Essays in Economic Development:
Education, Child Labour and Wage Inequality

Salma Ahmed

M.S.S in Economics, *University of Dhaka, Bangladesh*
MPhil in Economics, *Monash University, Australia*

A Thesis Submitted in Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

Department of Economics
Faculty of Business and Economics
Monash University
Australia

July 2012

Table of Contents

List of Figures

List of Tables

Abstract

This thesis presents four self-contained essays that explore issues that are crucial in improving human well-being in a developing country: improving health, minimising child labour and reducing gender inequality. The analysis is focused on Bangladesh where the prevalence of child labour and gender differences in several domains is still widespread.

The first essay aims to examine the gender wage gap along the entire wage distribution into an endowment effect and a discrimination effect, taking into account possible selection into full-time employment. Applying a new decomposition approach to the Bangladesh Labour Force Survey (LFS) datasets of 1999 and 2005, we find that women are paid less than men everywhere on the wage distribution and the gap is higher at the lower end of the distribution. Discrimination against women is the primary determinant of the wage gap. We also find that this gap has widened between 1999 and 2005.

The second essay examines whether gender differences in tertiary enrolment rates can be explained by wage premiums in returns from secondary to tertiary education levels. Using LFS data, we find that wage premiums do not have any significant effect on the gender gap in tertiary enrolment rates. We also note that wage premiums in returns from secondary to tertiary education significantly influence tertiary enrolment rates for males but not for females, once additional variables are added. We offer evidence that part of the explanation for low female enrolment in tertiary education is attributable to demographic factors.

The third essay investigates whether there is any trade-off between child labour hours and schooling. By drawing on the 2002 dataset of the Bangladesh National Child Labour Survey (NCLS), we find that working hours adversely affect child schooling from the very first hour of work. However, the marginal impact of child labour hours weakens when working hours increase; yet, working hours always negatively affect schooling when we use a non-parametric approach. We find that parents do not have identical preferences towards schooling decisions concerning boys and girls. Both mother and father show a significant preference for educating a female child. The same incentive effect is not found for a male child. These conclusions persist, even after allowing for sample selection in child labour.

The fourth essay tests the effect of child labour on child health outcomes in Bangladesh. We use self-reported injury or illness due to work as a general measure of health status. Using NCLS data, we find that child labour is positively and significantly associated with the probability of being injured or becoming ill, once the endogenous relationship between these factors is accounted for. These findings remain robust when we consider child labour hours and restrict our analysis to rural areas. Moreover, the intensity of injury or illness is significantly higher in construction and manufacturing than in other sectors.

Declaration

I hereby declare that this thesis contains no material that has been accepted for the award of any other degree or diploma in any university, and that to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of this thesis.

The thesis contains one co-authored chapter (Chapter 2) of which the candidate was the primary researcher and author. This declaration is followed by a statement signed by the candidate and a co-author of this chapter.

[redacted]

SALMA AHMED

[Amendments at the back]

Thesis-Related Research Outcomes and Achievement

Refereed Conference Papers

Ahmed, S. and Maitra, P. (2011), 'A Distributional Analysis of the Gender Wage Gap,' Econometric Society Australasian Meeting, Adelaide, Australia, 4-7 July.

Ahmed, S. and Ray, R. (2011), 'Trade-off between Child Labour and Schooling in Bangladesh – The Role of Parents' Education,' Australian Conference of Economists, Canberra, Australia, 11-13 July.

Ahmed, S. (2011), 'Health Consequences of Child Labour in Bangladesh,' The 33rd Australian Conference of Health Economists, Melbourne, Australia, 29-30 September.

Achievement

'Best Student Paper Award' – third place for the paper titled 'Health Consequences of Child Labour in Bangladesh' in 24th PhD Conference in Economics and Business, Brisbane, Australia, 19-20 November 2011.

Acknowledgements

I wish to thank my academic advisors Prof. Ranjan Ray, Prof. Pushkar Maitra and Dr. Liang Choon Wang for their invaluable advice and supervision. In particular, I am grateful to Ranjan Ray and Pushkar Maitra for their support and encouragement during my years of research at Monash University. I am also grateful to Prof. Paul Miller, Prof. Alison Booth, and the participants of the Econometric Society Australasian Meeting, the Australian Conference of Economists, the Australian Conference of Health Economists and the Australasian Development Economics Workshop for providing very useful suggestions and comments.

I would like to express my appreciation to Professor Russell Smyth, Dr. Stephen Miller, Barbara Cramer and Felicity Milne for providing academic and technical support. I would like to sincerely thank the Monash Research Graduate School (MRGS) for funding this research. I am greatly indebted to the officials at the Bangladesh Bureau of Statistics (BBS) for providing the Labour Force Survey data.

Finally, I would like to express my deep appreciation to my family. I would especially like to thank my husband Mohammad Badrul Muttakin, my father Ali Ahmed, and my sisters, Amina Alam and Sohana Ahmed, for their encouragement and moral support.

The thesis is dedicated to my mother, Anarkali Ahmed, 1940-2005.

Chapter 1

Introduction

1.1 Introduction

This thesis presents four self-contained essays that focus on education, child labour and gender inequality. Despite some social and economic progress, there are still significant challenges for developing countries as they seek to improve human well-being in the twenty-first century. For example, gender differences are noticeable in several spheres. In addition, there is still a significant use of child labour in many parts of the world in spite of international and national efforts to tackle this phenomenon.[1] Careful research on the elimination of child labour and its effect as well as promotion towards gender equality within the framework of the Millennium Development Goals is essential for policymakers to ensure well-being in a society.

All these essays use datasets from Bangladesh. This country was not chosen to be representative of all developing countries; indeed no country could be. However, Bangladesh is an interesting country to study since child labour and gender inequality is known to exist there. The issue of child labour in Bangladesh became prominent in discussions in the early 1990s when the United States and several other countries refused to import items (for example, garments) from Bangladesh that had significant involvement of child labour. As a result, many children have been withdrawn from the garments industries, but child labour has not been significantly reduced. Child labour is still prevalent in rural areas, especially in the agricultural sector.

Women are at a significant disadvantage in Bangladesh. Gender inequality prevails in education as well as in a number of areas, such as economic participation and

[1] In this thesis, we use the terms ‘child labour’ and ‘child work’ interchangeably.

opportunity, survival chances and longevity.[2] Bangladesh, with a Gender Gap Index of 0.65, ranks 90th out of 130 countries (Hausmann et al. 2008). Women incur the direct costs of these inequalities, but eventually the whole society faces these costs as the nation's capacity, economic growth and living standards are reduced.

Although this thesis focuses on Bangladesh, much of the evidence presented here has implications that are relevant to policymakers in other developing countries.

1.2 The Structure of the Thesis

This thesis is structured as follows. Chapter 2 examines the gender wage gap among full-time employees in Bangladesh. We not only examine the gap at the mean, but also examine variation in the gender wage gap across the entire distribution of wages. By focusing on the distributions we are able to comment on, and observe, multiple features of the wage distribution and not just the mean. It also reports the effects of sample selection (selection into full-time employment for both males and females) on the gender wage gap, as full-time workers might not be a random subset of all workers. Furthermore, we investigate whether the gender wage gap varies over time. We apply a new decomposition approach to quantify the extent of the gender wage gap across the entire distribution of wages, while explicitly accounting for selectivity bias.

Chapter 3 investigates gender differences in the demand for tertiary education in Bangladesh. Females are well behind their male peers in rates of enrolment and completion of tertiary education. In Bangladesh, the gender parity index (the ratio of females to males) in tertiary education is approximately 0.53 (Hausmann et al. 2008). This is significantly lower than other countries at the same level of development, such as Pakistan. Using Bangladesh Labour Force Survey (LFS) datasets of 1999 and 2005, we seek to understand whether labour market returns explain gender differences in tertiary enrolment rates. Nevertheless, this does not necessarily imply that the other possible explanations of gender differentials in the tertiary enrolment rate in Bangladesh are trivial (for example, social and cultural norms, opportunity costs and demographic factors), although there is no evidence for this claim. In this chapter, we control for

[2]Although it can be argued that schooling and education are distinct, we speak of them as being synonymous in this thesis.

demographic factors and compare the effect of demographic factors with the effect of variables that measure returns to education.

Chapter 4 examines the trade-off between child labour and schooling in Bangladesh and also explores how parents' education affects the work-schooling trade-off between the two genders. It has been argued that educational attainment is likely to be lower for children who combine work and schooling. Much of this evidence is based on the rate of child participation rather than the hours worked. This is surprising given that child workers who spend longer hours on work activities will have little time for school attendance and study and this will hamper their educational progress. While the negative association between children's working hours and their schooling is well maintained in the literature, the mechanism through which this relationship works is not as well understood. Therefore, a causal relationship is hard to justify. The problem is largely methodological and linked to the difficulties in the estimation of child school production functions. This is because the number of hours that a child works is endogenously chosen. In order to address the potential problem of endogeneity, we adopt an instrumental variable (IV) estimation strategy. This is similar to the most recent literature on developing countries (Beegle et al. 2009; Boozer and Suri 2001; Goulart and Bedi 2008; Gunnarsson et al. 2006) which uses an IV strategy to identify the effect of child labour. We also use a non-parametric approach for further investigation of the relationship between child labour hours and schooling. This empirical analysis is conducted using the Bangladesh National Child Labour Survey data for 2002-2003.

Chapter 5 investigates the effect of child labour on child health outcomes. Although there has been much discussion about the consequences of children's labour on their schooling, less is known about the causal relationship between child labour and health. The only related studies we are aware of that quantify the health impacts of child labour are by Beegle et al. (2009) and O'Donnell et al. (2005) in Vietnam, Kana et al. (2010) in Cambodia and Wolff and Maliki (2008) in Indonesia. The absence of research for South Asian countries is significant given the fact that the majority of Asian child workers come from South Asia (Ray 2004).

We use self-reported injury or illness as a measure of health status and employ the bivariate probit approach to quantify the causal effect of child labour on child health, considering the endogeneity problem of child labour. In addition, we adopt the semi-parametric approach (partially linear model) to understand the association between working hours and subjective child health, treating working hours as endogenous. The choice of the semi-parametric estimator is motivated by the fact that it allows for a more flexible relationship between hours worked and health outcomes. We implement these methods by exploring the ILO-sponsored National Child Labour Survey conducted in Bangladesh in 2002-2003, which was especially designed to provide data on the relationship between children’s work and health outcomes.

Finally, Chapter 6 concludes this thesis. It summarises the findings and explores their implications for future research.

Chapter 2

A Distributional Analysis of the Gender Wage Gap

with Pushkar Maitra

2.1 Introduction

There is now an extensive literature that analyses the extent of gender gap in wages. The specific aim of this literature is to try to understand how much of the gap is due to differences in productive characteristics (*the endowment effect*) or how much of it is due to discrimination, which is the gender gap in wages that persists even after the differences in endowments have been controlled for. This is termed *the discrimination effect*.[3] This is an important question as different implications and policy prescriptions need to be drawn depending on the source of the wage gap.

The most common method of decomposing the gender gap in wages has been to use the Oaxaca-Blinder decomposition method (Blinder 1973; Oaxaca 1973), which typically conducts the decomposition analysis at the mean of the wage distribution. However, looking at the effect at the mean might not provide the full story. Recent evidence using data from both developed and transitional economies suggests that the average wage gap and decomposition at the mean is not representative of the gaps, and is not representative of the factors that explain these gaps at different points of wage distributions for the population of interest. See, for example, Albrecht et al. (2003), Machado and Mata (2005), Miller (2005), Gupta et al. (2006), Arulampalam et al. (2007), Chzhen and Mumford (2011). The general consensus of this literature is that the gender wage gap varies over the distribution of wages. The literature also points to the

[3]In a perfectly competitive market (assuming homogeneity and perfect substitutability of the labour force), discrimination originates from employer prejudice (Becker 1971). This suggests that even in the presence of equal endowments of productive skills, wage inequality will persist if employers reward productive skills differently depending on the gender of the worker.

importance of gender differences in the propensity to participate in the labour market. Additionally, it also implies that in order to obtain unbiased estimates of the gender wage gap, we need to explicitly account for self-selection into employment. For example, if the women who stay out of employment are those who would have received the lowest returns from work, then ignoring the selection issue would result in a significant bias in the estimated gender wage gap across the wage distribution.[4]

There is also a sizeable literature using data from developing countries on decomposing the gender wage gap at different points of the wage distribution (see, for example, Ganguli and Terrell 2005; Millimet and Wang 2006; Nopo 2006; Pham and Reilly 2007; Popli 2012; Sakellariou 2004). With the exception of Popli (2012), none of these studies take into account non-random selection into employment. This is particularly relevant for developing countries, where the employment rate, the type of employment, and the choice of industry and occupation vary systematically by gender. One of the principal aims of this essay is to address this shortcoming by examining the extent of the gender wage gap among employees who choose to work full-time (i.e. to account for potential sample selection) and also decompose this gap at different points of the distribution.[5]

In this essay, we use two nationally representative unit record datasets (surveys conducted in 1999-2000 and 2005-2006) from Bangladesh to examine the following questions:

a. Does the gender wage gap vary over the entire wage distribution?
b. How much of the wage gap can be attributed to discrimination?
c. Did the gender wage gap change over time?
d. How are the results affected if we explicitly take selection into full-time employment into account?

[4] Indeed, Albrecht et al. (2009), using data from Netherlands, and Picchio and Mussida (2010), using data from Italy, find that after adjusting for sample selection and for gender differences in the distribution of characteristics, the average log-wage gap between male and female workers widens across the entire wage distribution.

[5] It is unlikely that the sample of full-time workers represents a random draw from the population as a whole. It is assumed that only individuals with wages exceeding reservation wages will enter the labour market, and these individuals may have attributes (for example, relative productivity in labour market and home activities, identity and stage in the life-cycle, and the attitudes and aspirations towards full-time work) that distinguish from other individuals (working part-time, self-employed or not employed). If such factors are observable, then they can be included in the regression model and this will allow us to correct for potential bias. However, the possibility that unobservable factors influence selection into full-time employment remains an obstacle.

We start by conducting the standard Oaxaca-Blinder decomposition at the mean. This provides a useful benchmark against which the extent of the gender wage gap at other points of distribution can be compared. The analytical framework that we adopt to compute and decompose the gender wage gap along the wage distribution is based on newly developed unconditional quantile regression models (Firpo et al. 2009). The advantage of the unconditional quantile regression over the traditional conditional quantile regression of Koenker and Bassett (1978) is that its estimated coefficients are explained as the impact of changes in the distribution of explanatory variables on the targeted quantiles of the unconditional marginal distribution of the dependent variable. Therefore, we can apply the Oaxaca-Blinder decomposition method directly to the estimation results from the unconditional quantile regression. More details of the unconditional quantile regression method that we use in this essay are provided in subsequent sections.

To investigate whether the gender wage gap varies over time, we conduct a decomposition analysis of changes in the gender wage gap along the wage distribution between two points in time (1999 and 2005). Several studies have shown that the factors that explain a gap at one point in time do not necessarily explain changes in this gap over a longer period of time and factors that are relevant at the lower end of the wage distribution may not be relevant at the upper end (see Kassenbohmer and Sinning 2010). We extend the procedure proposed by Wellington (1993) who decomposes changes in the gender wage gap at the mean to decompose changes in the gender wage gap over the entire distribution of wages.

To the best of our knowledge, there are two papers that have used unconditional quantile regressions to examine (and decompose) gender differences in wages across the entire distribution. Wei and Bo (2007), using data from urban China, find that the gender wage gap has increased from 1987 to 2004 across the wage distribution, and this increase has been greater at the lower end of the wage distribution. Kassenbohmer and Sinning (2010) examine changes in wage differentials between white men and women in the United States from 1994 to 2007 and find that the gender wage gap narrowed by more than 20 percent at the lowest decile and by less than 4 percent at the highest decile of the wage distribution. Our essay differs from these two earlier studies in a number of different ways. First, we perform decomposition across the entire distribution of wages, while explicitly accounting for selection into full-time employment. Second, we decompose changes in the gender wage gap from 1999 to 2005 to assess the changes in the contribution of individual covariates between the two time periods, again explicitly accounting for selectivity bias. We use the Heckman (1979) two-step approach to take account of the selection issue into full-time employment and extend the Oaxaca-Blinder decomposition to the unconditional quantile regression framework. To the best of our knowledge this is the first attempt at using the unconditional quantile regression model where the issue of selection bias is explicitly addressed. Third, our method allows us to go beyond previous analyses of the gender wage gap in Bangladesh that only consider average wage differentials, neglecting the remainder of the distribution.

We find that the extent of the gender wage gap varies significantly across the wage distribution after adjusting for gender differences in the distribution of characteristics, indicating that the mean gender wage gap disguises the variation across the wage distribution. These differences are not uniform across the wage distribution; the disparity is largest in the lowest quantile (reaching 65 percent in 1999 and 108 percent in 2005 using real hourly wages) and declines (though not monotonically) as we go up the wage distribution. Differences in characteristics (*the endowment effect*) are not uniform at all quantiles and are mostly in favour of males. Discrimination explains the major proportion of the wage gap at all quantiles. The gender wage gap, however, increased from 1999 to 2005 by about 26 percent in the lowest quantile and by about 20 percent in the highest quantile. Finally, sample selection into full-time employment has a significant impact on the gender wage gap and the results suggest that not controlling for sample selection is likely to over-estimate the observed wage gap.

2.2 Empirical Methodology

2.2.1 Oaxaca-Blinder Decomposition

As a first attempt to formally identify the underlying causes of the gender wage gap, we perform the Oaxaca-Blinder decomposition at the mean. Specifically, we start by estimating separate (log) hourly wage equations as follows:[6]

---

[6]Wage equations are estimated separately for men and women in order to allow for different rewards by gender to a set of productive characteristics or endowments. A Chow test (*F*-test) rejects the null hypothesis that explanatory variables have equal impacts on the wage rates of males and females for both years. The Chow test statistics for the 1999 survey year is $F(35, 5451) = 12.21$ (with a p-value of 0.000), and $F(35,18322) = 37.12$ (with a p-value of 0.000) for the 2005 survey year.

$$lnw_{ijt} = X'_{ijt}\beta_{jt} + \varepsilon_{ijt}; \qquad i = 1, \dots, n;\ j = m, f;\ t = 1999{,}2005 \qquad (2.1)$$

where i denotes the individual; j the gender group (male or female) and t the survey year (1999 or 2005); lnw_{ijt} is the (log) of hourly wages; X_{ijt} is the vector of explanatory variables (a set of individual characteristics) that affect the wages received and $\varepsilon_{ijt} \sim N(0, \sigma^2)$. Equation (2.1) is estimated using ordinary least squares (OLS).

We define D_t as the difference in the expected value of male and female wages in period *t* (raw difference) obtained by estimating Equation (2.1) separately for males and females. D_t can be decomposed into the component of the raw difference attributable to differences in observed characteristics or endowments (E) and to differences in coefficients (C). We can then write:

$$\begin{aligned} D_t &= \overline{lnw}_{mt} - \overline{lnw}_{ft} \\ &= [\bar{X}'_{mt}\hat{\beta}_{mt} - \bar{X}'_{ft}\hat{\beta}_{ft}] \\ &= (\bar{X}_{mt} - \bar{X}_{ft})'\hat{\beta}_{mt} + \bar{X}'_{ft}(\hat{\beta}_{mt} - \hat{\beta}_{ft}) \qquad (2.2) \\ &= E + C \end{aligned}$$

Here $\hat{\beta}_{jt}$ is the estimated value of β_{jt}. The first term in the right hand side of Equation (2.2), $(\bar{X}_{mt} - \bar{X}_{ft})'\hat{\beta}_{mt}$, is the explained component of the wage gap, which is the component of the gap that can be explained by differences in observed characteristics at the mean, weighted by coefficients attributable to men $(\hat{\beta}_{mt})$. This is E.The second term, $\bar{X}'_{ft}(\hat{\beta}_{mt} - \hat{\beta}_{ft})$, is the unexplained component or C. This is the wage difference that is due to differential reward for equal characteristics and is interpreted as a measure of the extent of discrimination in the labour market. It estimates the residual or the component of the wage difference that cannot be attributed to differences in observable characteristics and thus reflects unequal pay for the same characteristics (Oaxaca 1973).

However, we need to be careful when interpreting the model residual as discrimination. If there is any omitted variable that has a positive effect on wages, and if men are more highly endowed with this variable, then the results from the decomposition would over-estimate discrimination. Alternatively, if some of the factors in the model are themselves affected by discrimination, then the analysis could well under-estimate discrimination. For example, if women have less access to the types of schooling that are deemed more valuable by the market, then the decomposition may well under-estimate discrimination.

An alternative way of writing Equation (2.2) is to use the female wage structure as the reference category. In this case, the explained component can be written as $(\bar{X}_{mt} - \bar{X}_{ft})'\hat{\beta}_{ft}$ and the unexplained component can be written as $[\bar{X}'_{mt}(\hat{\beta}_{mt} - \hat{\beta}_{ft})]$. However in our analysis we present and discuss the results corresponding to the case where the male wage rate is the reference category.[7]

2.2.2 Distributional Decomposition Using Unconditional Quantile Regression

This section expands our analysis by examining the gender wage gap along the whole distribution of wages using the Oaxaca-Blinder decomposition approach based on unconditional quantile regression estimates (Firpo et al. 2009). They show that the Oaxaca-Blinder decomposition can be approximated for any distributional statistic, including quantiles. This method comprises of two stages. In the first stage, distributional changes are divided into a wage structure effect and a composition effect using a re-weighting method. The re-weighting method allows us to directly estimate these two components without having to estimate a structural wage determination model. In the second stage, both components are further divided into the contribution of each explanatory variable using re-centred influence function (RIF) regressions. These regressions directly estimate the impact of the explanatory variables on the distributional statistic of interest thereby generalising the Oaxaca-Blinder decomposition method by extending the decomposition to any distributional measure. See Firpo et al. (2009) for

[7] In doing so, we have abstracted from an important debate: which wage structure should we use as the reference category? The Oaxaca-Blinder method applied both male and female wage structures as the reference category. This creates an index number problem, since the estimates of the discrimination component differs depending on the choice of the reference category. Furthermore, the resulting levels of discrimination provide a range within which the actual level of discrimination falls. Reimers (1983) hypothesises that the correct procedure is instead to take an average of both male and female wage structures. Cotton (1988) suggests improving upon the procedure by employing a weighted average of the two wage structures, which should then provide us with an exact figure rather than a range. Neumark (1988), on the contrary, regards these benchmarks as unsatisfactory and argues that the choice of a non-discriminatory wage structure should be based on OLS estimates from a pooled regression of both males and females. However, Ginther and Hayes (2003) point out that a pooled wage structure (i.e. an average of the male and female wage structures) is not likely to be used in a legal framework concerned with equal opportunities for women and men. Rather the authors argue that men are the usual comparison group in legal proceedings concerning gender discrimination.

more details. Specifically, the predicted wage differential $D_t(\nu)$ measured at quantile ν can be decomposed as follows:

$$\begin{aligned} D_t(\nu) &= lnw_{mt}(\nu) - lnw_{ft}(\nu) \\ &= [\bar{X}'_{mt}\hat{\beta}_{mt}(\nu) - \bar{X}'_{ft}\hat{\beta}_{ft}(\nu)] \\ &= \left(\bar{X}_{mt} - \bar{X}_{ft}\right)'\hat{\beta}_{mt}(\nu) + \bar{X}'_{ft}\left[\hat{\beta}_{mt}(\nu) - \hat{\beta}_{ft}(\nu)\right] \\ &= E(\nu) + C(\nu) \end{aligned} \tag{2.3}$$

Here $\hat{\beta}_{jt}(\nu)$ is the parameter estimates of the RIF regression model, $\bar{X}_{jt}$is a vector of average characteristics of workers.[8] In our analysis, we apply this framework to the following quantiles $\nu = 0.10, 0.25, 0.50, 0.75, 0.90$.

