gender differentials in returns to education.

As can be seen from Figure 2, which shows the estimated returns differential (without STEP controls) and the female labour force participation rate for each country in our sample, our results do not demonstrate this relationship. Instead, we find that most of the countries' returns gaps are centred around zero, regardless of their female participation rates, while countries that experience a significant gap do not have similar female participation rates. This therefore provides an initial indication that selection of female workers may not be a very important factor in explaining returns gaps (or a lack of returns gaps) for countries in our sample.

Figure 2: Female participation rates and estimated returns gaps (without STEP controls)



Dark markers indicate statistically significant gaps (5% level), hollow markers are not statistically significant. Line of best fit included. Returns gap defined as return for females less return for males = $\hat{\beta}_5$ for OLS without STEP controls (see Equation 1). Female participation rate defined as the ratio of active females (employed and unemployed) in the labour market.

In the rest of this subsection, we discuss the results from the Heckman two-step procedure using the full specification (including the STEP variables as controls), and also compare these results to the full-specification OLS results. We have also run the Heckman procedure without the STEP controls, but only briefly refer to these results below since they do not materially affect our conclusions, and report them in Table 4 in the Appendix. We first describe the results from the first stage and the second stage regressions, and then provide an interpretation of the results.

The results from the first stage are reported in Table 5 in the Appendix, which includes the outcomes' joint F-tests of the relevance of the four excludable variables for employment status. As can be seen from the table, the excludable variables are clearly relevant for some countries and genders, while in other cases they are only marginally or not relevant. We discuss the implications of this in Subsection 5.4 below.

One important result from the first stage regression is that the education variables do have a highly significant (at the 1% level) and positive impact on the probability of being a wage worker for all countries and across both genders, except for Colombian males. This means that, ceteris paribus, higher education on average increases the probability of being employed as a wage worker in urban areas for these countries.

From the second stage regression results, reported in Table 2, we see that while Bolivia and Colombia displayed insignificant estimates of the returns gaps before controlling for selection, both countries now show a significant gender returns gap, Bolivia at a 5% and Colombia at a 1% significance level. The point estimates for the returns gap after correcting for selection are 11 percentage points in Bolivia and 7.5 percentage points in Colombia, while these estimates were centred closely around zero in the OLS estimation. Therefore, if we do not control for selection using the inverse Mills ratio, we underestimate the returns gap between females and males in these countries.

For Ukraine, controlling for selection does not alter the estimated returns gap materially, since it is now 7.3 percentage points, which is comparable to (albeit a little larger than) the OLS estimate of 5.9 percentage points. In the rest of the countries, we still find no significant gender gap in returns to education. Specifically, for Ghana we see that the gap remains insignificant when we control for selection. We also see that, if we do not control for the STEP variables (see Table 4 in the Appendix) controlling for selection results in a lower and insignificant point estimate of the returns gap for Ghana compared to the OLS estimate without STEP controls (Table 1). However, in this specification the coefficients for the IMR are not significant for either gender. Hence, for Ghana it seems that the gender returns gap can be partially explained by differences in the STEP variables but not by non-random selection.

Table 2: Second stage Heckman results (including STEP controls)

	Bolivia	Colombia	Ghana	Georgia	Ukraine	Vietnam
Returns gap	0.104**	0.072*	0.038	0.163	0.070**	0.006
	(0.047)	(0.041)	(0.039)	(0.216)	(0.036)	(0.028)
Female IMR	1.819***	1.415**	−0.119	0.619	0.081	0.670*
	(0.377)	(0.722)	(0.302)	(1.042)	(0.289)	(0.407)
Male IMR	0.542	−0.712**	−0.801*	−0.383	−0.097	0.401
	(0.358)	(0.344)	(0.412)	(0.937)	(0.374)	(0.558)
R^2	0.461	0.393	0.290	0.432	0.359	0.329
N	642	669	661	201	726	1052
PSUs	165	387	194	115	256	225

Note: *** p<0.01, ** p<0.05, * p<0.1 Standard errors in parentheses. Gap in returns defined as return for females less return for males = $\hat{\beta}_7$ for for Heckman with STEP controls (see Equation 5). Second stage results without STEP controls are reported in Table 8, and the first stage results for both specifications is reported in Table 9 and 10. See Table 5 and 6 in the Appendix for summary statistics (also on the excludable variables).