Two observations about the gender wage gap are worth mentioning. First, even though the gender gap can theoretically be negative, empirically it is always positive both at the mean and quantiles. This implies that men are paid more. Second, when women are more productive than men, but are still discriminated against, discrimination is greater than the gender wage gap. This is something that occasionally happens with our sample.

2.2.3 Decomposition of the Inter-Temporal Change in the Gender Wage Gap

We use the Wellington (1993) method to extend the single period Oaxaca-Blinder approach to analyse changes in the wage gap between 1999 and 2005. We want to examine how the changes in the characteristics and the returns to these characteristics combine to affect the gender wage gap over the relevant period. To do this, we start by subtracting the difference in (log) wages in period τ from the corresponding difference in period t. Specifically, we can write the change in the mean gender wage gap over time as follows:

$$\begin{aligned} D_t - D_\tau &= [\bar{X}'_{mt}\hat{\beta}_{mt} - \bar{X}'_{ft}\hat{\beta}_{ft}] - [\bar{X}'_{m\tau}\hat{\beta}_{m\tau} - \bar{X}'_{f\tau}\hat{\beta}_{f\tau}] \\ &= \left[(\bar{X}_{mt} - \bar{X}_{m\tau})'\hat{\beta}_{mt} - \left(\bar{X}_{ft} - \bar{X}_{f\tau}\right)'\hat{\beta}_{ft}\right] \\ &\quad + \left[\bar{X}'_{m\tau}\left(\hat{\beta}_{mt} - \hat{\beta}_{m\tau}\right) - \bar{X}'_{f\tau}\left(\hat{\beta}_{ft} - \hat{\beta}_{f\tau}\right)\right] \end{aligned} \tag{2.4}$$

[8]We follow García et al. (2001) and Mueller (1998) and use average characteristics to decompose the wage differentials at different quantiles.

where $D_t = \overline{lnw}_{mt} - \overline{lnw}_{ft}$. The first term of the decomposition $\left[(\bar{X}_{mt} - \bar{X}_{m\tau})'\hat{\beta}_{mt} - \left(\bar{X}_{ft} - \bar{X}_{f\tau}\right)'\hat{\beta}_{ft}\right]$ shows the change in the wage gap due to changes in the mean of the regressions (the explained portion) evaluated at the period t coefficients. The second term $\left[\bar{X}'_{m\tau}\left(\hat{\beta}_{mt} - \hat{\beta}_{m\tau}\right) - \bar{X}'_{f\tau}\left(\hat{\beta}_{ft} - \hat{\beta}_{f\tau}\right)\right]$ represents the portion of the change in the wage gap that can be explained by changes in the coefficients between the two periods, evaluated at the corresponding group's mean in period τ.

We can extend Equation (2.4) to decompose the wage difference at different quantiles over time as:

$$\begin{aligned} D_t(v) - D_\tau(v) &= [\bar{X}'_{mt}\hat{\beta}_{mt}(v) - \bar{X}'_{ft}\hat{\beta}_{ft}(v)] - [\bar{X}'_{m\tau}\hat{\beta}_{m\tau}(v) - \bar{X}'_{f\tau}\hat{\beta}_{f\tau}(v)] \\ &= \left[(\bar{X}_{mt} - \bar{X}_{m\tau})'\hat{\beta}_{mt}(v) - \left(\bar{X}_{ft} - \bar{X}_{f\tau}\right)'\hat{\beta}_{ft}(v)\right] \\ &+ \left[\bar{X}'_{m\tau}\left(\hat{\beta}_{mt}(v) - \hat{\beta}_{m\tau}(v)\right) - \bar{X}'_{f\tau}\left(\hat{\beta}_{ft}(v) - \hat{\beta}_{f\tau}(v)\right)\right] \end{aligned} \tag{2.5}$$

Changes in any of these above components over time would cause changes in the gender wage gap. In terms of mean characteristics, the explanations centre on changes in male-female productivity-related characteristics. For example, if women's work experience over time becomes similar to that of men's, then the male-female wage gap is likely to be reduced. On the other hand, there might be a number of different reasons as to why differences in the coefficients might change over time. For example, if there are changes in the returns to the explanatory variables, such as a change in the relative magnitudes of the coefficients that favour women, the gender wage gap will be reduced. However, this second component might also represent labour market discrimination that cannot be associated with any particular observed characteristics in the regression, thus making the interpretation of the unexplained component more difficult.

2.2.4 Selection into Employment?

The focus of our analysis is full-time workers. However, full-time workers might not be a random subset of all workers and indeed they might differ in terms of both observables and unobservables to those not employed full-time in the labour market (for example, those who are self-employed or who are employed in family businesses). One way to correct this selection bias is to employ the standard Heckman two-step estimation technique. We first estimate the inverse Mill's ratio (λ) from a probit equation that determines full-time participation in the labour market (choosing to become a full-time wage employee). To do this we estimate the following equation:

$$I_{ijt} = Z'_{ijt}\gamma_{jt} + \mu_{ijt}; i = 1, \dots, n;\ j = m, f;\ t = 1999{,}2005 \qquad (2.6)$$

where I_{ijt} is a dummy variable denoting full-time employment status ($I = 1$ if the individual is in full-time employment and 0 otherwise) and μ_{ijt}~IIDN(0,1). Estimation of Equation (2.6) allows us to compute the inverse Mill's ratio (λ), which is then added as an additional regressor in Equation (2.1), both in the OLS and at different quantiles. We include ownership of dwelling (home ownership), wealth quintile of the household, the number of young children in the household and the number of men and women in the household over 65 years of age to achieve identification of the selection term, the inverse Mill's ratio (see Table 2.2). These variables are assumed to affect the probability of full-time employment (for example, the presence of young children in the household may inhibit women's probability of full-time employment in the labour market, or provide additional impetus to male workers) but not to affect wages. Indeed, there is very little reason to expect that these variables will have an effect on the wage rate, which is market determined and is typically beyond the control of any individual.

We can now compute the extended gender wage gap (at the mean) as:

$$\begin{aligned} D_t &= \overline{lnw}_{mt} - \overline{lnw}_{ft} \\ &= \left(\bar{X}_{mt} - \bar{X}_{ft}\right)'\hat{\beta}_{mt} + \bar{X}'_{ft}\left(\hat{\beta}_{mt} - \hat{\beta}_{ft}\right) + \left(\hat{\theta}_{mt}\bar{\lambda}_{mt} - \hat{\theta}_{ft}\bar{\lambda}_{ft}\right) \qquad (2.7) \end{aligned}$$

where $\left(\hat{\theta}_{mt}\bar{\lambda}_{mt} - \hat{\theta}_{ft}\bar{\lambda}_{ft}\right)$ is the contribution of differences in the average selectivity bias.[9] This part may be viewed as the differences in unobservables, which influence wages. Selectivity bias results in the observed wage differential being different from the offered wage differential. If we rewrite Equation (2.7) as:

$$\begin{aligned} \widetilde{D}_t &= \left(\overline{lnw}_{mt} - \overline{lnw}_{ft}\right) + \left(\hat{\theta}_{ft}\bar{\lambda}_{ft} - \hat{\theta}_{mt}\bar{\lambda}_{mt}\right) \\ &= \left(\bar{X}_{mt} - \bar{X}_{ft}\right)'\hat{\beta}_{mt} + \bar{X}'_{ft}\left(\hat{\beta}_{mt} - \hat{\beta}_{ft}\right) \qquad (2.8) \end{aligned}$$

[9] θ_{jt} is the estimated coefficient of λ_{jt} from the extended wage regression where λ_{jt} is included as an additional explanatory variable.

Following Duncan and Leigh (1980) and Reimers (1983), the left hand side of Equation (2.8) can be interpreted as a measure of differences in the offered wage (the sum of the difference in the observed mean wages and the difference in average selectivity bias). The only difference between Equations (2.7) and (2.8) is that Equation (2.8) presents a decomposition of the selectivity-adjusted wage difference (difference in offered wages) as opposed to a decomposition of the observed wage difference, as in Equation (2.7). Equation (2.8) can also be estimated at different quantiles.

The decomposition of the change in the gender wage gap (at the mean) over time, taking into account selection into full-time employment, can be expressed as:

$$\begin{aligned}D_t - D_\tau &= [\bar{X}'_{mt}\hat{\beta}_{mt} - \bar{X}'_{ft}\hat{\beta}_{ft}] - [\bar{X}'_{m\tau}\hat{\beta}_{m\tau} - \bar{X}'_{f\tau}\hat{\beta}_{f\tau}] \\ &= \left[(\bar{X}_{mt} - \bar{X}_{m\tau})'\hat{\beta}_{mt} - \left(\bar{X}_{ft} - \bar{X}_{f\tau}\right)'\hat{\beta}_{ft}\right] \\ &+ \left[\bar{X}'_{m\tau}\left(\hat{\beta}_{mt} - \hat{\beta}_{m\tau}\right) - \bar{X}'_{f\tau}\left(\hat{\beta}_{ft} - \hat{\beta}_{f\tau}\right)\right] \\ &+ \left[\left(\bar{\lambda}_{mt} - \bar{\lambda}_{m\tau}\right)'\hat{\theta}_{mt} - \left(\bar{\lambda}_{ft} - \bar{\lambda}_{f\tau}\right)'\hat{\theta}_{ft}\right] \\ &+ \left[\bar{\lambda}'_{m\tau}\left(\hat{\theta}_{mt} - \hat{\theta}_{m\tau}\right) - \bar{\lambda}'_{f\tau}\left(\hat{\theta}_{ft} - \hat{\theta}_{f\tau}\right)\right]\end{aligned} \tag{2.9}$$

Taking into account sample selection, we can decompose the gender wage gap at different points of the wage distribution as follows:

$$\begin{aligned}\widetilde{D}_t(\nu) &= \left(lnw_{mt}(\nu) - lnw_{ft}(\nu)\right) + \left(\hat{\theta}_{ft}\bar{\lambda}_{ft}(\nu) - \hat{\theta}_{mt}\bar{\lambda}_{mt}(\nu)\right) \\ &= \left(\bar{X}_{mt} - \bar{X}_{ft}\right)'\hat{\beta}_{mt}(\nu) + \left[\bar{X}'_{ft}\left(\hat{\beta}_{mt}(\nu) - \hat{\beta}_{ft}(\nu)\right)\right]\end{aligned} \tag{2.10}$$

Finally, we can extend Equation (2.9) to decompose the wage differences between men and women over time at different quantiles as:

$$\begin{aligned}D_t(\nu) - D_\tau(\nu) &= [\bar{X}'_{mt}\hat{\beta}_{mt}(\nu) - \bar{X}'_{ft}\hat{\beta}_{ft}(\nu)] - [\bar{X}'_{m\tau}\hat{\beta}_{m\tau}(\nu) - \bar{X}'_{f\tau}\hat{\beta}_{f\tau}(\nu)] \\ &= \left[(\bar{X}_{mt} - \bar{X}_{m\tau})'\hat{\beta}_{mt}(\nu) - \left(\bar{X}_{ft} - \bar{X}_{f\tau}\right)'\hat{\beta}_{ft}(\nu)\right] \\ &+ \left[\bar{X}'_{m\tau}\left(\hat{\beta}_{mt}(\nu) - \hat{\beta}_{m\tau}(\nu)\right) - \bar{X}'_{f\tau}\left(\hat{\beta}_{ft}(\nu) - \hat{\beta}_{f\tau}(\nu)\right)\right] \\ &+ \left[\left(\bar{\lambda}_{mt} - \bar{\lambda}_{m\tau}\right)'\hat{\theta}_{mt}(\nu) - \left(\bar{\lambda}_{ft} - \bar{\lambda}_{f\tau}\right)'\hat{\theta}_{ft}(\nu)\right] \\ &+ \left[\bar{\lambda}'_{m\tau}\left(\hat{\theta}_{mt}(\nu) - \hat{\theta}_{m\tau}(\nu)\right) - \bar{\lambda}'_{f\tau}\left(\hat{\theta}_{ft}(\nu) - \hat{\theta}_{f\tau}(\nu)\right)\right]\end{aligned} \tag{2.11}$$

2.3 Background, Data, and Descriptive Statistics

During the 1990s Bangladesh embarked on an ambitious program of economic reforms, including political democratisation, macroeconomic stabilisation and trade liberalisation. During the next 15 years Bangladesh experienced an increase in the GDP growth rate in real terms, with the growth rate increasing from 4 percent per annum in 1991 to 6 percent per annum in 1999 and then to 6.6 percent per annum in 2005. It is hardly a coincidence that a switch to a higher growth regime in the second half of the 1990s happened at the same time as the implementation of economic reforms. Though there might be disagreements as to the extent to which this economic growth contributed to higher standard of living of the poor throughout this period, poverty rates declined from 50 percent in 1999 to 40 percent in 2005 (Sasin 2007). It has been argued that much of this poverty reduction was driven by an increase in wages and employment opportunities, particularly in non-agricultural sectors.

From a gender viewpoint, women have made important advances in the labour market during this period. Although still far behind men, the participation rate of women's labour force has increased from 24 percent in 1999 to 29 percent in 2005. This has been associated with an increased share of women in the urban labour force, particularly in the manufacturing sector (often in the ready-made garments industries). However, gender inequalities continue to persist in the labour market in Bangladesh. Women filled only 20 percent of the 5.6 million new jobs generated between 1999 and 2005. Despite a strong convergence in the distribution of characteristics (for example, in terms of educational attainment) from 1999 to 2005, wages of men and women have not converged to the same extent and a sizeable gender gap continues to persist. Ahmed (2007), for example, finds that women in Bangladesh earn 66 percent less per hour than men and a significant proportion of this wage gap is attributable to discrimination. The few other studies that now exist on the male-female wage gap in Bangladesh have yielded similar results(for example of early studies, see, Ahmed and Maitra 2010; Al-Samarrai 2007; Kapsos 2008). The situation is much worse in the rural sector because of tradition, culture and religious norms that tend to restrict women's participation in paid employment and their freedom to choose a particular occupation.[10] This often makes

[10]Women in rural areas are often employed in low productivity jobs which are usually concentrated in public food for work programs and in unpaid family businesses.

women in Bangladesh less likely to join the labour market and, compared to men, they are more likely to commit to family than to career opportunities.

Maternity leave provisions are mandated in the public sector but are rare in the private sector and formal childcare remains relatively uncommon in Bangladesh. Finally the rate of unionisation is quite low in Bangladesh (only 3 percent of all workers are unionised) and unions are also less likely to negotiate with employers over pay and work conditions. While labour laws stipulate that no one can be discriminated against based on gender, enforcement of this legislation has been weak so far. Moreover, there is no national minimum wage in Bangladesh.

The empirical analysis in this essay uses two nationally representative datasets. Specifically, we use data from two Labour Force Surveys conducted in 1999-2000 (henceforth LFS 1999) and 2005-2006 (henceforth LFS 2005). These two cross-sectional surveys were administered by the Bangladesh Bureau of Statistics (BBS). The questionnaire for these two surveys is almost identical, and therefore overall inter-temporal compatibility is very good. The data contains information on a range of individual (age, gender, marital status, educational attainment, employment status, hours worked, wages earned) and household-level characteristics (household size and composition, religion, land holding, location, asset ownership). However, the key difference between the two datasets is the sample size.

The estimating sample in the LFS 1999 dataset consists of 12,652 individuals from 9,790 households, while the LFS 2005 dataset consists of 57,074 individuals from 40,000 households. The main reason for this large difference in sample size is the extent of coverage. The LFS 1999 consists of 442 Primary Sample Units (PSUs), of which 252 are rural and the rest are urban. The LFS 2005 consists of 1,000 PSUs, of which 640 are in rural areas with the rest in urban areas. From each PSU, 40 households were randomly selected for a detailed interview in the LFS 2005, while only 20 households from each rural PSU and 25 from each urban PSU were randomly selected for the same in LFS 1999. The difference is largely due to the sampling frame used in these two surveys. The LFS 1999 is based on the Population Census 1991, while the LFS 2005 is based on the Population Census 2001. However, prior to the LFS 1999, BBS conducted a complete new listing of households in the sample area.

Our decomposition analysis is restricted to individuals aged 15-65 who are in full-time wage employment (specifically defined as individuals who work for 40 hours or more during the week). The official retirement age in Bangladesh is 60 for males and 55 for females. However, these retirement ages are enforced only in the public sector and a large proportion of men and women continue to work beyond the age of 60. The selected sample of full-time workers consists of 5,522 individuals (84 percent males) in the LFS 1999 dataset and 18,392 individuals (88 percent males) in the LFS 2005 dataset.

In the wage regression the dependent variable is the (log) of hourly wages. Hourly wages are computed by dividing monthly wages by the total hours of work per month. The survey collected information on the usual hours work per week but not the number of weeks worked during a month. Therefore, the monthly hours of work are computed by multiplying the usual hours of work per week by 52/12. All nominal wages are converted to real values using the national consumer price index, 1999 = 100.

A large number of explanatory variables that are expected to affect wages are included in the wage regressions. We include dummy variables for the highest educational attainment of the individual,[11]occupation categories[12] and industry of work,[13]different age groups, marital status, training[14] and region of residence (rural residence is the reference category).[15]

In Figure 2.1, we show the distribution of wages by gender for the two survey years. The mass of the distribution of wages for males is to the right of that for females. Figure 2.2 shows that the distribution of wages has shifted to the right for both males

---

[11] We use four dummy variables to capture the highest level of educational attainment: primary (Grade 1-5), secondary (Grade 6-10), post-secondary (Secondary School Certificate/Higher Secondary School Certificate or equivalent), graduate (graduate or higher degree). The reference category is no education.

[12]We include six occupational category dummies: professional, administrative, clerical, sales, agriculture, and production. The reference category is service.

[13] We include nine industry indicators: agriculture, manufacturing, health, public administration, transport, financial institutional, real estate, wholesale and retail, and education. The reference category is hospitality.

[14] Here training does not refer to any job-specific training. Also, given the nature of the training question in LFS, it was not possible to separate out training acquired before entering one's current position. This specification does give us any insight into the importance of training.

[15] One could argue that occupation and industry variables should not be included as explanatory variables in the wage regression because of the possibility that industry and occupation are endogenous. That is, it is possible that individuals might choose their jobs and industries based on earnings prospects. An additional reason for omitting these variables is that employers' discriminatory practices could be highly correlated with industry and occupation. On the other hand, it is believed that these occupational and industry controls might embody unmeasured occupation-specific and industry-specific human capital (Arulampalam et al. 2007). Therefore, we might overlook the potential effect of unobserved human capital if we exclude such controls from the analysis. Arulampalam et al. (2007) argue that estimates with these controls can be viewed as a lower bound of the extent of discrimination.

and females in 2005, relative to 1999. A more detailed picture of this evolution of wage rates of males and females and gender wage gaps from 1999 to 2005 could be seen from Table 2.1, which presents the (log) real hourly wages and the gender wage gap at different quantiles and at the mean for the two datasets. The estimated (log) real hourly wages for both males and females increased from 1999 to 2005. The increase in wage rates is greater for men than for women. This is true at the mean as well as at different quantiles. In addition, the gender wage gap has increased over the relevant period almost everywhere on the distribution. The increase in the gender wage gap has been higher at the lower end of the distribution, increasing from 0.5026 log points in 1999 to 0.7303 log points in 2005 at the 10th quantile ($\nu = 0.10$) compared to the upper end of the distribution, where it has increased from 0.2261 log points in 1999 to 0.4049 log points in 2005 at the 90th quantile ($\nu = 0.90$).

In addition to the differences in the (log) real hourly wages between males and females discussed above, there are substantial differences in the means of the observed characteristics. Gender-specific descriptive statistics for the full sample (comprising of full-time workers, self-employed and employed in family businesses) are presented in Table 2.2. Table 2.3 displays descriptive statistics for the sample of full-time wage employees. We also present *t*-tests for gender differences.

Table 2.2 shows that females are on an average younger and are generally less educated compared to males. Gaps in educational attainment between males and females are statistically significant at all levels of education from 1999 to 2005. A higher proportion of males are married in 1999 when compared to females and interestingly this pattern is reversed in 2005. More than 40 percent of men and women are likely to be in full-time wage employment in 1999, although the gender difference is not statistically significant. However, full-time wage employment has decreased for both males and females from 1999 to 2005. For men this decline is 6 percentage points (down from 43 percent in 1999 to 37 percent in 2005), while for women this decline is 28 percentage points (down from 45 percent in 1999 to 17 percent in 2005). These results perhaps suggest that female workers are more likely to be dependent on self-employment activities or employed in family businesses. The explanation in part is due to downsizing of the state-owned enterprises (SOEs) as part of economic reforms in Bangladesh in the 1990s, which is more likely to affect women than men because downsizing typically implies a reduction in low skilled and low-paid public sector jobs, many of which are held by women. Once retrenched, women may face larger obstacles in finding comparable formal-sector jobs, forcing them to turn to low-paying jobs or become self-employed.

Restricting ourselves to the sample of full-time wage employees (Table 2.3), we again find that women are in general younger, and more likely to be in full-time employment, if they reside in the urban region. The gender difference is statistically significant in each of the two survey years. Moreover, females are generally less educated except at the post-secondary and graduate levels in 2005 and the gender differences are statistically significant at the 1 percent level. Furthermore, gender differences in occupation level exist and take on different patterns over the years. Notably, both male and female employment in administrative occupations decreases from 1999 to 2005. However, the decrease is more abrupt for females than for males, from 13.8 percent in 1999 to 0.02 percent in 2005 for females, and from 7 percent in 1999 to 0.06 percent in 2005 for males. This is compounded by the fact that during 1999 to 2005, the public sector in Bangladesh has been actively recruiting women in professional categories. Nearly half of these women are teachers, and the rest are nurses and paramedics in the health sector. This is evident in our data. More females moved to a skilled occupation (for example, professional), from 6 percent in 1999 to 23 percent in 2005. In comparison, the percentage of males in professional occupations remained rather stable. The percentage working in production-related occupations also increased for females, from 28 percent in 1999 to 36 percent in 2005. But it increased rapidly for males, from 16 percent in 1999 to 34 percent in 2005.