It is important to understand how the inclusion of the IMR affects the results. The first stage probit regression shows that education has a positive impact on the probability of working. Since the IMR is, by construction, negatively related to the probability of working, it is smaller for relatively educated individuals. By the same reasoning, uneducated individuals who are wage workers show a larger value for the IMR, which implies that selection matters more for them relative to educated individuals.

The direction of the coefficients of the IMR in the second stage reveals whether men and women are positively or negatively selected. A positive coefficient implies positive selection, while a negative coefficient implies the opposite. Intuitively, positive selection means that individuals who are in the sample of workers have unobserved characteristics which are positively correlated with their wages relative to non-workers. By the same logic, negative selection means that these unobserved characteristics are negatively correlated with wages.

Taken together with the first stage results, this means that where we find positive selection for females, the selection effect is stronger for uneducated females than for educated females. In other words, the effect of selection is negatively related to education levels. In turn, this means that, by not accounting for selection, we would underestimate the returns gap.

This can be seen from Equation 6 below, where $wage_{f,h}$ are wages of highly educated females, $wage_{f,l}$ wages of low educated females and similarly for males. Not accounting for selection means that $wage_{f,l}$ is overestimated, which means that the returns gap is underestimated.

$$returns\ \ gap = (wage_{f,h} - wage_{f,l}) - (wage_{m,h} - wage_{m,l}) \tag{6}$$

The two countries where the results change substantially, Bolivia and Colombia, are the only ones where we find a significant selection effect for females. In both these countries, the coefficient of the IMR is positive and significant at the 1% level in the second stage regression, which means that females are positively selected in these countries. As set out above, this explains why the estimated returns gap increased after we controlled for selection.

We can also see from the IMR coefficient that males in Colombia are in fact negatively selected, and the same applies for Georgia. For Colombia, the underestimation of the gender returns gap without correcting for selection is therefore amplified by the fact that males show a negative selection pattern.

In summary, the results from the Heckman two stage regressions show that, while we do not find strong evidence that that selection plays an important role in Georgia, Ghana, Ukraine or Vietnam, selection does matter in Bolivia and Colombia, and it is driven mainly by the selection of females. Specifically, positive female selection leads to an underestimation of the returns gap because it is negatively correlated with education if it is not explicitly accounted for.

5.4 Limitations and potential extensions

The most general limitations of our findings relate to the nature of our sample and the data. First, due to the complex survey design of the data, sample variances and covariances are calculated based on the number of clusters, and not the number of individuals included in the sample for each regression. These relatively small effective sample sizes mean that in some cases we have obtained point estimates with large standard errors, which are not conducive to inference. A possible extension of our paper would therefore be to undertake the same analysis using a larger number of clusters.

Second, our sample of countries is heterogeneous in a number of respects, including geographical location, degree of development and political history. It is therefore to be expected that we do not find a clear pattern across all of the countries in our sample. One possibility for future research would therefore be to investigate our results in more detail for each country individually, for example by looking at which specific economic features of each country give rise to the results we obtain. As far as the results from the quantile regressions are concerned, it would also be interesting to investigate the different effects at play at different quantiles of the wage distribution of each country. Although we have provided some initial indications of what may be driving the results we obtain for countries such as Ghana and Ukraine, a detailed analysis of each country's economic and labour market dynamics falls outside the scope of this master's project.

Although it has not been the main focus of this paper, there are also some limitations specific to the Heckman procedure. First, as in all labour economics papers that use this approach, there is a risk that the exclusion restrictions do not hold for the four variables we have identified as affecting employment status, and being exogenous to the wage after controlling for the Mincerian and STEP variables. This can potentially introduce bias into the second stage Heckman regressions. Nevertheless, in our view, the excludable variables we have used appear to be more justifiable than other frequently used in the literature, such as the number of children interacted with marital status. Furthermore, they allow us to model selection for males. In using these four variables, we also follow other papers that have used the STEP data.