2.4 Estimation Results

We start with a discussion of the probit regression (selection equation) and then turn to (log) wage regression results. The OLS and the quantile wage regression estimates for the different specifications are presented in Tables A2 to A5 in Appendix A for males and females corresponding to survey years 1999 and 2005. We next describe the decomposition results.

2.4.1 Probit Results

We estimate probit regressions separately for males and females.[16] For ease of interpretation the results are presented as marginal effects in Table A1 in Appendix A. The marginal effects indicate the percentage point change in the estimated probability when the particular variable is changed and all other variables are evaluated at their mean (the marginal effects). A change for the dichotomous (dummy) variables reflects a change from 0 to 1, while for the continuous variables the change reflects the normal partial derivative.

Looking at the results of Table A1, we observe that males' probability of full-time work decreases with age from 1999 to 2005. In case of females, the likelihood of full-time employment increases with age, but decreases in the later years of life. In terms of education, both highly educated (i.e. graduate) men and women are more likely to work full-time. Both married and divorced men are more likely to work full-time than single men, though the results are only statistically significant for married men in 2005. On the other hand, married women are less likely to work full-time than single women, though the marginal effect of marital status is not statistically significant for women in 1999. This result confirms *a priori* expectations of the traditional role of women in the family unit. This finding is supported by the significant negative impact of the presence of young children on the probability of full-time employment for women, except for those women who have younger children (i.e. aged 0-5), in 1999. Finally, the probability of full-time work is higher for both men and women who live in urban areas compared with those who live in rural areas.

2.4.2 Wage Regression Results

The OLS and the quantile wage regression estimates, with and without conditioning on selection, are presented in Tables A2 to A5 in Appendix A. In this essay, our focus is on selectivity-corrected wage regression results (see Tables A4 and A5). We find that age is non-linear, rising to peak for males aged 55-59. The age effect, however, is not found for

[16] Separate regressions are estimated for males and females given the likelihood that different factors do not influence the employment decisions of men and women in the same manner. This is confirmed by the Wald test (χ^2 test). The Wald test statistics for the 1999 survey year is $\chi^2(28)$ = 202.6 (with a p-value of 0.000), and for the 2005 survey year it is $\chi^2(28)$ = 2703.80 (with a p-value of 0.000).

female employees in both survey years. This emphasises that work experience (proxied by age) rewards men more than women.

The estimated rates of return from education increase monotonically for both men and women at the mean than for the reference educational attainment group (i.e. individuals with no schooling) in 1999, though the returns to primary education are not statistically significant for females. It is important to note that in spite of the positive rates of return to education, the impact of education is stronger for men than for women, particularly for those men who hold post-secondary and graduate degrees. That is, regardless of the types of jobs better educated women obtain, they are rewarded less than men. A similar pattern of wage effect is not found in 2005. Furthermore, in 2005 the rate of returns from education decline except for males who hold graduate degrees and for females who hold post-secondary and graduate degrees. Looking at the contribution of education to the quantiles as we move along the distribution, note that while the returns to highest educational attainment (graduate) increase from 16 percent at the bottom quantile up to 100 percent at the 90th quantile of the male wage distribution in 1999, they only go up to 28 percent at the 90th quantile in the case of women. The opposite is true in 2005, where it increases from 56 percent at the bottom quantile to only 20 percent at the 90th quantile in the male wage distribution; and up to 58 percent at the 90th quantile in the case of women. These findings suggest that education is a significant source of overall wage variation in Bangladesh. Moreover, education contributes to the wage gap at the top of the wage distribution (i.e. 90th quantile) in 1999 and at the bottom of the wage distribution in 2005 between males and females. Overall, these results suggest that increasing the overall level of education for the population located at the bottom of the distribution would not help to reduce gender inequality in wages.

Being married is associated with larger premiums for men than for women at the mean, as well as at different quantiles from 1999 to 2005. This result is a common empirical finding and may be attributable to employer's beliefs that married men are stable and reliable employees compared to married women, who are considered more likely to have broken career paths (Kidd and Viney 1991). We also observe that the returns to marriage decrease at higher quantiles. Furthermore, we find that this decline is more pronounced for women than for men.

With respect to occupation and industry, there was no *a priori* expectation as to the pattern of signs. Both men and women earned more in professional and administrative occupations at the mean at and above the 50th quantile in 1999. Interestingly, females earned more than males in both types of occupations, though there is a large drop in the returns to professional jobs for women at the 90th quantile in 2005. It is also found that males and females working in financial institutions and public administration earned significantly higher wages at the mean and at all quantiles in 1999; however, women earned higher returns than men except at the 90th quantile. The returns for working in financial institutions, however, become statistically insignificant for both males and females at the top quantiles in 2005. These results are to some extent varied for the rest of the sectors.

In summary, these results suggest that the returns to labour market characteristics are different for men and women in both survey years.

2.4.3 Decomposition Results without Selection Correction

We start with a discussion of the results of the decomposition at the mean for both survey years (Table 2.4). This forms an interesting baseline to compare our results with the rest of the distribution. In Table 2.4, we specifically give more emphasis to the proportion of the total wage gap that is attributable to discrimination. It may also reflect the effect of unmeasured variables in the regression model at the mean. The results that are presented use male wages as the reference category.[17] Decomposition of the OLS estimates reveals that in 1999 the wage difference between males and females is 0.4542 log points, which corresponds to a wage differential of $(exp(0.4542) - 1) \times 100 = 57$ percent. Decomposition of this gap reveals that the explained component is considerably smaller compared to the component due to discrimination. After accounting for differences in productive characteristics, the discrimination component is 93 percent of the total wage gap; and only 7 percent of the total wage gap is explained by the superior endowment of the male. The wage gap between males and females increases to 0.6488 log points (or 91 percent) in 2005. Compared to the results for 1999, we find that while the discrimination component (as a proportion of the total wage gap) is lower in 2005,

[17]The results are robust when we use female wages as the reference category.

discrimination continues to account for the majority of the observed wage gap. See Table A6 in Appendix A for more details of the decomposition.

The decomposition results based on the unconditional quantile regressions by survey year are also presented in Table 2.4. We find that for both surveys, the estimated total gender wage gap is higher at the lower end of the distribution. The gender wage gap is systematically higher in the 2005 sample compared to the 1999 sample, with the gap ranging from 25 to 72 percent in 1999 and 50 to 133 percent in 2005. Notice that the wage gap is lower at the 90th quantile of the wage distribution compared to anywhere else on the distribution. For both survey years and almost everywhere along the distribution, discrimination accounts for the majority of the gender wage gap, ranging from 77 percent at the 90th quantile to 101 percent at the median in 1999 and from 73 percent at the 25th quantile to 104 percent at the 90th quantile in 2005. The proportion of the wage gap due to discrimination decreases monotonically in absolute terms for higher income quantiles in 1999.[18] In 2005, however, the effect is reversed at the highest quantile.

With the exception of the 90th quantile, the proportion of the gender wage gap due to discrimination is lower in the 2005 sample. It therefore appears that the relative position of women in the labour market has improved over time. These findings are similar to those reported by Jolliffe and Campos (2005) for Hungary, who showed a dramatic decline in discrimination from 1986 to 1998 following the introduction of the free-market system. As with Jolliffe and Campos (2005) our results could be representative of a decline in the 'taste for discrimination' on the part of the employers driven by the process of economic liberalisation and the emergence of a free-market economy.

Turning to the contribution of different characteristics (endowments) of men and women as a proportion of the wage gap reveals that differences in characteristics mostly are in favour of males both at the bottom and at the top end of wage distribution. While making up about 17 percent (0.0831 log points) at the lower end of the distribution, it accounts for 23 percent (0.0523 log points) of the difference between high-earning women and their male counterparts in 1999, highlighting the relevance of endowment

[18]A similar result is obtained by Wei and Bo (2007), who find that the wage gap due to discrimination is much lower at the top end of the wage distribution compared to the bottom end in urban China.

effect at the upper end of the wage distribution.[19] The pattern changes slightly in 2005: while the contribution of characteristics is in favour of males at the lower end of the wage distribution, it changed in favour of females at the 90th quantile. Thus, improvement in observed characteristics between 1999 and 2005 among high-earning women tended to reduce the gender wage gap, but discrimination against them completely wiped out these gains.[20]

We next turn to the decomposition of the change in wages from 1999 to 2005, using 2005 wage coefficients as the reference category. The choice of the reference category is arbitrary. An alternative decomposition could be obtained by taking 1999 as the reference category. These results are presented in Table 2.5. Almost everywhere (the exception being the 75th quantile), the wage gap has increased over the relevant period: from 36 percent at the 25th quantile to 20 percent at the 90th quantile. What is interesting is that at the lower end of the wage distribution ($\nu = 0.10, 0.25, 0.50$), the endowment effect is actually negative, indicating that in the absence of discrimination the gender wage gap would have been in favour of women and progressively so in lower quantiles, reaching -0.21 log points at the 25th quantile. The most remarkable finding is changes in educational attainment in favour of women that helps reduce the gender wage gap.[21] Discrimination against women, however, completely wipes out these beneficial effects.

At the upper end of the distribution ($\nu = 0.90$) less than 30 percent of the change in total wages is explained by discrimination. This result appears to suggest that once women have reached a position where they are at the higher end of the wage distribution, they do not face significant discrimination. In other words, the male premium is not particularly high at the upper end of the wage distribution (see Table 2.4). This could also be related to selection: women whose earnings place them at the higher end of the wage distribution might not be a random subset of the sample of women. We next turn to this issue of selection.

[19] It is found that men are heavily favoured in terms of industrial distribution of employment and age at the 90th quantile (see Table A6 in Appendix A).

[20] The main contributor to the narrowing gender wage differentials is declining occupational gap, which resulted mainly from the movement of female workers into the professional occupations (see Table 2.3).

[21] See Table A7 in Appendix A.

2.4.4 Decomposition Results with Selection Correction

How important is the selection bias? From Tables A4 and A5 in Appendix A, it appears that the answer to this question depends on the sample and the quantile under consideration. For the sample of women the coefficient estimate of the inverse Mill's ratio (λ) is positive and statistically significant in the OLS wage equation in 2005. What this implies is that women who actually work full-time have higher earning potential in full-time work than women in general. The selection correction term λ is, however, not statistically significant at selected quantiles in both survey years. Therefore, we cannot conclude that there is any correlation between the unobserved factors that affect the selection into full-time employment and a woman's wage across the wage distribution. For males, on the other hand, while the coefficient estimates of λ is positive and significant at the mean in both survey years, it becomes statistically significant only at the upper end of the wage distribution (see Figure 2.3). This indicates that the unobserved characteristics (such as 'innate intelligence' or 'drive') that result in men choosing to be in full-time employment tend to increase a man's wage at the upper end of the wage distribution.[22] Although λ is not always statistically significant, for the sake of consistency we compute and present (in Tables 2.6 and 2.7) the decomposition results adjusted for sample selection bias.

The decomposition results at the mean in both survey years reveal that the differences in productive characteristics are in favour of males (Table 2.6). Although the discrimination component is the major component of the wage gap in 2005, it becomes negative in 1999. This suggests that if men and women obtained a similar wage for their productive characteristics, women would have obtained a higher mean wage. The negative discrimination might be attributable to the labour market valuing the unobservable characteristics of women who self-select into full-time wage employment more often than that of men.[23]

The decomposition results using the selectivity-corrected quantile regression model paints a rather different picture, particularly with respect to the discrimination

[22] A number of studies use the Heckman estimator, but the estimates of the coefficient on the inverse Mill's ratio (λ) in the male wage equation differ significantly: Ahmed and Maitra (2010) reported a significant, negative coefficient for Bangladesh, Ashraf and Ashraf (1993) a positive significant coefficient for Pakistan.

[23] Interestingly, this is not a new result. Using LFS 1999, Ahmed and Maitra (2010) obtained similar results for the rural labour market of Bangladesh when they used male wage structure as the reference category.

effect. Women at the bottom and the top quantiles have benefitted from reduced discrimination in 2005, going from -0.03 log points at the 10^{th} quantile to -0.35 log points at the 90^{th} quantile.

Once sample selection bias is taken into account it is important to distinguish between observed wages and offered wages. The incidence of sample selection bias indicates that the observed wage differential will differ from the unobserved differential (differences in selection bias) in terms of wage offers. When the selection effect is positive (negative), the offered wage differential is lower (higher) than the observed wage differential. The extent of observed wage gap is therefore over-estimated (under-estimated), depending on the sign of the selection effect. In our analysis, the selection effect is positive except at the 25^{th} quantile in 2005 (see Table 2.6). Therefore, the extent of the observed wage gap is likely to be over-estimated if we ignore the selection issue into full-time employment. These results, however, should be interpreted cautiously as λ is never statistically significant (except at the mean in 2005) for females.

Turning next to decomposition results for the change in the wage gap from 1999 to 2005 (Table 2.7), we see that inclusion of the selection term does not change the results pertaining to the endowment effects. However, women benefitted from a decline in discrimination both at the lower and upper ends of the wage distribution. The selection effect is generally negative, which implies an increase in the offered wage gap from 1999 to 2005.

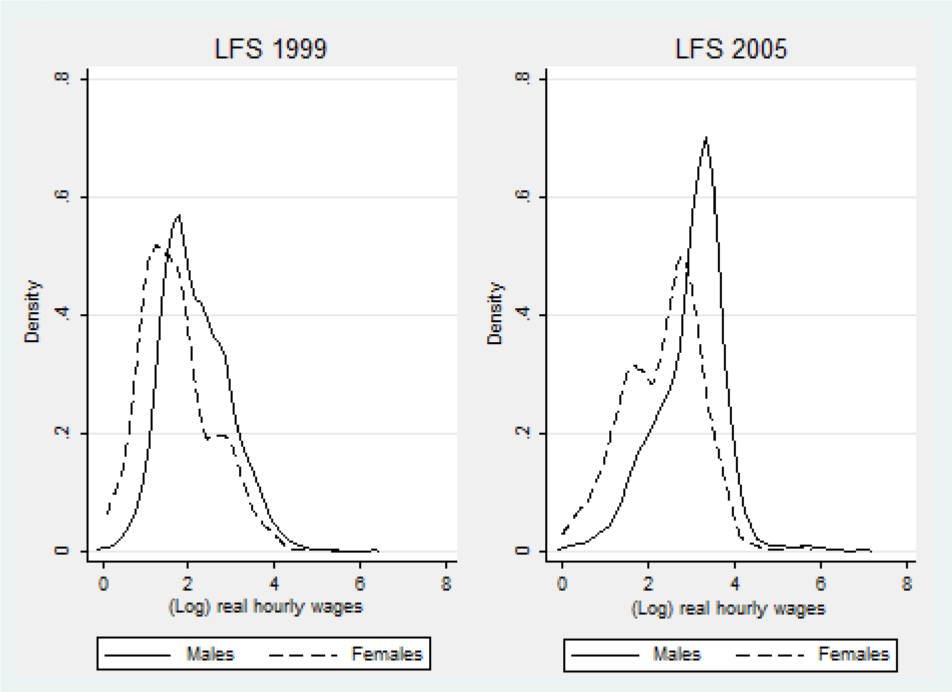
2.5 Summary and Conclusion

The main objective of this essay has been to examine whether the gender wage gap varies along the wage distribution. We also investigated whether the gender wage gap changes from 1999 to 2005 across the wage distribution to assess the contribution of different factors that may explain changes in the gender wage gap, both at the mean and also at other points of the distribution. Finally, we considered the effects of sample selection (selecting into full-time employment for both males and females) on the gender wage gap at different points of the distribution of wages.

Our decomposition results indicate that women employees are paid less on average compared to their male counterparts from 1999 to 2005 and the gap is greater at the lower end of the wage distribution. The major component of the wage gap is attributed to labour market discrimination against women and it is lower for high-wage earners compared to low-wage earners. However, the size of the endowment effect varies significantly over the period under consideration and is mostly in favour of men. Analyses of the changes in the gender wage gap by earnings percentile show that the gap widened much more at the lower end of the wage distribution than at the upper end over the survey years. A sizeable proportion of the increase in the gender wage gap at the lower end of the distribution is due to an increase in discrimination against females. Our results also show that not controlling for sample selection is likely to over-estimate the observed wage gap across the wage distribution. The selection-corrected wage gap (that is, the offered wage gap) is found to be predominantly due to discrimination against women.

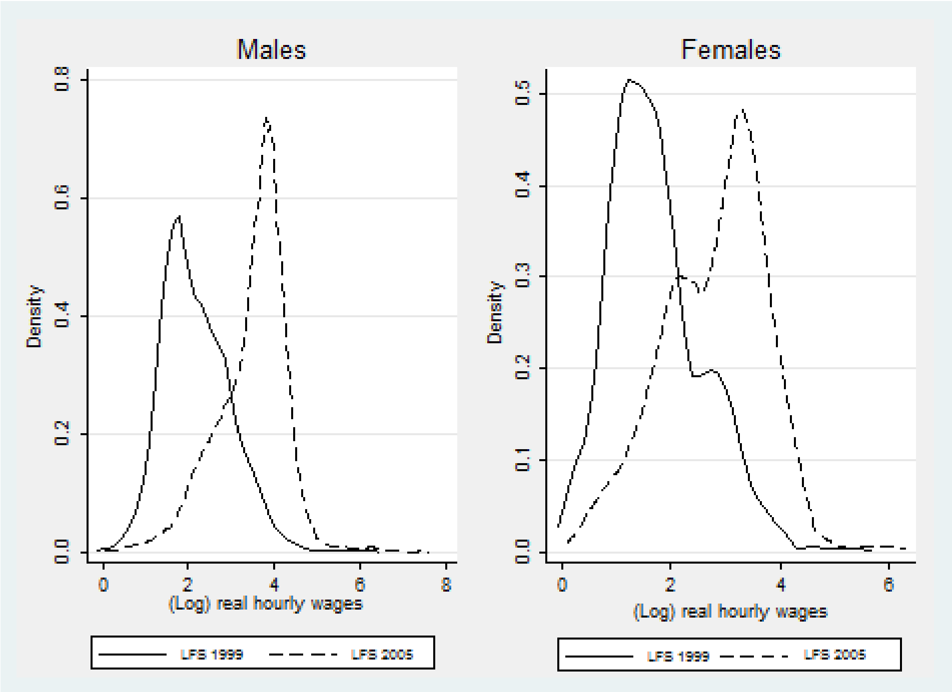
What causes the gender wage gap and why is the extent of the gender wage gap greater at the lower end of the distribution? It is possibly the result of a combination of a number of different factors such as low level of unionism, social norms, and the absence of structured childcare. Unfortunately we are unable to explore these hypotheses using the LFS datasets. The analysis of this essay shows that discrimination is a major part of the wage differential along the entire wage distribution. This strongly suggests that, although the Bangladeshi labour code stipulates equal pay and equal employment opportunity, there is considerable under-utilisation and under-appreciation of women's skills in the labour market. While laws have been passed and the legislature has accepted the role of gender-based affirmative action policies in reducing the gender wage gap, there is considerable lack of enforcement of these laws. To attain true gender equality we need policies that change employers' 'taste for discrimination,' and we need stronger legal enforcement.

Figure 2.1: Distribution of (Log) Real Hourly Wages, by Gender



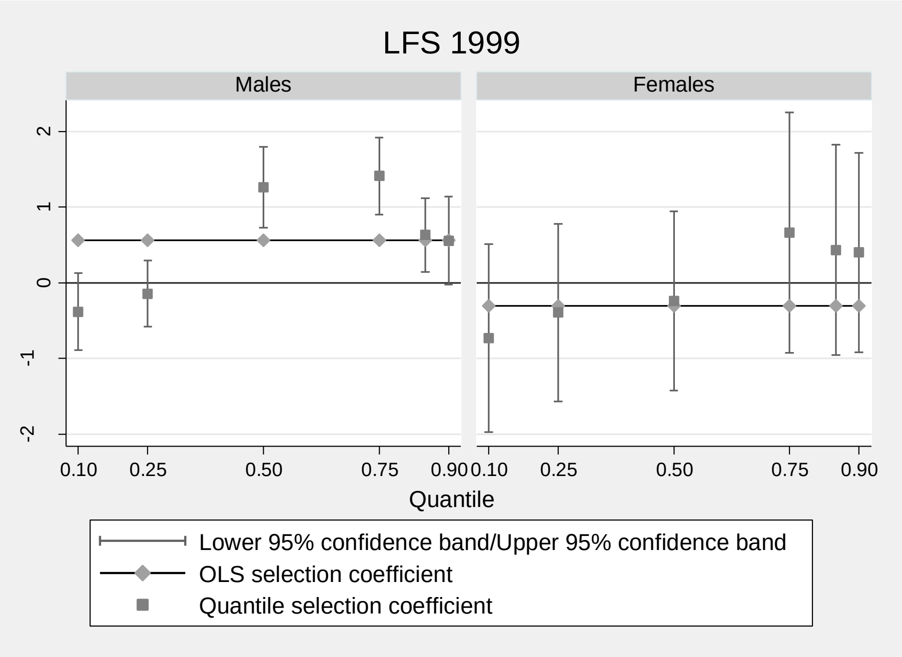
Source: Data are from LFS 1999 and LFS 2005.

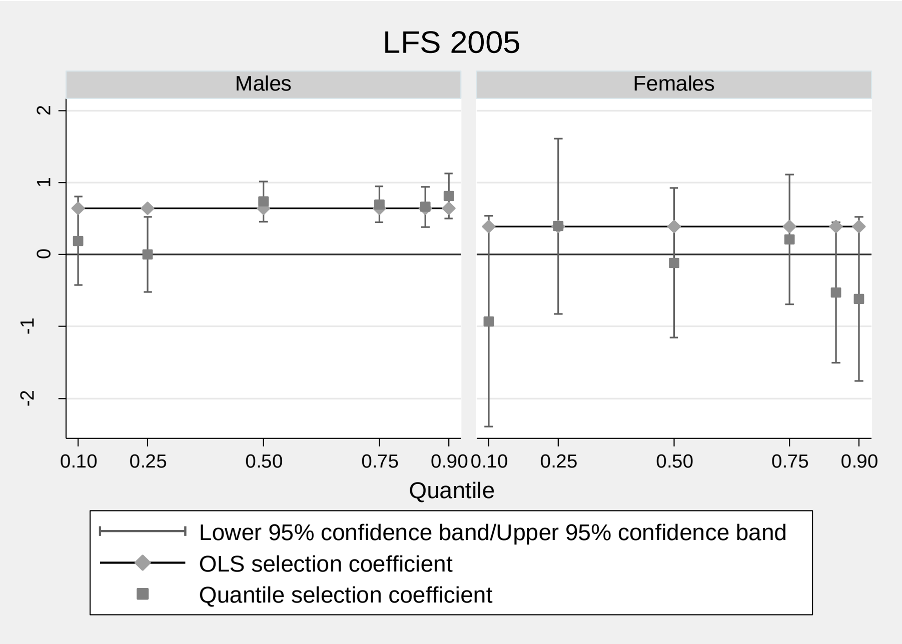
Figure 2.2: Changes in the Distribution of (Log) Real Hourly Wages, by Gender



Source: Data are from LFS 1999 and LFS 2005.

Figure 2.3: Is there a Selection Effect?





Source: Data are from LFS 1999 and LFS 2005.

Table 2.1: (Log) Real Hourly Wages and Gender Wage Gap over the Different Quantiles

Quantile	Males			Females			Gender Wage Gap		
	1999	2005	2005-1999	1999	2005	2005-1999	1999	2005	2005-1999
0.10	1.315	1.828	0.513	0.813	1.097	0.285	0.503	0.730	0.228
0.25	1.631	2.510	0.879	1.089	1.663	0.574	0.542	0.848	0.305
0.50	2.061	3.132	1.071	1.592	2.479	0.887	0.469	0.653	0.184
0.75	2.735	3.489	0.754	2.180	3.014	0.834	0.555	0.475	-0.080
0.90	3.221	3.804	0.583	2.995	3.399	0.404	0.226	0.405	0.179
Mean	2.173	2.973	0.800	1.719	2.324	0.605	0.454	0.649	0.195

Notes: Data are from LFS 1999 and LFS 2005.The wage gap is measured as the difference between the (log) male real hourly wages and the (log) female real hourly wages.

Table 2.2: Descriptive Statistics: Full Sample

	LFS 1999					LFS 2005				
	Males		Females		Difference	Males		Females		Difference
Variables	Mean	Std. Dev.	Mean	Std. Dev.	*t*-test	Mean	Std. Dev.	Mean	Std. Dev.	*t*-test
Employment - Reference: self-employed or employed in family businesses										
Full-time Employment	0.4340	(0.4956)	0.4499	(0.4976)	-1.32	0.3684	(0.4824)	0.1712	(0.3767)	43.36 ***
Age - Reference: 60 or higher										
Age 15 -19	0.0863	(0.2808)	0.1603	(0.3669)	-10.21 ***	0.1037	(0.3049)	0.0422	(0.2011)	21.90 ***
Age 20 -24	0.0926	(0.2898)	0.1780	(0.3826)	-11.39 ***	0.1148	(0.3188)	0.1489	(0.3560)	-10.50 ***
Age 25-29	0.1180	(0.3227)	0.1749	(0.3800)	-6.99 ***	0.1230	(0.3284)	0.1705	(0.3761)	-14.15 ***
Age 30-34	0.1366	(0.3434)	0.1441	(0.3513)	-0.89	0.1223	(0.3276)	0.1504	(0.3575)	-8.49 ***
Age 35-39	0.1535	(0.3604)	0.1178	(0.3224)	4.11 ***	0.1381	(0.3450)	0.1508	(0.3579)	-3.71 ***
Age 40-44	0.1333	(0.3399)	0.0763	(0.2656)	7.06 ***	0.1186	(0.3233)	0.1156	(0.3197)	0.95
Age 45-49	0.1045	(0.3059)	0.0693	(0.2540)	4.82 ***	0.1070	(0.3091)	0.0892	(0.2850)	5.93 ***
Age 50-54	0.0770	(0.2666)	0.0430	(0.2028)	5.40 ***	0.0754	(0.2641)	0.0621	(0.2413)	5.20 ***
Age 55-59	0.0502	(0.2184)	0.0197	(0.1391)	5.99 ***	0.0508	(0.2197)	0.0392	(0.1940)	5.52 ***
Education - Reference: No Schooling										
Primary school	0.2381	(0.4259)	0.1901	(0.3925)	4.66 ***	0.2380	(0.4259)	0.2290	(0.4202)	2.14 **
Secondary school	0.2051	(0.4038)	0.1663	(0.3725)	3.97 ***	0.2257	(0.4180)	0.1627	(0.3691)	15.64 ***
Post-secondary school	0.1236	(0.3291)	0.1036	(0.3049)	2.50 **	0.1276	(0.3337)	0.0688	(0.2530)	18.81 ***
Graduate	0.0672	(0.2503)	0.0460	(0.2096)	3.54 ***	0.0568	(0.2314)	0.0300	(0.1705)	12.40 ***
Marital Status - Reference: Single										
Married	0.7983	(0.4013)	0.6466	(0.4781)	14.96 ***	0.7666	(0.4230)	0.8065	(0.3951)	-9.67 ***
Divorced	0.0011	(0.0335)	0.0369	(0.1886)	-18.12 ***	0.0017	(0.0414)	0.0244	(0.1543)	-27.67 ***
Widowed	0.0027	(0.0521)	0.1107	(0.3139)	-33.18 ***	0.0065	(0.0803)	0.0985	(0.2980)	-58.02 ***

Continued on next page

Table 2.2 (continued): Descriptive Statistics: Full Sample

	LFS 1999					LFS 2005				
	Males		Females		Difference	Males		Females		Difference
Variables	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test
Presence of Children - Reference: If household has no children										
Number of children 0-5	0.7192	(0.8516)	0.6229	(0.7954)	4.67 ***	0.6283	(0.7693)	0.6399	(0.7999)	-1.51
Number of children 6-12	1.0194	(1.0280)	0.9009	(0.9691)	4.75 ***	0.8517	(0.9426)	0.8931	(0.9725)	-3.45 ***
Home Ownership - Reference: If household owns an accomodation										
Household pays no rent	0.0292	(0.1685)	0.0551	(0.2282)	-5.90 ***	0.0191	(0.1367)	0.0236	(0.1518)	-3.27 ***
Household pays rent	0.2119	(0.4087)	0.3140	(0.4642)	-9.98 ***	0.0962	(0.2948)	0.0768	(0.2664)	6.78 ***
Region - Reference: Rural										
Urban	0.5080	(0.0048)	0.6355	(0.0108)	-10.48 ***	0.4005	(0.4900)	0.3390	(0.4734)	12.79 ***
Number of males 65 or higher	0.0398	(0.1974)	0.0581	(0.2362)	-3.67 ***	0.0679	(0.2521)	0.0813	(0.2736)	-5.27 ***
Number of females 65 or higher	0.0318	(0.1754)	0.0298	(0.1702)	0.36	0.0626	(0.2440)	0.0571	(0.2330)	2.32 **
Quintile - Reference: Quintile1										
Quintile2	0.2098	(0.4072)	0.1547	(0.3617)	5.62 ***	0.1885	(0.3911)	0.2105	(0.4076)	-5.61 ***
Quintile3	0.2062	(0.4046)	0.1653	(0.3716)	4.18 ***	0.1964	(0.3972)	0.2106	(0.4078)	-3.61 ***
Quintile4	0.1956	(0.3967)	0.2184	(0.4133)	-2.33 **	0.2033	(0.4024)	0.1848	(0.3882)	4.67 ***
Quintile5	0.1924	(0.3942)	0.2396	(0.4270)	-4.83 ***	0.2134	(0.4097)	0.1562	(0.3630)	14.49 ***

Notes: Data are from LFS 1999 and LFS 2005. Std. Dev. is standard deviation. *t*-test for difference (Males-Females). *** p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 2.3: Descriptive Statistics: Sample in Full-time Employment

	LFS 1999					LFS 2005				
	Males		Females		Difference	Males		Females		Difference
Variables	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test
Age - Reference: 60 or higher										
Age 15 -19	0.1021	(0.0044)	0.1382	(0.0116)	-3.18 ***	0.1150	(0.0025)	0.1229	(0.0069)	-1.11
Age 20 -24	0.1025	(0.0046)	0.1708	(0.0126)	-5.89 ***	0.1168	(0.0025)	0.1330	(0.0071)	-2.24 **
Age 25-29	0.1246	(0.0048)	0.1742	(0.0127)	-4.00 ***	0.1299	(0.0026)	0.1640	(0.0077)	-4.48 ***
Age 30-34	0.1414	(0.0051)	0.1562	(0.0122)	-1.15	0.1293	(0.0026)	0.1518	(0.0075)	-2.98 ***
Age 35-39	0.1524	(0.0053)	0.1393	(0.0116)	1.00	0.1426	(0.0028)	0.1514	(0.0075)	-1.11
Age 40-44	0.1317	(0.0050)	0.0798	(0.0091)	4.32 ***	0.1190	(0.0026)	0.1142	(0.0067)	0.67
Age 45-49	0.1021	(0.0044)	0.0809	(0.0091)	1.94 *	0.1013	(0.0024)	0.0744	(0.0055)	4.05 ***
Age 50-54	0.0784	(0.0039)	0.0416	(0.0067)	3.88 ***	0.0717	(0.0020)	0.0464	(0.0044)	4.48 ***
Age 55-59	0.0784	(0.0039)	0.0416	(0.0067)	3.85 ***	0.0455	(0.0016)	0.0284	(0.0035)	3.75 ***
Education - Reference: No Schooling										
Primary school	0.2137	(0.4100)	0.1843	(0.3879)	1.98 **	0.2151	(0.4109)	0.1645	(0.3708)	5.57 ***
Secondary school	0.1775	(0.3821)	0.1000	(0.3002)	5.71 ***	0.1851	(0.3884)	0.1159	(0.3202)	8.14 ***
Post-secondary school	0.1289	(0.3351)	0.1022	(0.3031)	2.20 **	0.1418	(0.3489)	0.1732	(0.3785)	-3.99 ***
Graduate	0.0915	(0.2884)	0.0753	(0.2640)	1.56	0.1029	(0.3039)	0.1452	(0.3524)	-6.10 ***
Training - Reference: No Training										
Vocational	0.0354	(0.1848)	0.0191	(0.1370)	2.50 **	0.0130	(0.1134)	0.0118	(0.1081)	0.49
General	0.0261	(0.1595)	0.0270	(0.1621)	-0.14	0.0517	(0.2213)	0.0849	(0.2787)	-6.48
Marital Status - Reference: Single										
Married	0.7787	(0.4152)	0.6056	(0.4890)	11.05 ***	0.7539	(0.4308)	0.6220	(0.4850)	13.47
Divorced	0.0011	(0.0328)	0.0618	(0.2409)	-16.38 ***	0.0016	(0.0401)	0.0582	(0.2341)	-27.91
Widowed	0.0032	(0.0568)	0.1416	(0.3488)	-25.31 ***	0.0056	(0.0745)	0.1461	(0.3533)	-44.05

Continued on next page

Table 2.3 (continued): Descriptive Statistics: Sample in Full-time Employment

	LFS 1999					LFS 2005				
	Males		Females		Difference	Males		Females		Difference
Variables	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test
Region - Reference: Rural										
Urban	0.5477	(0.4978)	0.7483	(0.4342)	-11.23 ***	0.4478	(0.4973)	0.6129	(0.4872)	-14.89 ***
Occupation - Reference: Service										
Professional	0.1427	(0.3498)	0.0674	(0.2509)	6.13 ***	0.1076	(0.3099)	0.2332	(0.4229)	-17.23 ***
Administrative	0.0723	(0.2590)	0.1382	(0.3453)	-6.55 ***	0.0060	(0.0770)	0.0022	(0.0467)	2.28 **
Clerical	0.0579	(0.2335)	0.0169	(0.1288)	5.09 ***	0.0631	(0.2431)	0.0700	(0.2552)	-1.26
Sales	0.0663	(0.2488)	0.2944	(0.4561)	-21.32 ***	0.0720	(0.2584)	0.0232	(0.1505)	8.81 ***
Agriculture	0.3532	(0.4780)	0.1528	(0.3600)	11.88 ***	0.3160	(0.4649)	0.1325	(0.3392)	18.19 ***
Production	0.1628	(0.3692)	0.2831	(0.4508)	-8.58 ***	0.3365	(0.4725)	0.3640	(0.4812)	-2.59 ***
Industry - Reference: Hospitality										
Agriculture	0.3441	(0.4751)	0.1303	(0.3369)	12.81 ***	0.3104	(0.4627)	0.1133	(0.3170)	19.72 ***
Manufacturing	0.2416	(0.4281)	0.3393	(0.4737)	-6.12 ***	0.2911	(0.4543)	0.3740	(0.4840)	-8.10 ***
Wholesale and Retail	0.0661	(0.2484)	0.0157	(0.1245)	5.90 ***	0.0636	(0.2440)	0.0092	(0.0954)	10.54 ***
Transport	0.0715	(0.2576)	0.0090	(0.0944)	7.14 ***	0.0816	(0.2737)	0.0131	(0.1138)	11.81 ***
Financial institution	0.0240	(0.1530)	0.0089	(0.0942)	2.82 ***	0.0279	(0.1646)	0.0446	(0.2065)	-4.40 ***
Real estate	0.0047	(0.0688)	0.0056	(0.0748)	-0.34	0.0073	(0.0849)	0.0052	(0.0723)	1.08
Public administration	0.0775	(0.2674)	0.0371	(0.1891)	4.31 ***	0.0679	(0.2515)	0.0547	(0.2274)	2.37 **
Education	0.0468	(0.2113)	0.1045	(0.3061)	-6.87 ***	0.0707	(0.2563)	0.1767	(0.3815)	-17.26 ***
Health	0.0130	(0.1131)	0.0315	(0.1747)	-4.04 ***	0.0190	(0.1365)	0.0626	(0.2422)	-12.68 ***

Notes: Data are from LFS 1999 and LFS 2005. Std. Dev. is standard deviation. *t*-test for difference (Males-Females). *** p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 2.4: Decomposition of the Gender Wage Gap

Quantile	Total Observed Gap	Percentage Gap	Endowment	Discrimination	Proportion Due to Discrimination
1999					
0.10	0.5026	65.30	0.0831	0.4195	0.8347
0.25	0.5423	72.00	0.0194	0.5229	0.9642
0.50	0.4688	59.81	-0.0073	0.4761	1.0156
0.75	0.5551	74.21	0.0033	0.5518	0.9941
0.90	0.2261	25.37	0.0523	0.1738	0.7687
Mean (OLS)	0.4542	57.49	0.0318	0.4224	0.9300
2005					
0.10	0.7303	107.57	0.1691	0.5612	0.7685
0.25	0.8475	133.38	0.2301	0.6174	0.7285
0.50	0.6530	92.13	0.1324	0.5206	0.7972
0.75	0.4747	60.75	0.0359	0.4388	0.9244
0.90	0.4050	49.93	-0.0173	0.4223	1.0427
Mean (OLS)	0.6488	91.32	0.1231	0.5257	0.8103

Notes: Data are from LFS 1999 and LFS 2005.
Percentage gap computed as (exp(Total Observed Gap)-1)× 100.
Male wages is the reference category.
A positive entry indicates an advantage in favour of males.

Table 2.5: Decomposition of Change in Gender Wage Gap

Quantile	Total Observed Gap	Percentage Gap	Endowment	Discrimination	Proportion Due to Discrimination
0.10	0.2277	25.57	-0.1720	0.3997	1.7554
0.25	0.3052	35.69	-0.2090	0.5142	1.6848
0.50	0.1852	20.23	-0.1529	0.3371	1.8301
0.75	-0.0804	-7.73	0.0413	-0.1217	1.5137
0.90	0.1788	19.58	0.1273	0.0515	0.2880
Mean (OLS)	0.1946	21.48	-0.0973	0.2919	1.5000

Notes: Data are from LFS 1999 and LFS 2005.
Percentage gap computed as (exp(Total Observed Gap)-1)× 100.
Year 2005 is the reference category.
A positive entry indicates an advantage in favour of males.

Table 2.6: Decomposition of the Gender Wage Gap with Selection

Quantile	Total Observed Gap	Endowment	Discrimination	Selection Effect	Percentage Gap	Proportion Due to Discrimination
1999						
0.10	0.5026	0.0955	0.2461	0.1610	40.72	0.7204
0.25	0.5423	0.0239	0.3977	0.1207	52.44	0.9433
0.50	0.4688	-0.0482	-0.2673	0.7842	-27.05	0.8475
0.75	0.5551	-0.0424	0.1842	0.4133	15.23	1.2990
0.90	0.2261	0.0343	0.0978	0.0940	14.12	0.7403
Mean (OLS)	0.4542	0.0135	-0.0067	0.4474	0.68	-0.9853
2005						
0.10	0.7303	0.1659	-0.0338	0.5982	14.12	-0.2559
0.25	0.8475	0.2301	0.8270	-0.2096	187.80	0.7823
0.50	0.6530	0.1203	0.0548	0.4779	19.61	0.3130
0.75	0.4747	0.0244	0.1708	0.2795	21.56	0.8750
0.90	0.4050	-0.0308	-0.3493	0.7850	-31.61	0.9192
Mean (OLS)	0.6488	0.1125	0.3922	0.1441	65.65	0.7771

Notes: Data are from LFS 1999 and LFS 2005.
Percentage gap computed as (exp(Offered Wage Gap)-1)× 100.
Male wages is the reference category.
Proportion of discrimination is calculated from the offered wage gap.
A positive entry indicates an advantage in favour of males.

Table 2.7: Decomposition of Change in Gender Wage Gap with Selection

Quantile	Total Observed Gap	Endowment	Discrimination	Selection Effect	Percentage Gap	Proportion Due to Discrimination
0.10	0.2277	-0.1634	-0.0462	0.4373	-18.91	0.2204
0.25	0.3052	-0.2124	0.8479	-0.3303	88.80	1.3342
0.50	0.1852	-0.1498	0.6446	-0.3103	63.97	1.3035
0.75	-0.0804	0.0414	0.0121	-0.1339	5.5	0.2262
0.90	0.1788	0.1352	-0.6473	0.6909	-40.08	1.2640
Mean (OLS)	0.1946	-0.1051	0.6029	-0.3032	64.51	1.2111

Notes: Data are from LFS 1999 and LFS 2005.
Percentage gap computed as (exp(Offered Wage Gap)-1)×100.
Year 2005 is the reference category.
Proportion of discrimination is calculated from the offered wage gap.
A positive entry indicates an advantage in favour of males.

Chapter 3

Market Returns and the Gender Gap in the Demand for Tertiary Education: Evidence from Bangladesh

3.1 Introduction

There has been tremendous progress in female participation in higher education in the last three decades. However, the male-female gap in higher education enrolment rates remains significant in developing countries. For example, the gap is widest in South Asia (4 percent) and the Middle East and North Africa (6 percent). In contrast, women are increasingly well represented in Latin America (18 percent for both genders).[24] There are many possible explanations for the gender gap in education, with the most frequently cited being that if the labour market rewards male schooling more than female, then parents have incentive to invest more in male schooling (Aslam 2009; Kingdon 1998, 2002). On the other hand, if the reverse is true, the greater schooling of males relative to females may reflect an asymmetry in parental incentives to educate males and females due to son preference[25] (Kingdon 1998, 2002). Such differential treatment may arise if parents expect that more direct benefit could be obtained from investing in sons. This notion is compounded by the fact that, in some countries, sons typically provide for parents in their old age, while daughters leave and become part of a different household economic unit (following marriage). Other explanations of different parental treatment of sons and daughters in education could be that the opportunity costs of schooling, in terms of home production, are greater for older daughters in many

[24]All the statistics here are based on the Task Force on Higher Education and Society (2000) "Higher Education in Developing Countries: Peril and Promise", Washington DC: World Bank.
[25]The term 'son preference' refers to the attitude that sons are more important and more valuable than daughters.

developing countries. This may be because older daughters are generally responsible for housework, cooking and taking care of younger siblings. Finally, social and cultural factors may also play a role. For example, higher education may raise the difficulty a woman has in finding a spouse if social norms require that husbands have higher education levels than wives, giving further disincentive for parents to invest in a daughter's schooling before marriage. In addition, a woman's education is rarely valued in traditional marriage markets, and so would not directly improve her choice set.

Several studies find that returns to higher education are generally greater for males than for females, particularly in developing countries (Asadullah 2006; Duraisamy 2002; Gibson and Fatai 2006; Psacharopoulos and Patrinos 2004).[26] However, there is also some evidence that the economic returns to higher education appear to be as great (or greater) for women as for men in many parts of the developing world, where gender differences in higher education participation remains high (Aslam 2009; Behrman and Deolalikar 1995).

A limitation of some of this literature is that it examines gender differences in the effect of schooling on labour market returns. Thus, the paths have not been identified through which labour market signals enter into the schooling decision of parents and children (Anderson et al. 2003). Due to this exclusion, these studies fail to explain how the labour market signals and constrains schooling choices or how the benefits from education affect parental investment decisions regarding children's schooling across genders. A number of researchers have included estimated measures of returns to education as determinants of participation in school, and have established that returns to education affect schooling decisions (see, for example, Anderson et al. 2003; Chamarbagwala 2008; Kingdon and Theopold 2006; Yamauchi-Kawana 1997). However, the primary interest of all these studies has been primary and secondary education levels rather than tertiary education levels. Analysis by economists of the effect of labour market returns on gender differentials in the demand for tertiary education is especially limited. To our knowledge, the only related study is by Jacob (2002) in the United States. Therefore, there is a significant absence of studies for developing countries.

[26] 'Returns to education' means the proportional increase in earnings per year of schooling, following conventional usage.

The purpose of this essay is to provide a comprehensive analysis of gender differentials in the demand for tertiary education in a developing country setting. In doing so, this essay seeks to document the association between labour market returns and gender differentials in higher education in Bangladesh, a country characterised by well-documented gender inequality in regards to adult literacy and tertiary education, as well as in a number of other welfare measurements (Hausmann et al. 2008). The focus of this essay is on tertiary education because primary education has been compulsory in Bangladesh since 1993 and enrolment rates are quite high. There is also some evidence that females perform better than males at the secondary school level, in terms of enrolment status and years of schooling completed (Asadullah and Chaudhury 2009). However, women are under-represented in tertiary education. In Bangladesh, the gender parity index (the ratio of females to males) in tertiary education is approximately 0.53 (Hausmann et al. 2008). This is significantly lower than other countries at the same level of development, such as Pakistan.[27] This under-representation is of concern from a policy point of view. While the government of Bangladesh has recently begun focusing on tertiary education and has developed a National Strategic Plan for higher education, this focus has not been related to gender equity issues to increase female enrolment in tertiary institutions.