Second, further concerns may be raised by the results of the first stage probit regression, both because for some countries the excludable variables do not appear to have a high degree of relevance in explaining employment status, and because the signs of the coefficients of some of the excludable variables change across genders and across different countries. It is beyond the scope of this paper to examine the possible explanations of the signs and significance of the excludable variables for each of the countries in our sample, and as mentioned above, in using these four variables we have followed previous research using the STEP data. In order to further refine this analysis, future research could concentrate on each country individually in terms of identifying valid excludable variables, and examining which features of these countries' economies and labour markets explain the signs and significance of these variables in the first stage of the Heckman procedure.

6 Conclusion

This paper explores gender gaps in returns to education for seven developing and emerging countries. First, we investigate the existence of such a gap in a standard Mincerian framework. We find a significant returns gap in only two countries, namely Ukraine and Ghana, while the estimates for the other countries are centred relatively tightly around statistically insignificant point estimates close to zero. Using quantile regressions to dig deeper does not materially affect our findings, although it does allow us to specify that the returns gaps estimated for Ghana and Ukraine are significant at two out of three quartiles of the wage distribution, and that in Vietnam there is a small but significant returns gap at the upper two quartiles of the distribution. These findings are important in providing context for the existing literature, showing that returns premiums in favour of females are not universally prevalent in developing countries for urban wage workers. This suggests that where large, significant returns gaps have been found in the literature, this seems to be driven to a large extent by other segments of the labour market.

Second, we use our novel dataset to analyse the extent to which controlling for previously unobserved heterogeneity, namely cognitive skills, personality traits and family background, affect OLS estimates of the returns gap. We find that controlling for these STEP variables does not materially affect our baseline estimates for Bolivia, Colombia, Georgia, Kenya and Vietnam (where the estimated gap remains insignificant and close to zero), or for Ukraine, where the estimated gap is of similar magnitude and remains significant. Only in Ghana we find that adding the STEP controls has a material effect, reducing the point estimate of the gap substantially and rendering it insignificant. The results of the quantile regressions qualify this finding somewhat, showing that controlling for the STEP variables does make a difference for estimates of the gap at certain quantiles of the distribution in Ukraine and Vietnam. Overall, our finding regarding the importance of these sources of previously unobserved heterogeneity is cautiously negative: although they do appear to make a small difference for the level estimates and have an important effect in Ghana, they do not appear to be universal sources of endogeneity in estimating the returns gap for urban wage workers.

Third, we examine the importance of controlling for selection in estimating the returns differential using the Heckman two-step procedure, dropping Kenya from our sample due to missing data. Here we find that after controlling for selection, our point estimates of the returns gap remain insignificant in Ghana, Georgia and Vietnam, albeit with a relatively high point estimate in Georgia. Similarly, our estimate of the returns gap in

Ukraine does not change considerably and remains significant. In contrast, we obtain higher and significant point estimates of the returns gap in Bolivia and Colombia. As explained above, this somewhat counter-intuitive result is due to positive selection of females into employment in Bolivia and Colombia, and the positive relationship between education levels and probability of employment. Interestingly, in the two countries where selection appears to be important, we found earlier that controlling for the STEP variables did not have an observable effect. Our findings therefore suggest that it is likely to be important to control for selection when estimating returns gaps in developing countries, even if only to exclude the possibility of selection bias. In addition, our approach suggests that selection is likely to operate through channels other than cognitive or non-cognitive abilities, or parental background.