In this essay, the gender gap in the demand for tertiary education is tested by estimating differences in the determinants of enrolment in tertiary education of males and females using two nationally representative Bangladesh Labour Force Surveys (conducted in 1999-2000 and 2005-2006) datasets.[28] Of particular interest, we investigate whether differences in enrolment in tertiary education between males and females have any significant relationship with premiums in returns from secondary to tertiary education levels. Although the importance of labour market returns in explaining the gender gap in tertiary enrolment rates is not well established in Bangladesh, previous work indicates a large premium for men and women who possess tertiary education in Bangladesh, with the premium usually greater for men than for women (Asadullah 2006). This provides a first indication as to why female enrolment rates in tertiary education are lower than that of males in Bangladesh. However, this does not necessarily

[27] In Bangladesh, the gender parity index (the ratio of females to males) in primary and secondary education is 104 (Hausmann et al. 2008).

[28] The enrolment rates are ‘gross enrolment rates’ that are based on the ratio of school enrollers for a given grade level, relative to estimates of the population for the age ranges normally associated with that grade level.

imply that the other possible explanations of the gender gap in the tertiary enrolment rate in Bangladesh are trivial. In studies on the demand for education, the importance of social and demographic factors in the determination of school enrolment is generally accepted. These aspects include parent's education, size and composition of a household, age at first marriage or other factors that provides an explanation for the way a household values education. However, there is no evidence that social and demographic factors may explain gender differentials in tertiary enrolment rates between males and females. Hence, in this essay, to assess the importance of these factors, we include a range of pertinent characteristics along with economic returns to education.

The empirical analysis reaches three major conclusions. First, wage premiums do not have any significant effect on the gender gap in tertiary enrolment rates. Second, it provides empirical evidence that premiums in returns from secondary to tertiary education predict tertiary enrolment rates for males but not for females, once other potential explanatory variables are added. Third, this study also shows that parents' education is positively correlated with the tertiary enrolment rate for both genders, while demographic factors such as fertility preferences are negatively correlated with it. This result applies more to females than to males. Although this in itself is not a surprising result, no other study of developing countries has compared the effect of demographic factors with the effect of variables that measure returns to education.

3.2 Structure of Education in Bangladesh and Recent Developments

In Bangladesh, formal education is generally delivered through the government. However, a non-formal education system is offered by NGOs (non-government organisations) and the government to target disadvantaged children and young adults. The formal education system of Bangladesh is broadly divided into three major levels: primary, secondary and tertiary education. Primary education (i.e. Grade 1-5) consists of five years of schooling. Education at this stage normally begins at the age of six, which is the official age of entry into primary school according to the Primary Education Act, 1992. Primary education is free in all government schools and government-funded institutions. Secondary education consists of seven years of schooling divided into three cycles, Level 1 from Grades 6 to 8 (junior secondary), Level 2 from Grades 9 to 10 (secondary), and Level 3 from Grades11 to 12 (higher secondary/post-secondary).

There is a diversification of courses after three years of schooling at the junior secondary level: vocational and technical courses are offered in vocational and trade institutes. There are two nationwide public examinations. The first is the Secondary School Certificate examination (SSC), conducted at the end of Grade10. The other is the Higher Secondary Certificate examination (HSC), conducted at the end of Grade12. Higher secondary is followed by tertiary education, which comprises two to five years of full-time study. The minimum requirement for tertiary education is the Higher Secondary Certificate (HSC). HSC holders are qualified to enrol in bachelor degrees of three years (at pass level) or four years (at honours level) in colleges or universities. After successful completion of a pass/honours bachelor's degree course, students can enrol in a master's degree course. A master's degree lasts one year for holders of a bachelor's degree with honours, and for two years for holders of a bachelor's degree at the pass level. In 2011, however, a major reform of the education system in Bangladesh was undertaken. Under this reform, the duration of all degree courses became four years instead of the previous three years.

This study examines gender differences in the determinants of enrolment in tertiary education using nationally representative datasets collected before the recent reforms in education systems.

3.3 Market Returns and the Demand for Education

There is general consensus that the demand for schooling increases with increased economic returns to education. Therefore, the important research and policy questions are whether and how much of the expected economic returns to schooling affect demand for education. Generally, an individual's incentive to forego current income and invest in education depends on the wage difference between better and less educated labour (the returns to education) and the probability of finding employment that adequately rewards the skills achieved. If an individual expects a high return to education as a result of sufficient skill-incentive job opportunities in the future, then they are more likely to attend school.

The literature on this feedback effect of the labour market on schooling choice is mostly limited to developed countries. For example, Freeman (1975) shows that college graduates attend law school after considering the current wage received by starting lawyers, not expected wages. This leads to the assumption that students are myopic. If the current wage is higher in comparison with the past wage, a larger number of students will enrol in the law school, thus expanding the stock of lawyers in subsequent years. This larger stock will consequently lower the wage they will actually receive to the expected wage. In the next cycle of law school intake, this lower wage will lead to a decreasing number of enrolments. In the long run, equilibrium may or may not be attained. Will and Rosen (1979) find that the demand for college education in the United States is highly responsive to labour market returns. Using panel data on male Second World War veterans, the authors estimate earnings functions for those attending and those not attending college, adjusting for non-random selection into college. They find that the higher expected lifetime earnings (due to a higher level of initial earnings and faster earnings growth) of college graduates, relative to high school graduates, raise the probability of entering college.

Micklewright et al. (1990) examine the effects of unemployment on school leaving in the United Kingdom between 1978 and 1984. While previous studies in the United Kingdom found that higher unemployment discourages early school leaving, which implies that the opportunity cost effect dominated the anticipated returns and parental income effects, Micklewright et al. (1990) did not find evidence of this. Altonji (1991) finds that the *ex ante* rate of return to education is substantially higher for youths who actually start college than those who do not. Moreover, the *ex ante* probability for completing and obtaining a college degree is higher for those who choose to start college, thus ensuring large *ex post* payoffs. Jacob (2002) investigates the gender gap in college attendance rates in the United States and finds that non-cognitive skills rather than labour market returns have a substantial effect on the probability of enrolling in college.

Studies estimating the effect of labour market returns on educational choice have rarely been conducted in developing countries. Yamauchi-Kawana (1997) uses estimated village-level returns and finds a positive relationship between schooling returns and school enrolment in India. Anderson et al. (2003) estimate province-level returns to schooling for mothers and fathers in Malaysia, and find that only the mothers’ returns have a statistically significant positive effect on children’s educational attainment. In an unpublished paper, Gormly and Swinnerton (2004) showed that among South African liquidity-constrained households, children's school enrolment is positively correlated with regional returns to education. However, Kingdon and Theopold (2006) find that higher returns to education in the local labour market in India increases the opportunity cost of schooling for males, resulting in a negative relationship between returns and school participation. The authors find that the relationship becomes positive in the case of non-poor males and females. In contrast, Chamarbagwala (2008) provides evidence from India that higher regional returns to primary education not only increase the likelihood that males and females will attend school but also decreases the likelihood that they will work.

With the exception of Jacob (2002), none of the above studies tests whether premiums in returns from secondary to tertiary education affect educational decisions at the tertiary level. This essay differs from this earlier work in three crucial respects. First, this study also investigates whether gender differences in tertiary enrolment rates have any significant relationship with differences in the effect of premiums in returns from secondary to tertiary education levels between males and females. Second, we use a fixed-effect estimator to control for the presence of unobserved characteristics that may be correlated with returns to education. Third, this essay incorporates ideas from the literature that there is a link between demographic outcomes and a factor that is referred to as 'women's choice' (Dyson and Moore 1983; Sen 1985). Female participation in higher education is a relevant factor in this context. It is difficult, however, to determine beforehand the possible effect of demographic factors on female participation in higher education. Much of the literature indicates that demographic factors, such as marital status and age at marriage, are a significant deterrent for female enrolment (see, for example, Kingdon 2002). Field and Ambrus (2008) find a similar association in the context of Bangladesh. Thus, this enables us to examine not only the effects of labour market returns on the gender gap in tertiary enrolment rates, but also the effects of demographic factors that are acknowledged in the literature.

3.4 Data and Variables

Data

The empirical analysis in this essay uses two nationally representative datasets from Bangladesh. Specifically, we use data from two Labour Force Surveys conducted in 1999-2000 (henceforth LFS 1999) and 2005-2006 (henceforth LFS 2005). These two cross-sectional surveys were administered by the Bangladesh Bureau of Statistics (BBS). The questionnaire for both surveys is almost identical, and therefore overall inter-temporal compatibility is very good. The LFS 1999 and 2005 include household- and individual-level data: household size and composition, religion, income, assets, demographic variables (age, gender and marital status), education participation and attainment, and a detailed employment section on principal and subsidiary activities (industry, occupation, type and income earned, and intensity of each activity). Economic and schooling activities of all individuals aged six and above were recorded for the week prior to the survey.

For the purposes of analysing gender differences in higher education, a sample of persons who have enrolled in tertiary education is required. Traditionally, most persons in Bangladesh are enrolled in tertiary education by the age of 19. However, in many developing countries children delay their initial enrolment and start schooling after the age of six. Allowing for this, the sample was limited to adults aged 18-21 at the time of the survey. This yields a total of 3,389 individuals from 9,790 households in the LFS 1999 dataset and 13,000 individuals from 40,000 households in the LFS 2005 dataset. Households are selected from both rural and urban areas across all divisions in Bangladesh. These divisions are divided into 64 districts, which consist of sub-districts called ‘upazila parishad’ or ‘thana.’

Since some demographic variables, such as age at first marriage and fertility preferences, are not available in the LFS dataset, LFS 1999 and LFS 2005 are combined with Demographic and Health Survey (DHS) data for 1999 and 2004, respectively. Except for a number of demographic variables, both surveys are similar in many respects. The two surveys provide information at the district level. Moreover, DHS were conducted in November in both 1999 and 2004, while LFSs were conducted in September in both 1999 and 2005. We joined these two datasets by using district identifiers. Both datasets contain names and codes of divisions and districts that are included within them. We used this information to aggregate the information in the DHS up to the district level and matched them with the LFS datasets.

In this essay, enrolment decisions in tertiary education are defined as a binary dummy variable, indicating whether the individual is enrolled in Grade 13 (bachelor's or equivalent) at the time of the survey, conditional on completion of secondary education.[29] The data suggests that there is a significant gap in enrolment in tertiary education between males and females. Table 3.1 shows that male rates of participation in tertiary education are greater than female rates between the ages of 18 and 21 in both survey years. However, women have been obtaining tertiary schooling in large numbers from 1999 to 2005. The tertiary enrolment rate for women aged 18 jumped from 28 percent in 1999 to 42 percent in 2005, while the tertiary enrolment rate for women aged 20 increased from 21 percent in 1999 to 31 percent in 2005, thereby reducing the gender disparity in tertiary enrolments from 1999 to 2005. On the other hand, among the most mature students (those of at least 21 years), the enrolment rate for females in tertiary education is much lower than that of males. Consequently, the disadvantage of females in the tertiary enrolment rate is somewhat larger in the oldest age group, although it narrows between the two time periods.

Variables

Among the explanatory variables of the enrolment equation, wage premiums in returns from secondary to tertiary education deserve special mention. This variable is constructed as the (log) difference between average hourly wages of those who have completed tertiary education (Grade 13 or more: bachelor's or higher) and are in paid employment (working at least 40 hours per week) and those who have completed secondary education (Grade 10 or more) and are in paid employment (working at least 40 hours per week).[30] This measurement is calculated at district-rural-urban areas (henceforth 'district-area'; where area indicates either a rural or an urban area). The district-area is chosen to permit more variation in wage premiums. We constructed wage premiums for three age groups: 25-35, 25-45 and 25 and above as follows:

[29] The tertiary education category also includes those technical diploma and certificate courses achieved after secondary education.
[30]Jacob computes the college premium as the (log) difference between the median weekly earnings of college graduates (bachelor's or higher) and high school graduates at the state level.

$$R_{jt} = \overline{lnw}^{g}_{Tjt} - \overline{lnw}^{g}_{Sjt}$$

where R_{jt} is wage premiums from secondary to tertiary education calculated at the district-area level j at time t (1999 or 2005) of the gender $g(g = \text{male or female})$; $\overline{lnw}^{g}_{Tjt}$ and $\overline{lnw}^{g}_{Sjt}$ are the natural (logs) of mean hourly wages of tertiary school graduates (T) and secondary school graduates (S) calculated at the district-area level j at time t of the gender $g(g = \text{male or female})$. Hourly wages are computed by dividing monthly wages by the total hours of work per month. The survey collected information on the usual hours work per week but not the number of weeks worked during a month. Therefore, the monthly hours of work is computed by multiplying the usual hours of work per week by 52/12.

We find that, in both surveys, males earn significantly more than females at secondary and tertiary education levels (Tables 3.2 and 3.3). While at the secondary level in 1999 the disparity is greatest among the oldest cohorts (aged 25 and above) and smallest among the youngest cohort (aged 25-35), in 2005 the opposite is true. These results suggest that there has been a remarkable increase in labour market returns for males and females with secondary schooling among older cohorts between 1999 and 2005. Furthermore, both men and women with tertiary schooling earn more than those with secondary schooling across cohorts from 1999 to 2005. Thus, there has been a wage premium from secondary to tertiary education levels for both genders, consistent with the literature (Asadullah 2006, Bangladesh study; Aslam 2009, Pakistan study). It is also interesting to note that the wage premium is often greater for women than for men of different age cohorts. These findings are similar to those reported by Aslam (2009) for Pakistan,[31] however, they are not consistent with findings reported by Asadullah (2006) for Bangladesh. These results could be due to an inverse relationship between years of schooling and labour market discrimination, tastes and circumstances.[32] These higher labour market returns may provide an incentive to invest in education among females.

[31] Aslam (2009) showed a premium in returns from bachelor's to master's degree for females than for males.

[32] Dougherty (2005), in fact, suggests that the better educated a woman is, the more likely she is to be capable of resisting discrimination. At the same time, it is also possible that the better educated a woman is, the more likely she is to be willing to seek employment outside the low-paying traditionally female occupations. Also, it is likely that the better educated a woman is and the greater her potential earnings, the more capable she is of paying for childcare and other services that allow her to seek a wage offer that fully values her characteristics.

Figure 3.1 shows that gender differences in wage premiums in favour of women are greatest among the older cohort (aged 25-45) in 1999; however, in 2005 this was notable in the youngest cohort (aged 25-35). It is also interesting to observe that the wage premiums from secondary to tertiary education have declined from 1999 to 2005 for both males and females across different age cohorts. It appears that men and women with secondary schooling in 2005 are rewarded substantially more compared to those in 1999. This may reflect the growth of employment opportunities between 1999 and 2005 that requires low-skilled labour, especially for those with secondary education.

Other Variables

Together with wage premiums, several independent factors are expected to affect an individual's higher education decisions. In this study, these factors are single-year age dummy variables to control for the possibly non-linear relationship between schooling progression and age. In addition, we include household composition and parental characteristics. Household composition includes the number of adult males and females (aged 15-60), and the number of elderly males and females (over 60 years old) in the household. While older dependents may increase the number of chores that need to be done, especially if they are infirm or unwell, they may also be able to assist with household chores. This would be especially true for older female dependents. Since older children more often care for younger siblings, a sample of pre-school children under the age of four is included. Additionally, we include the number of school children aged 5-17 to take into account whether there is competition for the limited household resources in schooling decisions.[33]

Parental characteristics, such as parents' education and their employment status, are also important potential determinants of schooling decisions. Parents' education can be seen as a proxy variable for a child's ability. In this essay, we assume that parents' education is correlated with personal ability, which in turn is partially inherited by their children. Parents' education is reported in the data in terms of their highest educational attainment. We include three dummy variables for the highest education level attained by the father and mother, using 'no education' as the reference category. These three

[33]This is highlighted in Maitra (2003), who suggests that the presence of children of different age categories significantly reduces a female's grade attainment at higher educational levels in rural Bangladesh. No such effects are identified in the case of males.

levels of education are: primary education, secondary education and a graduate equivalent degree. Precise knowledge of how parent employment benefits children's long-term outcomes, such as educational attainment, is crucial from a policy perspective. If, for example, parent employment benefits children's education, then the government can substantially expand the employment opportunities for parents on the ground that increases in family income may also benefit children. To control for parent employment status, we include a dummy variable that indicates if at least one parent is employed.

Additionally, a dummy variable for whether the religion of a household is Muslim (the reference category being non-Muslim) is included to capture different values of education across households.[34]

Table 3.4 presents means and proportions of the variables used in the enrolment equation separately for the 1999 and 2005 survey years. On average, 49 percent of the individuals in the 1999 sample are below 20 years of age, while it is 52 percent in the 2005 sample. Conditional on completing secondary education, enrolment rates in tertiary education are higher in 2005 than in 1999. We also find that in both survey periods many males and females have completed primary school and are employed. A majority of the individuals in both samples belong to Muslim households. We also find that fathers are more educated than mothers. The higher level of education among males probably reflect the patrilineality that exists in Bangladeshi society.

As previously discussed, another channel through which female participation in higher education may be interrupted is via demographic factors. In settings such as Bangladesh, traditional customs that encourage early marriage may be responsible for low rates of female education, because women rarely continue school once they are married. In addition, traditional norms and attitudes towards the formation of families place an additional barrier on women who wish to make choices about their own lives. In Bangladesh, for example, up to 2004, the median age at first marriage is around 15 years for women aged 18-21 from 1999 to 2004 (see Table B1 in Appendix B). Furthermore, fertility preferences are still relatively high in Bangladesh, and the data shows that proportion of women who want more children has not significantly changed

[34]Another potentially important determinant is direct costs of schooling (transport costs and other expenses). Data on direct costs of schooling are not available for this study. Thus, we focus on the socio-economic characteristics of households (see also Glick and Sahn 2000 for more discussion on this issue).

from 1999 to 2004. Therefore, it can be speculated that first marriage and formation of families is likely to increase the opportunity cost of female schooling. However, other literature suggests that women's education levels are likely to be lower in a society that has higher sex ratios (Dyson and Moore 1983). Due to a longstanding tradition of son preference, sex ratio at birth (gender bias in natality) is considerably high in Asian countries (for example, India, China and South Korea), ranging between 112 and 113 per 100 female births (WDB 2012).[35] Interestingly, Bangladesh is close to this trend. The argument is that an adverse sex ratio, which reflects parents' desire for a high ratio of sons to daughters, will lead to a shortage in the supply of women and so places on them greater restriction and control. This may apply in Bangladesh, where higher female mortality rates shortly after birth through to childbearing ages (i.e. aged 15-44) is found (D'Souza and Chen 1980). Furthermore, from these female mortality rates, it may be inferred that similar gender differentials also pertain to access to education.

3.5 The Model

3.5.1 Conceptual Framework

Let us assume that individuals decide to participate in tertiary education if the benefits exceed the costs. Here, we consider only the financial benefits of enrolling in tertiary education measured in terms of premiums in returns from secondary to tertiary education. Given these costs and benefits, a utility-maximising decision-maker will enrol in tertiary education when the net benefit to enter tertiary education is positive. Let Y^*_{ijt} be the net benefit to enter tertiary education of individual i belonging to the district-area level j at time t (1999 or 2005). That is:

$$Y^*_{ijt} = \alpha M_{ijt} + R_{jt}\beta + X'_{ijt}\delta + \varepsilon_{ijt;}\ i = 1, \dots, \mathrm{N} \qquad (3.1)$$

where M_{ijt} is a dummy variable for male and X'_{ijt} is a vector of explanatory variables defined at the individual or household level that may affect tertiary enrolment. The coefficient on males (M_{ijt}) reflects the magnitude of the gender gap. R_{jt} is the wage

[35] In most human populations, more males than females are born as a result of a biological phenomenon (Waldron 1985). In an analysis based on different countries with complete and reliable data, Visaria (1971) found that in the absence of intervention, the number of male births per 100 female births ranged between 103 and 107.

premium that differs for males and females, which is calculated at the district-area level j for three age groups. ε_{ijt} are idiosyncratic errors.

As Y^*_{ijt} is unobserved, we define instead a dummy variable $Y_{ijt} = 1$ when an individual is observed to be enrolled in tertiary education conditional on completion of secondary education at the time of the survey. Persons below tertiary education, that is, those who at least completed secondary education but are not enrolled belong to the reference category:

$$Y_{ijt} = \begin{cases} 1 \ \ if \ \ Y^*_{ijt} > 0 \\ 0 \ otherwise \end{cases} \qquad i = 1, \dots, \mathrm{N} \qquad (3.2)$$

We examine the demand for tertiary education (Y_{ijt}) by estimating the following linear probability model that includes both men and women (assuming that 'prices' are equal between the two genders):[36]

$$Y_{ijt} = \alpha M_{ijt} + R_{jt}\beta + X'_{ijt}\delta + \varepsilon_{ijt;} \ i = 1, \dots, \mathrm{N} \qquad (3.3)$$

Since an important objective of this essay is to test for differences in the effect of wage premiums between genders, we consider interactions between wage premiums (R_{jt}) and a dummy variable for male (M_{ijt}) in the model and estimate the following equation:

$$Y_{ijt} = \alpha M_{ijt} + R_{jt}\beta_0 + \left(M_{ijt} * R_{jt}\right)\beta_1 + X'_{ijt}\delta + \varepsilon_{ijt;} \ i = 1, \dots, \mathrm{N} \quad (3.4)$$

A potential problem of estimating Equation (3.4) is that unobserved district-specific factors common to all individuals within a district, such as unavailability of educational facilities or a 'taste for education' may affect the quantity and quality of education. To account for this, we use a fixed-effect (FE) estimator at the district-area level. Therefore, Equation (3.4) can be rewritten as follows:

$$Y_{ijt} = \alpha M_{ijt} + R_{jt}\beta_0 + \left(M_{ijt} * R_{jt}\right)\beta_1 + X'_{ijt}\delta + \omega_j + \varepsilon_{ijt;} \ i = 1, \dots, \mathrm{N} \quad (3.5)$$

where ω_j captures the unobserved characteristics at the district-area level.

In addition, despite inclusion of ω_j, any observed correlation between wage premiums and educational choices may be driven by several factors. The most important

[36] The Linear probability model is used for ease of interpretation.

factor is the level of a district's wealth or poverty. For example, participation in tertiary education may be higher in wealthier districts. If wealthier districts also have higher returns to education, there may be a spurious positive correlation between returns to education and educational choices, driven by a district's wealth and not by returns to education. However, it is not obvious that wealthier districts will have higher returns to education. If demand for and supply of educated labour jointly determine returns to education, wealthier districts will have a higher supply of skilled labour, which *ceteris paribus* should lower returns to education. In addition, the quality and availability of educational facilities also differ markedly by districts in Bangladesh. If wealthier districts have higher returns to education and more educational institutions, we may find a spurious positive correlation between returns to education and participation in tertiary education. To address the issue of poverty or wealth in districts, we include a district's head count ratios. Unfortunately, we cannot control for the quality of educational institutions at the tertiary level, as data on the quality of educational institutions are not currently available.

Another problem we face is that returns to education could be affected by the unemployment rate of a district. In fact, higher unemployment rates reduce the returns to education. For example, higher unemployment rates of those with post-secondary schooling may reduce the incentive to invest in education because they might reduce expected wages of higher education graduates. In X_{ijt}', we include control for district youth unemployment rate because the problem of unemployment mostly affects youth labour in Bangladesh. In defining youth, we follow the definition of LFS; that is, we include individuals between the ages of 15 and 29 years. We also report unemployment rates for the youngest cohort aged 15-19. We find that young women gradually perform better than their male counterparts with respect to wage premiums from secondary to tertiary education. However, they tend to perform worse in terms of unemployment, irrespective of age group or educational qualifications, between 1999 and 2005. This evidence is presented in Table 3.5. Additionally, we find that unemployment is quite high in the youngest cohort (aged 15-19) and that this is especially the case for males. Furthermore, the relationship between unemployment and educational qualifications appears to be non-linear, irrespective of gender. More specifically, the unemployment rates are lower for tertiary education graduates aged 15-29 compared to primary and secondary education graduates aged 15-29 in 1999. However, in 2005, the reverse is true.

Aside from these major empirical problems, there are several minor empirical issues. First, since aggregate variables such as wage premiums are used to estimate individual outcomes (enrolment in tertiary education) the standard errors of the coefficients of aggregate variables will be biased upward. The FE estimator does not control for this heteroskedasticity. Standard errors are therefore made robust to heteroskedasticity and adjusted for clustering at the district-area level. Second, the data only includes children who still reside at home and excludes those who have left home either for work or marriage. It may be possible that the decision to attend advanced schools and to remain in the parental household are related. For example, if the least motivated children tend to leave home at an earlier age for work then the correlation between errors in the home-leaving and school-participation equations will lead to potential sample selection bias when schooling is estimated for home-residing children only. These biases may differ by gender if, for example, only the less motivated males who leave school and join the workforce or stay at home, while the less motivated females who leave school, choose to get married and leave home. The likelihood of leaving home either for work or marriage increases with age, and it is more likely that our estimates for educational choices are biased as we restricted the analysis to individuals aged 18-21 years old. This bias could be more pronounced for females since, by the age of 20, they are more likely to leave home for marriage in Bangladesh.[37] However, note that LFS is a nationally representative sample of individuals, and children not residing at their parents' home are enumerated at their new residence. For this reason, we have estimated the enrolment equation at the individual and not at the household level.

3.6 Estimation Results

Initially, gender-pooled linear probability models (FE-OLS) are estimated by including wage premiums of three age groups (25-35, 25-45 and 25 and above). This may provide insight into the significance and direction of wage premiums of each age group on the

[37]An estimated 75 percent of rural girls in Bangladesh marry before the age of 16 (Barkat and Majid 2003).

probability of enrolment choice at the tertiary education level. As a reference, we have also provided the estimates gained from the gender-pooled OLS model. On the whole, we find that males have a higher probability of enrolment at the tertiary level, though the gender effect is not statistically different from zero at conventional levels of significance. This is true for all regressions with one exception: when we use wage premiums of the 25-35 cohort (Table 3.6). Wage premiums from secondary to tertiary education have positively influenced the current demand for tertiary schooling of males or females, but not in all the specifications that we have considered. However, its effect is not statistically different from zero at conventional levels of significance. These findings are similar to those reported by Jacob (2002) in the United States, who finds that returns to schooling have less significant effect on college attendance rates of males and females. When we allow wage premiums to differ by gender, we find a gender difference in tertiary enrolment rates. However, differences in the effect of wage premiums between males and females are not significantly associated with enrolment in tertiary education of males and females in all regression models.

It is possible to hypothesise that, in addition to wage premiums in returns from secondary to tertiary education, educational choice might also be influenced by other potential covariates. Including additional covariates in the enrolment equation (Equation 3.5) allow us to determine whether the effects of wage premiums on the enrolment decision are altered by controlling for additional covariates. Similar to the baseline results above, we note that wage premiums do not affect the tertiary enrolment rate of males and females when potential covariates are included in the model (see Table B2 in Appendix B). It is also worth noting that the gender gap in tertiary enrolment rates is not reduced, but is no longer statistically significant.

It is naive to believe that all coefficients are the same for males and females in the enrolment function. However, this issue may be of particular relevance in the present context because of the existence of gender inequality in tertiary enrolment rates. For this reason, and to complete the earlier findings, we rerun FE-OLS for each gender group separately using wage premiums of the 25-35 cohort. Female enrolment in tertiary education is based on female wage premiums, while male enrolment is based on male wage premiums.

3.6.1 Gender

It appears that wage premiums do not have any effect on the gender gap in tertiary enrolment rates in Bangladesh; other factors may be undermining female enrolment in tertiary education. To examine this, we begin by comparing the means of important variables between the male and female samples.

These results are summarised in Table 3.7. We also present *t*-tests for gender differences in Table 3.7. With the exception of only a few cases, there is a significant difference between males and females in the sample. We find that females are relatively better off compared to males when their parents are more educated. This finding is consistent with the literature that suggests that education helps modernise individuals and cultures and decreases the role of the patriarchy (Kambhampati and Rajan 2008). Furthermore, there is a significant difference in terms of parental employment status between males and females in the sample. We further examine the robustness of these results by running a gender-pooled regression, allowing all the coefficients of variables to differ by gender. Detailed results are reported in Table B3 in Appendix B. They suggest that there is a significant difference in the factors determining male and female enrolment in tertiary education. This is confirmed by a Chow test ($F = 5.20$, $p = 0.0000$). We therefore conclude that we can best capture gender differences in tertiary enrolment rates by running separate regressions for males and females. Detailed results are presented in Table 3.8.

There are some noteworthy differences between males and females. It appears that wage premiums are a stronger determinant for male enrolment (see Columns 1, 2 and 4 of Table 3.8). These findings differ to findings by Kingdon and Theopold (2006) in India, who show that if household is credit constrained, male children in particular are withdrawn from school to take advantage of the higher returns to their (existing) levels of schooling. Our results perhaps suggest that higher wage premiums to males who possess tertiary education in Bangladesh may lead parents to invest in a child's education due to higher expected future earnings. However, in the female sample, the effect of the wage premium is positive, although it is not statistically different from zero at conventional levels of significance except in Column 2 of Table 3.8. Likewise for males, the positive relationship (although insignificant) between wage premiums and the tertiary enrolment rates in the female sample suggests that the female opportunity cost of education is smaller. This is not impossible in the context of Bangladesh, as we find widespread graduate unemployment among females between 1999 and 2005. An alternative explanation may be that female choices are limited due to the gender division of labour in Bangladesh. Females are less likely to be wage employees and are more likely to engage in household activities. There is some support in our data for this argument. For example, in the age group 18-21, approximately 75 percent of males and only 17 percent of females are employed in waged work.

The effect of other covariates reveals some useful insight into the significance and direction of the explanatory variables on the probability of enrolment decision at the tertiary level across gender groups. The age variables suggest that the propensity of enrolment decreases with age for both males and females. In case of males, the results perhaps suggest that older males are most likely to leave school as they grow up due to household resource constraints or labour market opportunities. In regard to females, one possible explanation is that females marry very young in Bangladesh and so are less likely to continue their education because this raises the opportunity cost.

We find that the probability of enrolment at the tertiary level is also affected by household composition. Notably, the presence of a number of school children aged 5-17 exerts a negative effect on male enrolment in tertiary education. This finding may imply that large numbers of school-aged children demand that more resources be invested into their education, which in turn forces older male children to be employed, due to parental resource constraints, to make schooling possible for themselves and for their siblings. This may have a negative effect on their schooling outcomes. The household composition effects, however, are considerably weaker for females. Most of the variables are not statistically significant at conventional levels, except for the number of adult females in the household. We find that additional adult women aged 15-60 in the household may reduce the opportunity cost of a woman’s time by providing substitutes for household work or through economies of scale in the household production, thereby raising the likelihood of enrolment in tertiary education among females.

Another finding is parental education has a positive and statistically significant effect on tertiary enrolment rates of males and females. The effects are quite strong and it is worth noting that parental education has a stronger effect on the tertiary enrolment rate of females compared with males. This is consistent with the literature that suggests that educated parents are more likely to value education and so are less likely to constrain females at home, sending them to school instead (Kambhampati and Rajan 2008). This result has important policy implications in regard to reducing gender inequality in education at the tertiary level.

Columns 4 and 8 of Table 3.8 include controls for demographic factors such as median age at first marriage, fertility preferences and sex ratio.[38] While age at first marriage and fertility preferences may be endogenous to the schooling decision, the results provide a useful description of the enrolment process at the tertiary level.[39] We note that fertility preferences exert a negative influence on the probability of enrolment in tertiary education for males and females. However, this result is only significant in the female sample. This finding has important implications for female enrolment, in both quantitative and qualitative terms. Moreover, the age at first marriage variable is highly important in explaining the probability of enrolment for males. As mentioned earlier, another point of interest is whether gender preference proxied by sex ratio at birth explains low female enrolment in tertiary education in Bangladesh. We find that sex ratio at birth is negatively correlated with tertiary enrolment rates of males and females, although the effect is only significant in the male sample. We therefore failed to confirm any parental discrimination against women in educational investment in Bangladesh.

3.7 Robustness

3.7.1 Using Median Wage Premiums

In this sub-section, we extend our analysis to check the robustness of our results by using median wage premiums, following Jacob (2002). We re-estimate baseline regressions, both gender-pooled OLS and FE-OLS, using wage premiums of three age groups: 25-35, 25-45 and 25 and above. The results are presented in Table 3.9. In

[38] We use the median age at first marriage of the sample of population aged over 21 years in a district, while ‘fertility preferences’ imply desire for more children by both males and females aged over 21 years in a district. ‘Sex ratio at birth’ is defined as the ratio of male to female children born in a district in a specific period.

[39] For some individuals ‘age at first marriage’ and ‘fertility preferences’ may be partly determined by educational decisions. This is more common for females as they are most likely to be guided by marriage practices and by the gender division of labour, particularly in the South Asian region. We are unable to deal with any potential endogeneity of these variables because the data do not yield instruments of acceptable quality for these demographic factors.

general, we find that there is a significant gap in enrolment in tertiary education between males and females and the gap increases in magnitude when we use wage premiums of younger (aged 25-45) and the oldest (25 and above) cohorts than the youngest cohort (aged 25-35). However, the effect of the wage premium is not consistent with our previous results. We find evidence that wage premiums from secondary to tertiary education do affect the tertiary enrolment rate, and this is especially the case when we use wage premiums of the 25-35 cohort under the FE estimation. We again note that differences in the effect of wage premiums between males and females are not significantly associated with enrolment in tertiary education of males and females in all regression models.

We further examine the robustness of our results by running a separate regression by gender, using median wage premiums of the 25-35 cohort. In general, we find that wage premiums are a stronger determinant for males' enrolment in tertiary education but not for females, which reiterates our findings from Table 3.8. Detailed results are reported in Table B4 in Appendix B.

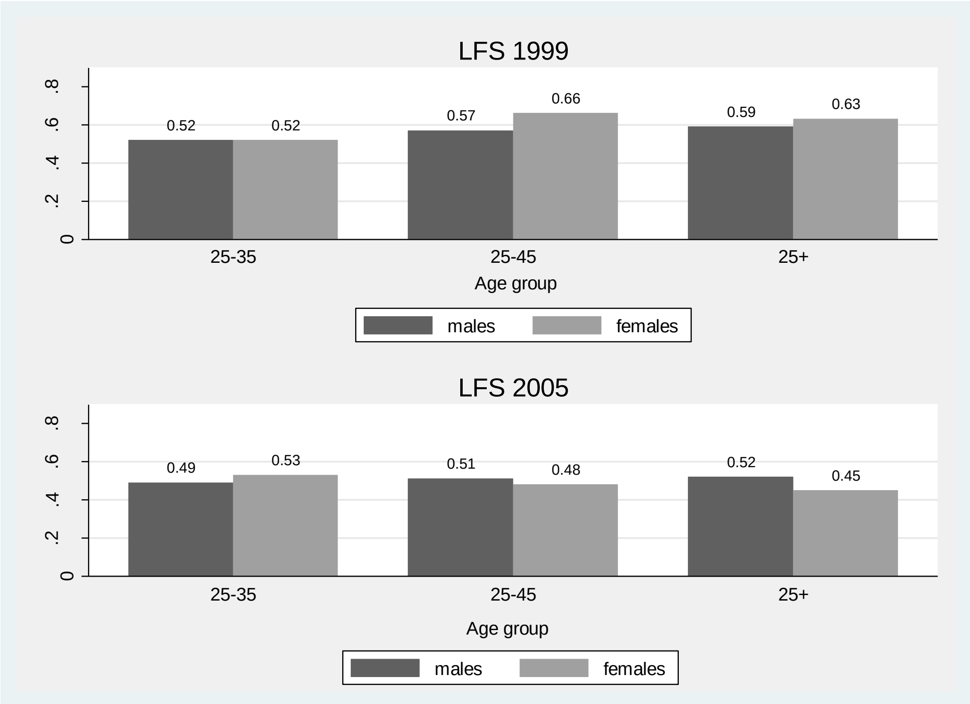
3.8 Summary and Conclusion

In this essay, we examine whether wage premiums from secondary to tertiary education affect the gender gap in tertiary enrolment rates in Bangladesh. The results uncovered significant differences in tertiary enrolment rates between males and females when we control for mean wage premiums of the 25-35 cohort. This result appears to be reasonably robust when we control for median wage premiums of the 25-35 cohort. While wage premiums do not significantly affect the tertiary enrolment rate in gender-pooled regressions, it appears to be a stronger determinant in our separate regression models. Overall, our results suggest that wage premiums positively affect tertiary enrolment rates of males and females, although the effect is only significant in the male sample.

We also find that males and females substantially benefit when their parents are educated, even after controlling for wage premiums. Furthermore, we note that fertility preferences are negatively associated with enrolment and are higher among females, even after controlling for labour market returns and parents' education. These results suggest that females will receive lower schooling in Bangladesh, because considerations involved in schooling decisions at the tertiary level are more likely to be guided by demographic factors.

Finally, the findings presented suggest that wage premiums are not a significant causal factor underlying the low female enrolment in tertiary education in Bangladesh. Therefore, any labour market reforms (for example, the employment of women with tertiary education in more rewarding sectors) to increase returns to tertiary education for females will be unlikely to increase their enrolment in tertiary education. On the other hand, we find that parental characteristics, such as parents’ education, are highly important for females with regard to their access to higher education, even after controlling for wage premiums. Therefore, the alternative policy may be to raise the economic benefits of education of adult members of the household, which could greatly improve human capital investment among young females in households. It may also be important to initiate innovative policies, such as night school and monetary incentives (for example, post-graduate scholarships and loans to qualified and interested women), if low female enrolment is due to higher costs (opportunity costs and/or direct costs).

Figure 3.1: Wage Premiums of Males and Females, by Age Group, 1999-2005



Source: Data are from LFS 1999 and LFS 2005.

Table 3.1: Enrolment Rates in Tertiary Education among High School Graduates and Raw Gender Difference in Enrolment in Bangladesh, 1999-2005

Age	Males (1)	Females (2)	Gender Difference (3)=(1)-(2)	Percentage of Female Disadvantage (3)/(1)*100
1999				
18	43.25	28.36	14.89	34.43
19	40.0	31.03	8.97	22.43
20	40.34	21.37	18.97	47.03
21	41.49	24.47	17.02	41.02
2005				
18	49.2	41.6	7.60	15.45
19	42.6	41.52	1.08	2.54
20	47.22	30.96	16.26	34.43
21	49.13	30.23	18.90	38.47

Source: Data are from LFS 1999 and LFS 2005.

Table 3.2: Mean of (Log) Hourly Wages of Men and Women, by Age Group and Education Levels, Bangladesh, 1999

	Secondary				Tertiary			
Age group (years)	Males (1)	Females (2)	Gap (M-F) (3) = (1)-(2)	*t* - test (M - F) (4)	Males (5)	Females (6)	Gap (M-F) (7)= (5)-(6)	*t* - test (M - F) (8)
25 - 35	2.5716	2.3634	0.2082	2.80**	3.0937	2.8851	0.2086	1.90*
	(0.5890)	(0.8606)			(0.5922)	(0.7171)		
25 - 45	2.6929	2.4010	0.2919	4.84***	3.2644	3.0574	0.2070	2.36**
	(0.5929)	(0.8818)			(0.6244)	(0.6721)		
25+	2.7528	2.4304	0.3224	5.54***	3.3445	3.0566	0.2879	3.34***
	(0.6210)	(0.8724)			(0.6553)	(0.6608)		

Notes: Data are from LFS 1999. (log) hourly wages for the secondary education levels are calculated for those individuals who have completed Grade 10 or more and are in paid employment while (log) hourly wages for the tertiary education levels are calculated for those individuals who have completed Grade 13 or more (bachelor's or higher) or any other technical programs and are in paid employment. *** p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 3.3: Mean of (Log) Hourly Wages of Men and Women, by Age Group and Education Levels, Bangladesh, 2005

	Secondary				Tertiary			
Age group (years)	Males (1)	Females (2)	Gap (M-F) (3) = (1)-(2)	*t* - test (M - F) (4)	Males (5)	Females (6)	Gap (M-F) (7)= (5)-(6)	*t* - test (M - F) (8)
25 - 35	2.7675	2.6538	0.1137	2.44**	3.2592	3.1821	0.0771	1.39
	(0.6254)	(0.7584)			(0.6423)	(0.6366)		
25 - 45	2.8790	2.8174	0.0616	1.68*	3.3921	3.2952	0.0969	2.15**
	(0.6241)	(0.7974)			(0.6595)	(0.6454)		
25+	2.9606	2.8807	0.0799	2.35**	3.4854	3.3263	0.1591	3.68***
	(0.6445)	(0.7869)			(0.6829)	(0.6578)		

Notes: Data are from LFS 2005. (log) hourly wages for the secondary education levels are calculated for those individuals who have completed Grade 10 or more and are in paid employment while (log) hourly wages for the tertiary education levels are calculated for those individuals who have completed Grade 13 or more (bachelor's or higher) or any other technical programs and are in paid employment. *** p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 3.4: Descriptive Statistics of Variables Used in Enrolment Equation

	LFS 1999					LFS 2005				
Variables	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Individual/Household covariates										
Age 18	3,389	0.3588	0.4797	0	1	13,000	0.3168	0.4652	0	1
Age 19	3,389	0.1295	0.3358	0	1	13,000	0.2032	0.4024	0	1
Age 20	3,389	0.4087	0.4917	0	1	13,000	0.2954	0.4562	0	1
Age 21	3,389	0.1030	0.3040	0	1	13,000	0.1846	0.3880	0	1
Male (1 = yes)	3,389	0.4647	0.4988	0	1	13,000	0.5033	0.5000	0	1
Muslim (1 = yes)	3,389	0.9150	0.2789	0	1	13,000	0.8718	0.3344	0	1
Married (1 = yes)	3,389	0.4001	0.4900	0	1	12,998	0.3389	0.4734	0	1
Education experience, aged 18-21										
less than primary education (Grade 1 - 5) (currently attending = 1)	142	3.2535	1.1388	1	5	2,265	4.4450	0.8181	1	5
Completed primary education (Grade 5) and employed	2,336	2.0167	0.9964	1	5	9,348	2.5816	1.2533	1	5
Completed secondary education (Grade 10 or more) and employed	327	5.3303	0.9371	6	11	292	6.3630	0.9480	6	11
Completed secondary education and enrol in tertiary education (bachelor's or equivalent) (1= yes)	584	5.5377	1.1590	6	11	1,095	6.2201	0.7142	6	11
Number of children 0-4	3,389	0.0608	0.2778	0	2	13,000	0.0658	0.2754	0	3
Number of school children 5-17	3,389	0.9675	1.3785	0	7	13,000	0.8840	1.2149	0	7
Number of males 15 -60	3,389	2.1596	1.1729	0	7	13,000	2.1648	1.1110	0	10
Number of females 15-60	3,389	1.8404	0.9777	0	6	13,000	1.8762	0.9460	0	9
Number of males over 60	3,389	0.1526	0.3629	0	2	13,000	0.1959	0.3983	0	2
Number of females over 60	3,389	0.0670	0.2524	0	2	13,000	0.1046	0.3093	0	2
Father has no education (1 = yes)	3,389	0.1688	0.3746	0	1	13,000	0.2237	0.4167	0	1
Father has primary education (1 if father has completed Grade 5)	3,389	0.1089	0.3115	0	1	13,000	0.1232	0.3286	0	1
Father has secondary education (1 if father has completed Grade 10 or more)	3,389	0.1608	0.3674	0	1	13,000	0.1861	0.3892	0	1
Father has graduate equivalent degree (1 if father has completed Grade 13 or more)	3,389	0.0531	0.2243	0	1	13,000	0.0328	0.1782	0	1
Mother has no education (1 = yes)	3,389	0.2726	0.4454	0	1	13,000	0.3212	0.4669	0	1
Mother has primary education (1 if mother has completed Grade 5)	3,389	0.1139	0.3177	0	1	13,000	0.1495	0.3566	0	1
Mother has secondary education (1 if mother has completed Grade 10 or more)	3,389	0.1328	0.3394	0	1	13,000	0.1405	0.3475	0	1
Mother has graduate equivalent degree (1 if mother has completed Grade 13 or more)	3,389	0.0083	0.0905	0	1	13,000	0.0056	0.0747	0	1
At least one parent employed (1 = yes)	3,389	0.2496	0.4329	0	1	13,000	0.4131	0.4924	0	1

Notes: Individuals aged 18-21 only. Data are from LFS 1999 and LFS 2005. Std. Dev. is standard deviation.