Taken together, our findings show that, at least for urban wage workers in the countries in our sample, a returns premium for females may not be as prevalent as previously suggested. We also find that controlling for potential sources of endogeneity, such as unobserved heterogeneity and selection, substantially changes the estimates of the gender returns gap in three out of seven of the countries in our sample. This highlights the importance of considering these channels to avoid the risk of biased estimation. This paper therefore represents a starting point for more detailed research, which could zoom in on the existence and drivers of returns differentials in individual countries, and overcome some of the limitations of this paper by extending it to rural areas and using samples with a larger number of clusters. These findings are also relevant to policy makers, since they demonstrate the importance of understanding the characteristics and dynamics of each country's individual labour market prior to making policy proposals.

References

Acosta, P., Muller, N., and Sarzosa, M. A. (2015). Beyond qualifications: Returns to cognitive and socio-emotional skills in colombia. Technical report, World Bank Group.

Anukriti, S. and Shatanjaya, D. (2017). Marriage markets in developing countries. *IZA Discussion Paper No. 10556.*

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education*. New York : National Bureau of Economic Research: Distributed by Columbia University Press.

Behrman, J. R. and Wolfe, B. L. (1984). The socioeconomic impact of schooling in a developing country. *The Review of Economics and Statistics*, 66(2):296.

Card, D. (1999). *The Causal Effect of Education on Earnings*, chapter 30, pages 1801–1863. North Holland.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1260.

Costa, P. and McCrae, R. (1999). *Handbook of Personality: Theory and Research*, chapter 3 - The Five-Factor Theory of Personality, pages 193 – 153. Guilford, New York.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men? *Journal of Human Resources*, XL(4):969–988.

Ghana Statistical Service (2014). Ghana living standard survey round 6 (GLSS6). http://www.statsghana.gov.gh/docfiles/glss6/GLSS6_Labour%20Force%20Report.pdf.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):10911119.

Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–668.

Harmon, C., Walker, I., and Westergaard-Nielsen, N. (2001). *Education and earnings in Europe: a cross country analysis of the returns to education*. Cheltenham: Edward Elgar.

Heckman, J. and Li, X. (2004). Selection bias, comparative advantage and heterogeneous returns to education: Evidence from china in 2000. *Pacific Economic Review*, 9(3):155–171.

Heckman, J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Heckman, J. J. and Hotz, V. J. (1986). An investigation of the labor market earnings of panamanian males evaluating the sources of inequality. *Journal of Human Resources*, 21(4):507.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in germany. *Labor Economics*, 17(3):535–546.

Koenker, R. and Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143156.

Kupets, O. (2010). Gender mainstreaming at the labour market of ukraine and role of the public employment service. International Labour Office.

Kupets, O. (2015). Skill mismatch and overeducation in transition economies. *IZA World of Labor*, 2015(224).

Lam, D. and Schoeni, R. F. (1993). Effects of family background on earnings and returns to schooling: Evidence from brazil. *Journal of Political Economy*, 101(4):710–740.

Mincer, J. (1974). *Schooling, experience, and earnings*. National Bureau of Economic Research.

Montenegro, C. E. and Patrinos, H. A. (2014). Comparable estimates of returns to schooling around the world. Technical report, World Bank Group.

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

Peet, E. D., Fink, G., and Fawzi, W. (2015). Returns to education in developing countries: Evidence from the living standards and measurement study surveys. *Economics of Education Review*, 49:69–90.

Psacharopoulos, G. and Patrinos, H. (2004). Returns to investment in education: a further update. *Education Economics*, 12(2):111–134.

Svejnar, J. (2010). Ukraine - labor market and policy response. World Bank.

Valerio, A., Sanchez Puerta, M. L., Tognatta, N., and Monroy-Taborda, S. (2015). Are there skills payoffs in low and middle income countries? empirical evidence using step data. World Bank.

Verick, S. (2014). Female labor force participation in developing countries. *IZA: World of Labor*, 2014:87.

Wichert, L. and Pohlmeier, W. (2010). Female labor force participation and the big five. Technical report, ZEW (Zentrum für Europäeische Wirtschaftsforschung)/Center for European Economic Research.

World Bank (2012). The step skills measurement program. http://microdata.worldbank.org/index.php/catalog/step.