Table 3.5: Youth Unemployment Rates, by Age Group, Education Levels and Gender, Bangladesh 1999-2005

		All	Primary	Secondary	Tertiary
15-19	1999				
	All	53.4	53.3	53.9	
	Males	16.1	16.5	16.7	
	Females	35.0	34.4	34.0	
	2005				
	All	52.5	52.2	52.9	
	Males	16.2	17.7	15.6	
	Females	32.5	27.1	35.4	
15-29	1999				
	All	54.1	54.4	53.9	53.3
	Males	13.9	13.8	15.6	18.7
	Females	40.9	41.9	36.1	26.8
	2005				
	All	51.8	51.9	52.2	53.2
	Males	13.6	12.7	13.8	18.1
	Females	38.5	41.2	38.6	28.3

Notes: Data are from LFS 1999 and LFS 2005. In LFS, youth refers to persons between the ages of 15 and 29 years. The youth unemployment rate is defined as the number of unemployed youths in a relevant age group divided by the number of youth labour force (employed and unemployed) of the same age group.

Table 3.6: Enrolment in Tertiary Education using Mean Wage Premiums for Different Age Groups (Sample: 18-21 years in both survey years)

Variables	(1) OLS	(2) FE-OLS	(3) OLS	(4) FE-OLS	(5) OLS	(6) FE-OLS
Male	0.0551*	0.0528*	0.0503	0.0438	0.0476	0.0526
	(0.0313)	(0.0287)	(0.0341)	(0.0316)	(0.0352)	(0.0332)
Premium 25-35	-0.0569	0.1000				
	(0.0485)	(0.0660)				
Male*premium 25-35	0.0166	0.0374				
	(0.0542)	(0.0495)				
Premium 25-45			-0.0328	0.0315		
			(0.0535)	(0.0895)		
Male*premium 25-45			0.0173	0.0421		
			(0.0561)	(0.0547)		
Premium 25+					-0.0275	-0.0329
					(0.0593)	(0.0858)
Male*premium 25+					0.0104	0.0136
					(0.0608)	(0.0586)
Year	0.1307***		0.1451***		0.1461***	
	(0.0280)		(0.0284)		(0.0279)	
Constant	0.6827***	0.6933***	0.6599***	0.7202***	0.6590***	0.7511***
	(0.0378)	(0.0320)	(0.0415)	(0.0455)	(0.0437)	(0.0427)
N	1,927	1,927	2,047	2,047	2,086	2,086
R-squared	0.0287	0.0114	0.0327	0.0068	0.0326	0.0051

Notes: Data are from pooled LFS 1999 and LFS 2005. FE-OLS is fixed-effect OLS. Figures in parentheses under FE-OLS are robust standard errors, corrected for clustering at the district-area level. Sample sizes in different specifications are reduced because of missing data in wages in different districts. *** p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 3.7: Gender Differences in Selected Variables
(Sample: 18-21 years in both survey years)

Variables	Sample	N	Mean	Std. Dev.	*t*-test for equality of means *t*-statistic
Age 19	Males	1,078	0.1818	0.3858	-1.91*
	Females	849	0.2167	0.4122	
Age 20	Males	1,078	0.3432	0.4750	0.73
	Females	849	0.3274	0.4696	
Age 21	Males	1,078	0.2226	0.4162	1.32
	Females	849	0.1979	0.3986	
Mother has primary education	Males	1,078	0.1892	0.3919	1.73*
	Females	849	0.1590	0.3659	
Mother has secondary education	Males	1,078	0.4499	0.4977	0.98
	Females	849	0.4276	0.495	
Mother has graduate equivalent degree	Males	1,078	0.0362	0.1868	-0.57
	Females	849	0.0412	0.1989	
Father has primary education	Males	1,078	0.1187	0.3236	2.88**
	Females	849	0.0789	0.2697	
Father has secondary education	Males	1,078	0.4332	0.4957	4.76***
	Females	849	0.3274	0.4695	
Father has graduate equivalent degree	Males	1,078	0.1736	0.3788	-1.44
	Females	849	0.1991	0.3995	
At least one parent employed	Males	1,078	0.5751	0.4966	3.01***
	Females	849	0.5065	0.5003	

Notes: Data are from pooled LFS 1999 and LFS 2005. Std. Dev. is standard deviation. ***p<0.01, ** p<0.05, * p<0.1. p denotes p-value.

Table 3.8: Enrolment in Tertiary Education by Gender: Linear Probability Models using Mean Wage Premiums
(Sample: 18-21 years in both survey years)

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
Premium 25-35	0.1604**	0.1416*	0.1352	0.1458*	0.1102	0.1016*	0.1047	0.0918
	(0.0708)	(0.0825)	(0.0831)	(0.0780)	(0.0694)	(0.0596)	(0.0702)	(0.0676)
Age 19	-0.0445	-0.0499	-0.0474	-0.0515	-0.0396	-0.0375	-0.0419	-0.0496
	(0.0421)	(0.0390)	(0.0392)	(0.0392)	(0.0344)	(0.0319)	(0.0321)	(0.0337)
Age 20	-0.0249	-0.0336	-0.0316	-0.0435	-0.1985***	-0.1366***	-0.1345***	-0.1403***
	(0.0265)	(0.0299)	(0.0303)	(0.0293)	(0.0370)	(0.0354)	(0.0342)	(0.0342)
Age 21	-0.0285	-0.0445	-0.0426	-0.0578**	-0.1019**	-0.0670	-0.0699	-0.0864
	(0.0312)	(0.0305)	(0.0313)	(0.0287)	(0.0469)	(0.0500)	(0.0502)	(0.0526)
Muslim	0.0312	0.0607	0.0627	0.0664	-0.0280	-0.0041	0.0038	-0.0004
	(0.0441)	(0.0427)	(0.0420)	(0.0423)	(0.0537)	(0.0471)	(0.0462)	(0.0439)
Number of children 0-4		-0.0368	-0.0366	-0.0397		-0.0857	-0.0976	-0.1010
		(0.0651)	(0.0650)	(0.0630)		(0.0766)	(0.0752)	(0.0766)
Number of school children 5-17		-0.0284**	-0.0286**	-0.0268**		-0.0102	-0.0071	-0.0030
		(0.0126)	(0.0127)	(0.0127)		(0.0128)	(0.0120)	(0.0130)
Number of males 15-60		-0.0047	-0.0064	-0.0047		-0.0104	-0.0095	-0.0077
		(0.0155)	(0.0153)	(0.0154)		(0.0147)	(0.0143)	(0.0141)
Number of females 15-60		-0.0132	-0.0137	-0.0143		0.0374*	0.0405**	0.0401**
		(0.0177)	(0.0177)	(0.0183)		(0.0201)	(0.0196)	(0.0194)
Number of males over 60		0.0391	0.0361	0.0331		-0.0082	-0.0151	-0.0149
		(0.0372)	(0.0376)	(0.0374)		(0.0474)	(0.0469)	(0.0464)
Number of females over 60		0.0518	0.0492	0.0383		0.0102	0.0075	-0.0057
		(0.0426)	(0.0429)	(0.0424)		(0.0509)	(0.0506)	(0.0533)
Mother has primary education		0.0150	0.0143	0.0146		0.1664***	0.1673***	0.1611***
		(0.0325)	(0.0324)	(0.0324)		(0.0467)	(0.0442)	(0.0449)
Mother has secondary education		0.1086***	0.1031***	0.0987***		0.2124***	0.2097***	0.2096***
		(0.0367)	(0.0371)	(0.0369)		(0.0439)	(0.0424)	(0.0429)
Mother has tertiary education		0.0940**	0.0921**	0.0736*		0.2187***	0.2158***	0.2152***
		(0.0410)	(0.0411)	(0.0395)		(0.0711)	(0.0723)	(0.0731)

Continued on next page

Table 3.8 (continued): Enrolment in Tertiary Education by Gender: Linear Probability Models using Mean Wage Premiums (Sample: 18-21 years in both survey years)

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
Father has primary education		0.0388	0.0396	0.0301		0.2316***	0.2347***	0.2280***
		(0.0414)	(0.0416)	(0.0433)		(0.0685)	(0.0670)	(0.0692)
Father has secondary education		0.0653*	0.0668*	0.0562		0.1147**	0.1146**	0.1151**
		(0.0361)	(0.0360)	(0.0360)		(0.0467)	(0.0458)	(0.0475)
Father has tertiary education		0.1292***	0.1298***	0.1327***		0.1734***	0.1719***	0.1660***
		(0.0378)	(0.0379)	(0.0359)		(0.0448)	(0.0453)	(0.0475)
At least one parent employed		0.0330	0.0368	0.0321		0.0659	0.0564	0.0459
		(0.0250)	(0.0242)	(0.0234)		(0.0411)	(0.0394)	(0.0390)
Mean head count ratios in district			0.0010	0.0038			0.0060	0.0088*
			(0.0025)	(0.0034)			(0.0039)	(0.0046)
Mean youth unemployment rate in district			8.4781*	5.5280			-15.0937***	-8.4739*
			(4.3959)	(4.4408)			(4.8555)	(4.8788)
Mean of fertility preference in district				-0.7890				-1.1232***
(aged over 21 years)				(0.5139)				(0.3856)
Median age at first marriage in district				0.0316**				0.0043
(aged over 21 years)				(0.0155)				(0.0223)
Mean sex ratio at birth in district				-0.0036**				-0.0005
				(0.0014)				(0.0014)
Constant	0.7274***	0.6453***	-1.9137	-1.4772	0.8113***	0.4495***	11.9167***	6.7282*
	(0.0614)	(0.0752)	(1.3022)	(1.2125)	(0.0534)	(0.0853)	(3.6885)	(3.7278)
N	1,078	1,078	1,078	1,078	849	849	849	849
R-squared	0.0093	0.0633	0.0663	0.0805	0.0386	0.2158	0.2267	0.2365

Notes: Data are from pooled LFS 1999 and LFS 2005. Analyses of Columns 4 and 8 are derived from pooled LFS 1999 and DHS 1999 and LFS 2005 and DHS 2004. FE-OLS is fixed-effect OLS. Figures in parentheses are robust standard errors, corrected for clustering at the district-area level. ***p<0.01, ** p<0.05, * p<0.1.
p denotes p-value.

Table 3.9: Enrolment in Tertiary Education using Median Wage Premiums for Different Age Groups (Sample: 18-21 years in both survey years)

Variables	(1) OLS	(2) FE-OLS	(3) OLS	(4) FE-OLS	(5) OLS	(6) FE-OLS
Male	0.0642**	0.0647**	0.0665*	0.0704**	0.0674*	0.0697*
	(0.0303)	(0.0288)	(0.0356)	(0.0343)	(0.0369)	(0.0356)
Premium 25-35	-0.0181	0.1353*				
	(0.0472)	(0.0746)				
Male*premium 25-35	-0.0037	0.0131				
	(0.0553)	(0.0515)				
Premium 25-45			0.0146	0.0809		
			(0.0474)	(0.0968)		
Male*premium 25-45			-0.0171	-0.0112		
			(0.0574)	(0.0581)		
Premium 25+					0.0323	0.0590
					(0.0532)	(0.0902)
Male*premium 25+					-0.0303	-0.0230
					(0.0581)	(0.0575)
Year	0.1290***		0.1457***		0.1488***	
	(0.0279)		(0.0278)		(0.0272)	
Constant	0.6653***	0.6802***	0.6369***	0.6967***	0.6286***	0.7073***
	(0.0365)	(0.0337)	(0.0350)	(0.0484)	(0.0427)	(0.0472)
N	1,927	1,927	2,047	2,047	2,086	2,086
R-squared	0.0272	0.0126	0.0324	0.0070	0.0326	0.0054

Notes: Data are from pooled LFS 1999 and LFS 2005. FE-OLS is fixed-effect OLS. Figures in parentheses under FE-OLS are robust standard errors, corrected for clustering at the district-area level. Sample sizes in different specifications are reduced because of missing data in wages in different districts. ***p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Chapter 4

Trade-off between Child Labour and Schooling in Bangladesh: the Role of Parents' Education

4.1 Introduction

Child labour is not a new phenomenon in a low-income economy where children are an economic resource for poor parents. Poverty is considered to be the major driving force for child labour in low-income countries (Ersado 2005; Maitra and Ray 2002). In under-developed economies, where labour markets are usually quite imperfect, poorer households are more likely to send their children to work in order to escape extreme poverty. This phenomenon is what Basu and Van (1998) call the Luxury Axiom.[40] Others argue that factors such as 'bequest constraint'[41] of parents play a significant role in sending children to work (Baland and Robinson 1998). On the other hand, studies also exist, which suggest that household poverty does not necessarily increase child labour. Swaminathan (1998), in fact, finds a weak relationship between the incidence of child labour and the incidence of poverty in India (see also Ray 2000a). A similar consensus is also found in other studies that show an inverted-U relationship between household wealth (for example, land) and child labour (Basu et al. 2010; Bhalotra and Heady 2003).[42]

[40] The phrase 'Luxury Axiom' refers to a family that will send a child to the labour market only if the family's income from non-child labour sources drops very low.

[41] The phrase 'bequest constraint' refers to the parents' understanding of possible future financial benefit that might impact on their present decisions concerning their child schooling.

[42]It is suggested that 'wealth effects' tend to reduce child labour but also improving the credit-worthiness of a household and hence mitigating the adverse effects of imperfect credit and insurance markets. At the same time, if it is difficult to hire productive labour (or to buy and sell assets), the children of asset-rich households will be more likely to work on the family farm, and less likely to attend school.

A key concern about child labour is whether the work activities of children hamper their school performance. This is an important question from a policy perspective because when working children attend school their educational attainment is likely to be lower than full-time students. In the long run, this will eventually reduce the accumulation of human capital. In addition to the negative effects of work, child labour may provide the household with sufficient income to keep children in school. Indeed, many studies cited in the literature reviews by Basu (1999) and Edmonds (2007) find zero or a positive effect of child labour on school enrolment and educational attainment. Therefore, the effect of child labour on schooling is ambiguous. However, it is generally accepted that the negative effects are more likely to outperform the positive ones.

The principal aim of this essay is to investigate whether the number of working hours adversely affects schooling outcomes of children. *Ceteris paribus*, child workers who work longer hours will have little time for school attendance and studying. Exhaustion from longer hours of work could also prevent children from being attentive inside and outside the classroom, which has implications for educational performance.[43] In this setting, we contend that parental attitudes and preferences may affect the work-schooling trade-off, and we use the educational levels of both mothers and fathers as an indicator of attitudes and parental preferences.

Additionally, this essay addresses sample selection bias. As we focus our analysis on child's working hours, this may introduce the well known problem of sample selection bias. Hours of work are only available for working children; children who work different hours might have unobserved characteristics that could be correlated with the unobservables in the outcome equation (in this case, school), causing our estimates to be biased and inconsistent.

Specifically, in this essay, we examine the following questions:

a. Is there any trade-off between work and schooling?
b. How does the parents' education affect the work-schooling trade-off?
c. How are the results affected if we explicitly take sample selection into child employment into account?

[43] Of course, a number of other school-related factors may also have a role to play.

We employ an instrumental variable (IV) estimation strategy because we presume the endogeneity of hours worked of children in the structural equation of schooling outcomes. A non-parametric method is also employed for further exploration of the relationship between child labour hours and child schooling. We approach the problem of selectivity bias by adopting the sample selection model. A double-hurdle model is also employed for comparison with the sample selection model.

To preview our results, we first note that our parametric results are generally consistent with the literature. We find that the working hours adversely affect child schooling from the very first hour of work, but the marginal impact of child labour hours weakens when working hours increase. However, note that when we control for different degrees of polynomials in working hours we find that hours worked by children decrease their schooling. The results from a non-parametric approach are consistent with the parametric findings when we control for different degrees of polynomials in working hours. We also find that parents' education plays an important role in influencing the child's work-schooling trade-off. While the schooling of both mothers and fathers shifts the work-schooling trade-off in favour of education, the mother's educational attainment has stronger marginal effects on the work-schooling trade-off than the father's education. We note that the mother's and father's education (as measured by the highest grade attained) shows a significant preference for educating a female child. The same incentive effect is not found for a male child, which suggests that male children are more likely to work. Finally, a sample consisting of only working children introduces a significant bias of IV estimates of the work-schooling trade-off.

4.2 Literature Review

4.2.1 The Child Labour-Schooling Trade-off

Following the seminal paper by Basu and Van (1998), a number of studies explicitly consider the trade-off between child labour and schooling (Basu 1999; Basu and Tzannatos 2003; Edmonds 2007; Emerson and Souza 2003; Udry 2006). There are also studies, which have attempted to test Basu and Van's (1998) model. For example, Ray (2000a) applied Basu and Van's (1998) model to Peruvian and Pakistani data. His results mostly confirm to those of Basu and Van in the context of Peru but not in Pakistan.

Falling adult male wages in Peru lead to increased participation of children in the labour market. The author also shows that Peruvian working children tend to combine employment with schooling while Pakistani children, especially older girls, drop out of school completely to participate in the labour market.

Other empirical research has demonstrated that work and school are competing activities that might also be combined, and offer valuable insights on the nature of the child labour-schooling trade-off (Ganglmair 2006; Jensen and Nielsen 1997; Pal 2004; Rammohan 2012; Sedlacek et al. 2009). Heady (2003), however, derives slightly different results for Ghana and concludes that child labour does not affect school attendance but substantially interferes with the quality of schooling with respect to children's reading and mathematics ability. A similar result is obtained by Parikh and Sadoulet (2005), who argue that child labour and school are 'not necessarily incompatible' but understand that due to children's 'dual commitment' at work and in school their educational progress may be impaired. By contrast, Ray (2000a), Maitra and Ray (2002) and Ersado (2005) argue that children tend to combine school and labour without detrimental effects on their school performance, and it might well be the case that their earnings from work make their schooling possible. Beegle et al. (2009), however, find that educational attainment is lower for children who combine work and schooling. Most of these analyses are based on children's participation rate rather than hours worked. Exceptions are Boozer and Suri (2001), Rossati and Rossi (2003), Ray and Lancaster (2005), Han (2008), and Kana et al. (2010). All these studies, except Kana et al. (2010), find a negative association between hours worked and educational attainment.

The issue of how school performance is affected by child labour has been further examined by Patrinos and Psacharopoulas (1995). Using Paraguayan data, they conclude that children can both attend school and work without affecting their schooling progress (i.e. whether or not children have to repeat grades). In a later paper on Peru, Patrinos and Psacharopoulos (1997) find that child labour significantly influences children's 'age grade distortions,' and hence hampers their educational progress. Similarly, in a study on Tanzania, Akabayashi and Psacharopoulos (1999) conclude that children's reading competence decreases with child labour hours. Furthermore, Gunnarsson et al. (2006), using a rich dataset from eleven developing countries, find that child labour significantly reduces test scores in every country. Other research focusing on school performance arrive at similar conclusions, including Goulart and Bedi (2008) in Portugal and Zabaleta (2011) in Nicaragua.

4.2.2 Evidence from Bangladesh

While many studies find evidence of the trade-off between child labour and schooling, only a limited number of studies have investigated this, especially in the context of Bangladesh (Amin et al. 2006; Arends-Kuenning and Amin 2004; Ravallion and Wodon 2000; Shafiq 2007). Existing studies on child labour have mainly explored whether child work is a deterrent or a complement to school attendance and/or enrolments (see, for example, Amin et al. 2004). Apart from the literature on the potentially negative effect of child labour on school attendance and performance, there has been a quite different strand of literature, which documents that children tend to combine schooling with work (Khanam 2008). This is specially the case for rural girls.

A study by Ridao-Cano (2001) is the only paper that examines the intra-household bargaining power on child labour and child schooling. Using data from rural Bangladesh, the author analyses the determinants of child labour (for example, participation in farm work) and schooling (participation in school) and concludes that mothers have a higher preference for child schooling than fathers. This difference is mainly revealed through the relative bargaining power of mothers proxied by access to credit from group-based credit programs. On the other hand, the mother's and the father's access to credit has no significant effect on child labour. In a different paper on rural Bangladesh, Canals-Cerdá and Ridao-Cano (2004) analyse the long-term consequences of child labour on school performance and find that child labour significantly reduces schooling progress and the intensity of its impact is more severe the earlier a child begins to work.

In a recent study of child labour, Islam and Choe (2011) analyse the impact of household participation in group-based credit programs on children's education and child labour. The authors conclude that household access to credit increases child labour and reduces school enrolment.

4.2.3 Contribution of this Study

This essay contributes to the literature in the following ways. First, we focus on the trade-off between child labour hours and schooling, and explore how parents' educational levels affect the work-schooling trade-off between genders. Though there has been much discussion on the role of parents' education on child labour and child schooling (see, for example, Coulombe 1998; Deb and Rosati 2002; Ganglmair 2006; Grootaert 1999; Han 2008; Shafiq 2007), there is a little empirical evidence on their impact on the trade-off between child labour hours and child schooling. Arguably, if parents' education positively influences parental preferences for children's education, then an increase in parents' educational levels should result in more schooling and less child labour, even in poor households. Alternatively, parents' education may increase the efficiency or effectiveness of the time spent interacting with children (for example, directly helping with schoolwork), and more educated parents may thus forgo some time spent working to make greater time investments in their children's human capital, which in turn reduces child's working hours. However, it is often posited that more educated parents in poor households without access to credit may face a trade-off between education and current consumption; this does not necessarily mean that children of more educated parents are more likely to go to school. Indeed, depending on circumstances, caring parents might insist on their children working, and use the additional income to improve their children's nutrition rather than increasing expenditure on education. However, another possibility is that even during income shocks (for example, unemployment and natural disasters), a household with educated parents is less likely to pull a child out of school, or to send a child to work, or both, because educated parents are more likely to have safety nets (for example, insurance).

Second, in a model of schooling, child labour hours are endogenous in the sense that they may be jointly determined with schooling. If so, treating child labour hours as exogenous could result in biased estimates. A small number of studies (Beegle et al. 2009; Bhalotra 2007; Han 2008; Ray and Lancaster 2005) have tried to control for endogenous child labour, mainly because of the unavailability of valid instruments in their dataset. We address the potential endogeneity problem by adopting an IV approach. These instruments are unique and have not been used previously.

Third, we implement a non-parametric approach for further investigation on the relationship between child labour hours and child schooling. To the best of our knowledge this is the first attempt to use a non-parametric approach to investigate the trade-off between the number of working hours and schooling.

Fourth, we explore the selectivity bias, which has received very little attention in most studies that estimate the effect of the number of working hours on child schooling outcomes (Ravallion and Wodon 2000; Ray and Lancaster 2005). We address the selectivity problem by using the sample selection model. Moreover, a double-hurdle model is employed for a robustness check.

Finally, the empirical analysis is carried out by utilising the Bangladesh National Child Labour Survey data for 2002-2003. This dataset has not been used in the previous literature to investigate the trade-off between child labour hours and child schooling in Bangladesh.