Appendix

A1 Summary statistics

Table 3: Labour market status, proportions by gender

	Employed		Unemployed		Inactive	
	Female	Male	Female	Male	Female	Male
Bolivia	0.68	0.78	0.07	0.05	0.25	0.17
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Colombia	0.58	0.77	0.12	0.09	0.30	0.14
	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Georgia	0.27	0.37	0.20	0.28	0.53	0.35
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Ghana	0.68	0.73	0.05	0.05	0.27	0.22
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Kenya	0.49	0.69	0.18	0.10	0.33	0.21
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ukraine	0.52	0.65	0.05	0.10	0.44	0.26
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
Vietnam	0.65	0.71	0.02	0.03	0.33	0.26
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)

Note: Linearised standard errors in parentheses. The employed group includes both wage workers and self-employed individuals.

Table 4: Worker type, proportions by gender

	Wage-worker		Self-employed	
	Female	Male	Female	Male
Bolivia	0.46	0.57	0.54	0.43
	(0.02)	(0.02)	(0.02)	(0.02)
Colombia	0.48	0.57	0.52	0.43
	(0.03)	(0.03)	(0.03)	(0.03)
Georgia	0.88	0.80	0.12	0.20
	(0.02)	(0.03)	(0.02)	(0.03)
Ghana	0.21	0.51	0.79	0.49
	(0.01)	(0.03)	(0.01)	(0.03)
Kenya	0.49	0.61	0.51	0.39
	(0.02)	(0.02)	(0.03)	(0.03)
Ukraine	0.93	0.83	0.07	0.17
	(0.01)	(0.03)	(0.01)	(0.03)
Vietnam	0.51	0.58	0.49	0.42
	(0.02)	(0.02)	(0.02)	(0.02)

Note: Linearised standard error in parentheses. The sample for this table is only employed individuals.

Table 5: Summary Statistics, OLS variables

	Bolivia		Colombia		Georgia		Ghana		Kenya		Ukraine		Vietnam	
	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male
Hourly wages (USD)	3.73	4.50	3.64	4.95	3.63	4.96	3.13	3.44	3.17	3.33	3.17	4.65	3.51	3.96
	(0.26)	(0.34)	(0.26)	(0.39)	(0.20)	(0.38)	(0.50)	(0.65)	(0.25)	(0.22)	(0.12)	(0.22)	(0.15)	(0.18)
Years of education	12.79	12.56	11.26	11.08	15.75	15.18	14.42	13.61	9.77	10.59	13.83	13.51	12.62	12.18
	(0.28)	(0.32)	(0.25)	(0.19)	(0.14)	(0.21)	(0.29)	(0.29)	(0.37)	(0.28)	(0.12)	(0.20)	(0.21)	(0.24)
Experience	11.94	13.42	15.52	15.24	19.21	17.29	10.03	11.55	15.52	15.69	22.81	18.75	17.01	19.61
	(0.78)	(1.07)	(1.03)	(0.74)	(0.64)	(0.99)	(1.25)	(0.88)	(0.83)	(0.65)	(0.65)	(0.98)	(0.54)	(0.81)
Av. literacy score	−0.13	−0.05	0.36	0.41	0.48	0.35	−0.01	−0.15	−0.41	−0.39	0.75	0.76	0.54	0.47
	(0.07)	(0.08)	(0.05)	(0.04)	(0.03)	(0.04)	(0.08)	(0.08)	(0.07)	(0.06)	(0.03)	(0.04)	(0.04)	(0.04)
Grit	3.00	2.95	2.96	2.92	2.97	2.83	2.69	2.83	2.69	2.72	2.82	2.76	2.75	2.77
	(0.04)	(0.05)	(0.06)	(0.05)	(0.03)	(0.04)	(0.09)	(0.06)	(0.04)	(0.03)	(0.04)	(0.07)	(0.02)	(0.03)
Decision making	3.06	2.96	3.23	3.05	3.43	3.33	3.17	3.15	3.16	3.14	3.23	3.09	2.98	2.95
	(0.04)	(0.05)	(0.06)	(0.04)	(0.02)	(0.04)	(0.07)	(0.06)	(0.03)	(0.03)	(0.03)	(0.06)	(0.03)	(0.03)
Extraversion	3.04	3.05	3.04	3.03	2.65	2.53	2.66	2.57	2.84	2.86	2.73	2.56	2.75	2.75
	(0.05)	(0.05)	(0.05)	(0.05)	(0.02)	(0.04)	(0.07)	(0.05)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)
Conscientiousness	3.16	3.15	3.39	3.29	3.24	3.17	3.28	3.36	3.28	3.25	3.16	3.01	2.83	2.88
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.06)	(0.05)	(0.03)	(0.03)	(0.03)	(0.06)	(0.02)	(0.03)
Openness	3.23	3.21	3.19	3.20	3.12	3.00	3.20	3.20	3.00	3.04	3.17	3.12	2.89	2.93
	(0.03)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.06)	(0.05)	(0.04)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)
Stability	2.36	2.60	2.50	2.75	2.59	2.74	2.64	2.82	2.67	2.76	2.43	2.82	2.77	3.12
	(0.05)	(0.05)	(0.06)	(0.05)	(0.04)	(0.05)	(0.06)	(0.05)	(0.03)	(0.02)	(0.05)	(0.05)	(0.02)	(0.02)
Agreeableness	3.03	3.10	3.15	3.22	3.21	3.11	3.19	3.06	3.88	2.85	3.01	2.85	3.03	3.00
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.07)	(0.06)	(0.04)	(0.03)	(0.03)	(0.05)	(0.02)	(0.03)
Father educ. (av. ISCED)	1.51	1.41	1.38	1.44	2.66	2.57	2.15	2.18	1.62	1.46	2.27	2.34	1.41	1.31
	(0.10)	(0.09)	(0.06)	(0.05)	(0.03)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	(0.04)	(0.04)	(0.05)	(0.05)
Mother educ. (av. ISCED)	1.06	1.10	1.33	1.38	2.61	2.53	1.70	1.72	1.40	1.21	2.27	2.34	1.16	1.00
	(0.10)	(0.09)	(0.07)	(0.05)	(0.03)	(0.04)	(0.10)	(0.08)	(0.07)	(0.06)	(0.04)	(0.07)	(0.05)	(0.05)
N	349	351	346	383	467	259	109	199	396	705	527	280	610	501

Note: Linearised standard errors in parentheses. Sample includes only wage workers, and is limited to the observations included in the OLS regressions (with and without STEP controls) and the Heckman regressions. In other words, any observations with one or more missing values for any of the variables included in those regressions are omitted from the sample for the whole of the analysis to ensure a consistent sample.

Table 6: Summary Statistics, excludable variables

	Bolivia		Colombia		Georgia		Ghana		Kenya		Ukraine		Vietnam	
	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male
Labour dependency ratio	0.56	0.57	0.58	0.58	0.57	0.52	0.60	0.67	-	-	0.66	0.65	0.62	0.62
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	-	-	(0.01)	(0.02)	(0.01)	(0.01)
Economic shocks before 15	1.34	1.38	0.90	0.78	0.21	0.30	0.50	0.69	0.98	1.05	0.24	0.26	0.44	0.42
	(0.11)	(0.15)	(0.10)	(0.08)	(0.03)	(0.05)	(0.09)	(0.11)	(0.09)	(0.07)	(0.03)	(0.06)	(0.04)	(0.04)
Asset index	0.00	0.14	−0.08	−0.04	0.16	0.10	0.79	0.57	0.20	0.02	0.09	0.09	0.08	−0.05
	(0.08)	(0.09)	(0.06)	(0.06)	(0.06)	(0.08)	(0.13)	(0.10)	(0.08)	(0.07)	(0.08)	(0.10)	(0.06)	(0.08)
Socioeconomic status at 15	1.86	1.74	1.87	1.80	2.48	2.19	2.23	2.06	1.93	1.87	2.06	2.05	1.85	1.76
	(0.05)	(0.04)	(0.05)	(0.06)	(0.03)	(0.04)	(0.06)	(0.05)	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)

Note: Linearised standard errors in parentheses. Sample includes only wage workers, and is limited to the observations included in the OLS regressions (with and without STEP controls) and the Heckman regressions. In other words, any observations with one or more missing values for any of the variables included in those regressions are omitted from the sample for the whole of the analysis to ensure a consistent sample.

A2 Results of quantile regression

Table 7: Quantile estimates of gender gap in returns to education

	25th percentile		50th percentile		75th percentile	
STEP controls	No	Yes	No	Yes	No	Yes
Bolivia	0.001	0.004	0.010	0.002	0.023	0.020
	(0.023)	(0.033)	(0.016)	(0.021)	(0.023)	(0.027)
Colombia	0.026	0.004	0.014	0.011	0.004	0.007
	(0.019)	(0.022)	(0.024)	(0.023)	(0.018)	(0.022)
Georgia	0.048	0.035	0.028	0.013	−0.014	0.000
	(0.031)	(0.039)	(0.024)	(0.026)	(0.026)	(0.025)
Ghana	0.060	0.012	0.089**	0.114	0.116***	0.096**
	(0.056)	(0.133)	(0.039)	(0.070)	(0.044)	(0.039)
Kenya	0.011	0.001	0.022	0.007	0.007	0.018
	(0.021)	(0.030)	(0.015)	(0.025)	(0.016)	(0.016)
Ukraine	0.061**	0.059	0.090***	0.053***	0.060*	0.039
	(0.026)	(0.039)	(0.022)	(0.025)	(0.035)	(0.032)
Vietnam	0.016	−0.012	0.025**	0.022	0.027**	0.024
	(0.012)	(0.018)	(0.010)	(0.016)	(0.012)	(0.015)

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors in parentheses. Returns gap defined as return for females less return for males = $\hat{\beta}_5$ for OLS without STEP controls (see Equation 1) and $\hat{\beta}_6$ for OLS with STEP controls (see Equation 2).

A3 Heckman two stage estimation

Table 8: Second stage Heckman results (without STEP controls)

	Bolivia	Colombia	Ghana	Georgia	Ukraine	Vietnam
Returns gap	0.070***	0.081***	−0.001	0.200	0.083*	0.026
	(0.030)	(0.031)	(0.050)	(0.158)	(0.045)	(0.023)
Female IMR	1.757***	3.081***	−0.338	0.775	0.138	0.485
	(0.338)	(1.027)	(0.357)	(2.026)	(0.354)	(0.378)
Male IMR	0.676**	−0.851*	−0.505	0.377	−0.222	0.379
	(0.314)	(0.469)	(0.424)	(0.927)	(0.368)	(0.313)
R^2	0.400	0.344	0.219	0.245	0.287	0.278
N	642	669	661	201	726	1052
PSUs	165	387	194	115	256	225

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors in parentheses. Gap in returns defined as returns for females less returns for males = $\hat{\beta}_6$ for Heckman without STEP controls (see Equation 4). The first stage results for both specifications is reported in Table 9 and 10. See Table 5 and 6 for summary statistics (also on the excludable variables).

Table 9: First stage Heckman results, marginal effects (without STEP controls)

	Bolivia		Colombia		Ghana		Georgia		Ukraine		Vietnam	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Labour dep.	−0.006	−0.089	0.014	−0.167***	0.107***	0.108**	0.054*	0.010	0.026	0.094	0.025	0.037
	(0.066)	(0.068)	(0.058)	(0.060)	(0.036)	(0.048)	(0.030)	(0.062)	(0.045)	(0.058)	(0.046)	(0.058)
Asset index	−0.052***	−0.048*	−0.018	−0.002	0.016	0.021	−0.009	−0.041	0.016	0.043	−0.050***	−0.059***
	(0.016)	(0.024)	(0.020)	(0.024)	(0.012)	(0.018)	(0.014)	(0.028)	(0.018)	(0.027)	(0.016)	(0.016)
Econ. shocks	−0.002	0.029*	−0.003	−0.011	−0.028	0.021	0.000	−0.012	0.052**	0.003	−0.007	−0.018
	(0.008)	(0.017)	(0.015)	(0.018)	(0.018)	(0.026)	(0.010)	(0.015)	(0.025)	(0.040)	(0.013)	(0.020)
Soc. econ. status	−0.066**	−0.096**	−0.017	−0.025	0.015	−0.031	−0.022	−0.067**	0.009	−0.030	0.002	−0.023
	(0.030)	(0.043)	(0.028)	(0.043)	(0.018)	(0.030)	(0.014)	(0.032)	(0.034)	(0.041)	(0.020)	(0.028)
Years education	0.034***	0.028***	0.012**	0.006	0.027***	0.035***	0.023***	0.034***	0.054***	0.036***	0.029***	0.024***
	(0.004)	(0.006)	(0.006)	(0.007)	(0.004)	(0.005)	(0.003)	(0.006)	(0.006)	(0.011)	(0.004)	(0.004)
F-stat.	5.29	5.12	0.44	2.07	3.82	2.25	1.45	2.23	1.40	0.45	2.54	3.83
Prob.	0.001	0.001	0.7809	0.084	0.005	0.065	0.220	0.068	0.235	0.219	0.041	0.005

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The first four variables are excluded from the second stage regressions, the coefficient estimates for years of education are included for information purposes. F-statistics and probabilities are for a joint F-test of the coefficient estimates of the excludable variables only.

Table 10: First stage Heckman results, marginal effects (including STEP controls)

	Bolivia		Colombia		Ghana		Georgia		Ukraine		Vietnam	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Labour dep.	−0.052	−0.083	−0.042	−0.153**	0.100***	0.133***	0.045	−0.037	0.013	0.110*	0.019	0.034
	(0.074)	(0.078)	(0.054)	(0.062)	(0.036)	(0.050)	(0.057)	(0.081)	(0.048)	(0.062)	(0.047)	(0.061)
Asset index	−0.034	−0.025	0.001	0.026	0.015	0.024	−0.029	−0.046	0.011	0.049*	−0.048***	−0.036*
	(0.023)	(0.032)	(0.021)	(0.025)	(0.012)	(0.018)	(0.024)	(0.030)	(0.019)	(0.026)	(0.017)	(0.019)
Econ. shocks	0.002	0.028*	0.004	−0.021	−0.033*	0.023	−0.011	−0.019	0.070***	−0.016	−0.001	−0.028
	(0.009)	(0.017)	(0.015)	(0.018)	(0.019)	(0.027)	(0.016)	(0.023)	(0.024)	(0.042)	(0.014)	(0.022)
Soc. econ. status	−0.062**	−0.103**	−0.041	−0.041	0.005	−0.027	−0.049*	−0.016	0.012	−0.021	0.003	−0.034
	(0.031)	(0.043)	(0.028)	(0.040)	(0.019)	(0.031)	(0.027)	(0.047)	(0.032)	(0.039)	(0.021)	(0.031)
Years education	0.041***	0.030***	0.020***	0.010	0.020***	0.034***	0.052***	0.047***	0.054***	0.035***	0.031***	0.024***
	(0.005)	(0.007)	(0.007)	(0.009)	(0.004)	(0.006)	(0.007)	(0.007)	(0.008)	(0.011)	(0.004)	(0.005)
F-stat.	1.862	3.589	0.752	2.354	3.605	2.727	1.735	0.850	2.177	1.730	1.945	1.660
Prob.	0.120	0.008	0.557	0.053	0.007	0.031	0.145	0.495	0.072	0.144	0.104	0.160

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. The first four variables are excluded from the second stage regressions, the coefficient estimates for years of education are included for information purposes. F-statistics and probabilities are for a joint F-test of the coefficient estimates of the excludable variables only.