4.3 Bangladesh Perspective

This essay's focus upon child labour in South Asia is motivated by the basic fact that the majority of Asian child workers come from South Asia (Ray 2004). However, the child labour participation rate in South Asia lags behind that of Africa. As regards Bangladesh, a National Child Labour Survey (NCLS) conducted in 2002-2003 by the Bangladesh Bureau of Statistics (BBS) under the auspices of the ILO-sponsored International Program on the Elimination of Child Labour (IPEC), found that approximately 5 million (14 percent) of the total 35 million children between the ages of 5 and 14 were economically active.[44] Of this total 3.5 million (71 percent) were boys and 1.5 million (29 percent) were girls.[45] Except for India, this is a strikingly high rate, especially in comparison with other countries in the South Asian region (see Table 4.1). Official statistics have shown that the total working population between the ages of 5

[44] A person who works one or more hours for pay or profit or working in a family farm or found not working but had a job or business from which he or she is temporarily absent during the last week of the survey.

[45] We wish to note in advance that gender differences in labour market participation should be regarded with some caution, as the survey does not properly capture domestic chores and this is the type of work that female children most often do.

and 17 was approximately 7.9 million, of which 5.8 million (73 percent) were boys and 3.1 million (27 percent) were girls.[46]

The widespread prevalence of child labour in Bangladesh, despite the government's programs and laws prohibiting work by children, suggests that additional policy measures to curb child labour are warranted.[47] It is widely believed that there is a lack of harmony among laws that uniformly prohibit the employment of children or set a minimum age for employment. Under the current law, the legal minimum age for employment varies, between 12 and 16, depending on the sector (Khanam 2006). This is because these laws focus mainly on the employment of children in the manufacturing, retail and establishment sectors, but they ignore the employment of children in the agricultural sector, which absorbs approximately 56 percent of the total child labour force. In addition, the informal sector and domestic work are exempted from these laws. Thus, more than 80 percent of the economic activity of children falls outside the protection of the labour laws.

In 1995, Bangladesh signed a Memorandum of Understanding (MOU) which had been undertaken by the ILO and UNICEF to eliminate child labour from the garments industry. As reported by Rahman et al. (1999), this approach neither reduced child labour among these children nor increased their schooling. A second MOU was undertaken by the same parties in 2000 to reaffirm the agreements of the first MOU and to develop a long-term and sustainable response to monitoring child labour in the garments industry (Khanam 2006).

Bangladesh also adopted school subsidy provisions to improve schooling and thereby attract and retain children. The innovative program of this type is Food for Education (FFE), which was introduced in 1993 and made available to rural children.[48] Ravallion and Wodon (2000) find that the FFE program has been successful in increasing school enrolment (from approximately 75 to 90 percent) but did not have a significant effect on child labour. Thus, the authors conclude that participation in the child labour force may not be very responsive to education-related policy measures.

[46]All these statistics in this section are based on the Report on the National Child Labour Survey 2002-2003, Bangladesh Bureau of Statistics, Dhaka, Bangladesh.
[47]There are 25 special laws and ordinances in Bangladesh to protect and improve the status of children in Bangladesh (see Khanam 2006 for more details).
[48]The main feature of the program is to provide a free monthly food ration contingent on the family being judged as poor and having at least one primary-school-age (at least six years old) child attending school that month.

Another educational incentive program that encouraged girls to increase their (junior) secondary schooling (i.e. Grade 6-10) was found to be effective in increasing secondary school attendance (see Arends-Kuenning and Amin 2004).[49] However, no research thus far attempts to shed light on whether this particular subsidy reduces participation in child labour.

As previously mentioned, the empirical analysis of this essay is based upon the individual-level data from the 2002-2003 National Child Labour Survey (NCLS). The NCLS considers a child (aged 5-17) to be employed if he or she worked at least one hour during the reference week (the week preceding the day of the survey). However, the survey does not consider child participation in domestic work to be labour. To enable our empirical analysis, we focus on children between the ages of 5 and 17 who worked at least one hour during the reference week as a paid (wage) employee (paid in cash or in kind), or who was self-employed,[50] or who worked as an unpaid employee (for example, work on the family farm or in family businesses) related to the household head.[51] This is especially important as globally only a relatively small fraction of children work for wages. Furthermore, we follow the definition of work similar to the NCLS, that is, we exclude domestic work. For the estimation of child labour, five years may be considered extreme because this is the cut-off age between infancy and childhood. However, it is not unusual in case of Bangladesh, particularly in rural areas. On the other hand, although the official enrolment age in Bangladesh is six years, there are some children who start school at the age of five (and therefore, the start of the potential trade-off between schooling and child labour). We extend the age limit to 17 years in order to capture a better interaction between child labour and education. According to the education system in Bangladesh, students at the age of 17 should be at the beginning of higher secondary school. However, the data suggests that there are some children in the age group of 5-17 who are still in primary school. This is very common, especially in rural areas.

---

[49] To promote school enrolment of girls in secondary schools, the government of Bangladesh, with the help of the World Bank, undertook the Female Secondary School Assistance Project (FSSAP) in 1993. The primary component of the project was to provide monthly stipends to female students from Grade 6-10. The tuition fees, as a part of the stipend, were directly issued to the school where the children were enrolled. The stipend increased by grade.

[50] A self-employed or own-account worker is officially defined as a person who works for his or her own farm or non-farm enterprise for profit or family gain.

[51] NCLS classified children as sons and daughters if they are the son or daughter of the head of the household or spouse. The father is called the head of the household if the head is identified as male and the mother is called the spouse (if listed as the opposite sex), and the mother is called the head if the head is identified as female and the father is called the spouse (if listed as the opposite sex).

We use two measures of children's schooling: school attendance and grade-for-age. However, it is often posited that a more accurate assessment of the impact of child labour on human capital development should focus on measures of learning outcomes, such as test scores, rather than school enrolment or attendance. We depart from this practice for two main reasons: a) test scores are not available for children in the dataset considered here, and b) the test scores which measure the reading, language and mathematical skills do not always provide a complete picture of learning achievements. This is especially true in a developing country such as Bangladesh where enrolling all school-aged children in school is still a major development challenge. In this survey, each child was asked whether he or she is attending school (full-time or part-time) at the time of the survey. Grade-for-age (GAGE) is defined as follows (Psacharopoulos and Yang 1991):

$$\text{GAGE} = [G/(A-E)] * 100$$

where G is the highest grade of formal schooling attained by a child, A is the child's age, E is the entry age to school. All those with a score under 100 are considered to be below normal progress in the school system because of grade repetition or late enrolment.[52,53] The formula for GAGE presented above highlights several issues when using data on very young children. For children who are in their first year of schooling, a strict interpretation of GAGE will give an infinite value since the denominator is zero (since $A-E = 0$). Furthermore, if a child starts school before the official school entry age (i.e. six years), then GAGE could potentially be greater than 100. Therefore, for GAGE specifications, we confine our sample to children between the ages of 7 and 17.

[52]Usually, GAGE is a real number between 0 and 100, where GAGE = 100 means that the individual has a good performance, and he or she has not repeated any year or dropped out of school. However, in a few cases, it could be the case that GAGE>100 because some children might start their education at an earlier age than the entry age. If GAGE is low (close to zero), this is a sign that this child has stopped studying for some years or has had a low performance. It is desirable that GAGE be close to 100.

[53]Ray and Lancaster (2005) employ the 'schooling for age' (SAGE) variable that measures schooling attainment relative to age. It is given by SAGE = years of schooling/(A-E) x100 where E represents the entry age to school. SAGE could not be calculated in this essay because NCLS does not report, 'years of schooling' as a continuous variable.

4.4 Data and Descriptive Statistics

The empirical analysis is based on the data drawn from the Bangladesh National Child Labour Survey (NCLS), conducted by the Bangladesh Bureau of Statistics (BBS) in 2002-2003 (henceforth NCLS 2002). This survey has been designed in the context of the commitments made by the Government of Bangladesh, following the ratification of the ILO Worst Forms of Child Labour Convention (No. 182) 1999. The NCLS 2002 was designed to provide reliable estimates of child labour at national, urban and rural levels, as well as by districts. The NCLS included a child population between the ages of 5 and 17 from 40,000 households. However, NCLS excluded children living in the streets or in institutions such as prisons, orphanages or welfare centres.

The analysis is performed upon a full sample of 14,062, 5-17 year old children drawn from the survey's urban and rural respondents. In this sample, 9,404 (67 percent) are male children and 4,658 (33 percent) are female children. Out of this sample, 2,801 males and 1,439 females reside in urban areas, while 6,603 males and 3,219 females reside in rural areas. 8,900 children are actively participating in the labour force, consisting of 6,750 males and 2,150 females. Of the children who are working, 2,508 reside in urban areas and 6,392 reside in rural areas. Approximately, 76 percent of both the urban and rural samples are male.

Table 4.2 presents definitions and descriptive statistics of the explanatory variables used in the analysis by child work status (i.e. working and non-working children). Child-specific characteristics include the child's age, education, and working hours (per week). Descriptive statistics conditional on work status suggest that working children are on average older and generally combine school with work more than their non-working counterpart. This difference is generally statistically significant at conventional levels.[54] In addition, a child gender dummy is included to capture the gender disparities in education and work that may arise due to differences in parental preferences. Moreover, the statistics indicate that there is a negative relationship between the labour supply and female children. It may be surmised that the demand for female child labour is high at home. At the household level, household compositions and household assets are included. Household compositions include a number of adult males and females aged over 17 years, which may reduce pressures upon the individual child.

[54] We computed this result using a standard *t*-test.

There is a difference between working and non-working children with respect to the number of adult males and females, and this difference is statistically significant at the 1 percent level. Since children must often care for their younger siblings, the number of younger children aged 0-4 is also included. It is evident from these statistics that children who supply labour tend to come from families with a smaller land holding and higher number of school-aged children. The difference between a working and a non-working child is generally statistically significant. We also included a set of parental characteristics that may influence parental decisions with regard to child labour. The statistics suggest that an improvement in parents' education will reduce child labour supply. This has important policy implications. However, there is little difference between working and non-working children in the effect of the father's education, but the difference is never statistically significant.

The remaining measure includes a set of community variables that may influence the demand for child labour. Hence, location (rural or urban areas) of the household is included as a regressor. The descriptive statistics suggest that the rate of incidence of child labour varies by urban and rural areas. Besides location, a policy measure to reduce child labour is an improvement in school quality. While we do not have any clear information about student achievement, we explore the effect of school input by including a set of dummy variables that capture the quality of school. These are the formal schools administered by the government, and the NGO schools run by non-government organisations. On average, approximately 59 percent of children who work go to a formal school, while the corresponding number for non-working children is approximately 21 percent. This difference is statistically significant at conventional levels.

Table 4.3 presents time allocation of children to schooling, work, or both, by urban and rural areas and by gender. A significant proportion of children simultaneously undertake both schooling and work activities. Interestingly, however, the proportion of children who only work is higher than that of those who both work and go to school. This finding differs from the findings of Shafiq (2007),who, using the HIES (Household Income and Expenditure Survey) 2000 from Bangladesh, found that children mostly attend school only whereas the proportion of children who both attend school and also work is relatively small. Rural children are more likely to work and go to school at the same time than are their urban counterparts. However, urban children are more likely to become idle (i.e. reportedly involved neither in schooling nor in child labour). The higher proportion of idle children in urban areas indicates that a household may pull children out of school before completion. However, the findings about idle children may be under-reported because NCLS 2002 does not consider domestic work as child labour or because households are hesitant to report child labour practices to survey collectors. With respect to gender, more female than male children attend school full-time (i.e. those attend school and avoid child labour), and fewer females than male children are employed full-time and combine work with schooling.

Table 4.4 presents the incidence of child labour force participation and school attendance for children between the ages of 5 and 17. In all areas, the child participation rate in the labour market increases with age, though not monotonically. Ray (2000a, 2002) also provides similar evidence from Nepal, Peru and Pakistan. In the case of child schooling, the attendance rate peaks around 12 years in urban and rural areas, and then falls. The gender picture is similar in both urban and rural areas with respect to child labour, with males registering a higher participation than females. This finding is consistent with Ray (2000b), who provides similar evidence from India. However, the situation differs sharply with respect to child schooling with a more even gender imbalance in the attendance rate between male and female children in the later age groups of 12-17 years. This observation is also borne out by Ray (2000a), who finds that Pakistani girls' schooling falls to nearly half that of boys in the age groups of 14-17 years. In the context of Bangladesh, there are several possible reasons for this drop-off. Girls are separated away from male contact at an early age (based on religion). Since there are few primary schools, and even fewer secondary schools reserved for girls, young females have to leave school on reaching adolescence. Another possible explanation is that it is customary for girls to marry early, which tend to further curtail schooling. Additionally, female children are often asked to do domestic chores, which further discourage educational advancement.

Interestingly, the school attendance rates of rural children in almost all age groups are consistently larger than their urban counterparts, with the former registering figures approximately 60 percent for males and 50 percent for females around 12 years and falls off sharply beyond 14 years. In the case of urban areas, the attendance rate rarely goes above 50 percent and falls off sharply beyond 14 years.

Table 4.5 shows the employment status of children by urban and rural areas and by gender. In rural areas 52 percent of working males were unpaid employees and 40 percent were paid employees. In urban areas, however, approximately 48 percent of working males were paid workers and 43 percent were unpaid workers (work without pay in family farms or in the family business). Similar patterns are not found for female working children. A large proportion of female child workers worked without pay in family farms or in the family business, and this is relatively high in rural areas.

These patterns suggest that opportunities for child workers are quite different in rural and urban areas. In rural areas, children are more likely to engage in agricultural activities and become unpaid workers, especially female children. In urban areas, children are more likely to find opportunities for some paid work. The gender difference in employment status among child workers is also significant in Bangladesh. Young females are more likely than young males to be unpaid workers in both urban and rural areas. This may imply that male children are increasingly entering the formal wage labour market rather than working as unpaid workers, and thus allowing female children to substitute into the unpaid activities.

Table 4.6 shows that industries that employ child labour are quite different for males and females in urban and rural areas. In urban areas, 44 percent of working males are employed in the construction industry, followed by 23 percent in the manufacturing industry and 22 percent in the agricultural industry. However, approximately 42 percent of working females were employed in the manufacturing industry, followed by 18 percent in construction. In rural areas, 63 percent of working males are employed in agriculture, followed by 21 percent in the wholesale and retail industry. Approximately 71 percent of working females were employed in agriculture, followed by 15 percent in manufacturing.

Figure 4.1 shows the interaction between the number of working hours and school attendance. This is important for the purposes of this essay because of the direct effects of the number of working hours on school outcomes. The more children have to work, the more tired they will be when in school and the less time they will have for study. Consequently, hours worked may have an adverse effect on learning while in school, even if they do not have a large effect on enrolment. In Bangladesh, there appears to be a threshold beyond which the number of hours worked is strongly associated with reduced school attendance. The number of working hours appears to have a relatively small impact on school attendance up to the 15-29 hours cohort, but attendance falls off dramatically when children work more than 29 hours per week. The decline is more gradual for girls than for boys, perhaps because some kinds of chores and subsistence work are more compatible with school.

4.5 An Analytical Framework

In general, a unitary model (or common preference model) is typically used to analyse the economic contribution of children.[55] Here, a household maximises the joint utility function of all its members. Quantity and quality of children, consumption of leisure and other market goods all enter the household utility function.

There is an extensive literature that suggests that male and female heads of the household may have different utilities, reservation utilities, and budget constraints and that they therefore may make different decisions (see, for example, McElroy and Horney 1981). The resolution of the preference difference of the male and female household heads may depend on the relative bargaining power of each individual, and this power may depend on (a) control over assets, both current and those brought into marriage; (b) unearned income or transfer payments and welfare receipts; (c) access to social and interpersonal networks; and (d) attitudinal attributes. In this essay, we allowed for the fact that parental preferences may not be conjugal and that differences may exist in parental preferences regarding child's work and schooling. In the absence of direct information on parental preferences, as well as the non-labour income of each spouse, we shall, in this essay, consider the different role of the mother's and father's education. We propose that women's education increases women's bargaining power in household decisions and if the mother obtains a higher marginal disutility from child labour than the father, an increase in the bargaining power of a mother will reduce child labour and increase child schooling. Indeed, Basu and Ray (2001) and Basu (2006) have shown that when women's bargaining power (as measured by level of education) increases, child labour initially falls but beyond a point it will tend to rise again.

[55] Unitary models assume that either all household members share the same preference function or that a single decision-maker takes all decisions. In any case, the upshot is that the household behaves as if it were a single or a unitary agent (Becker 1981).

Let us assume that a household is comprised of a mother and father and some number of children. Fertility is assumed to be exogenous. In general, each parent is considered altruistic in that they care about the consumption of each member of the household and the quality (educational attainment) of their children. All decisions are made by altruistic parents, and children are treated as recipients of parents' decisions.[56] The parents allocate the child's total endowment of time between school attendance and work. Leisure is not included in this analysis for simplification. Household income must meet the cost of household consumption and schooling. Household income is generated by a typical household production function with decreasing returns. We assume it is a function of the parent's non-labour income, the parent's labour income and the child's income. We shall maintain the strong assumption that non-labour income is exogenous, and we therefore ignore the fact that current non-labour income probably reflects past labour supply decisions. While primary education is almost free in Bangladesh, schooling costs can be significant in terms of costs on transportation, school uniform, utensils and so on, especially for the poor. Parental investment for a given level of schooling would depend on a vector of child, parental, household and communal characteristics. Child labour supply equals child's total endowment of time less school attendance and depends on a vector of child, parental, household and communal characteristics.

This simplified framework forms the basis of our empirical analysis, which is discussed in the following section.

4.6 Empirical Model

Let ${S_i}^*$ denote the latent variable that describes household decisions to enrol a child i in school.The equation for ${S_i}^*$is written as follows:

$$
\begin{aligned}
{S_i}^* = {} & \alpha_0 + \alpha_1 L_i + \alpha_2 L_i^2 + \alpha_3 G_i + \alpha_4 R_m + \alpha_5 (G_i * R_m) + \alpha_6 E_m + \alpha_7 (G_i * E_m) \\
& + \alpha_8 E_f + \alpha_9 \left(G_i * E_f\right) + \alpha_{10} Z + \mu_i \qquad (4.1)
\end{aligned}
$$

[56] In line with the literature, it is assumed that children do not bargain with their parents because they do not have a valid fallback option (Bhalotra 2007).

where L_i represents a child i's $(i = 1, \ldots, \mathrm{K})$ weekly working hours in the reference week and its quadratic term L_i^2 captures the non-linear effects of hours worked,[57] G_i is the dummy for a female child. E_m and E_f denote the mother's and father's education measured by the highest grade attained.[58] Emerson and Souza (2007) conclude that the mother's and father's education may affect investments in the male and female children differently. The gender of the sampled child is, therefore, interacted with each parent's education $E_j(j = \text{mother or father})$. We also include a dummy variable taking the value of 1 if the mother's level of education is higher than that of fathers and 0 otherwise (R_m).[59] The mother's level of education relative to the father's is assumed to influence her ability to negotiate in defence of her preferences with respect to her children's schooling. Z is a vector of exogenous child, household and community characteristics that determine S_i^* and μ_i is the random factor.

In practice, however, we do not observe S_i^*. For school attendance, one could observe the following binary measures:

$$S_i = \begin{cases} 1 & if\ S_i^* > 0 \\ 0 & if\ S_i^* = 0 \end{cases}$$

Alternatively, for GAGE:

$$S_i = \begin{cases} GAGE & if\ S_i^* > 0 \\ 0 & if\ S_i^* = 0 \end{cases}$$

Thus, the estimating equation is:

$$S_i = \alpha_0 + \alpha_1 L_i + \alpha_2 L_i^2 + \alpha_3 G_i + \alpha_4 R_m + \alpha_5 (G_i * R_m) + \alpha_6 E_m + \alpha_7 (G_i * E_m) + \alpha_8 E_f + \alpha_9 \left(G_i * E_f\right) + \alpha_{10} Z + \mu_i \qquad (4.2)$$

While we estimate Equation (4.2) by probit if the variable of interest is school attendance, we use the Tobit model if the variable of interest is GAGE because GAGE has observed zero values for approximately 46 percent of children aged 7-17.

[57] Ray and Lancaster (2005) and Han (2008) also employed the quadratic term of the number of working hours in their study of child labour and schooling.

[58] Ravallion and Wodon's (2000) study on child labour and child schooling in Bangladesh also used binary variables to indicate level of education completed by the mother and the father.

[59] Alternatively, we could have included the difference between the mother's and father's level of education.

One potential concern of estimating Equation (4.2) is that child labour hours and child labour hours (squared) may not be exogenous. Specifically, certain unobserved factors that determine school attainment may also explain hours worked, causing our probit or Tobit model estimates to be biased and inconsistent.[60] To avoid this endogeneity problem, we use a two-stage instrumental variable (IV) strategy. There is, however, one other problem. Note that household income included in variable Z is most likely to be endogenous because whether or not children work is likely to influence parents' reservation wages and their labour market participation (Wahba 2006). To avoid endogeneity of this variable, we include an occupation status of the father to proxy wealth.[61]

Thus, the reduced-form work equation is written as follows:

$$L_i = \beta_0 + \beta_1 G_i + \beta_2 Z + +\beta_3 V_i + \beta_4 R_m + \beta_5 (G_i * R_m) + \beta_6 E_m + \beta_7 (G_i * E_m) + \beta_8 E_f + \beta_9 (G_i * E_f) + \nu_i \tag{4.3}$$

where V_i is the instrumental variable and ν_i is the random factor. Instruments are described and justified in the following section.

The instrumental variable (IV) procedure can be described as follows. In the first stage, we estimate Equation (4.3) by OLS and obtain the residual (υ). We follow the same procedure when the child labour hour (squared) is the dependent variable. In the second stage, we include predicted residuals from first-stage regressions into Equation (4.2) as additional regressors. The significant coefficients on residuals imply that the null hypothesis of exogeneity of child labour hours is rejected. Equation (4.2) can now be written as follows:

$$S_i = \alpha_0 + \alpha_1 L_i + \alpha_2 L_i^2 + \alpha_3 \hat{\upsilon}_{L_i} + \alpha_4 \hat{\upsilon}_{L_i{}^2} + \alpha_5 G_i + \alpha_6 R_m + \alpha_7 (G_i * R_m) + \alpha_8 E_m + \alpha_9 (G_i * E_m) + \alpha_{10} E_f + \alpha_{11} (G_i * E_f) + \alpha_{12} Z + \mu_i \tag{4.4}$$

where $\hat{\upsilon}_{L_i}$ and $\hat{\upsilon}_{L_i^2}$ are predicted residuals from the OLS estimates of child labour hours and child labour hours (squared) equations, respectively.

---

[60] The unobserved determinants could be person-specific or individual attributes, such as 'motivation' or 'energy' that might drive certain children to both work more and study more.

[61] We assume that the contribution of children's income to overall household income is not large. In doing so, we have abstracted from another significant endogeneity problem.

4.7 Results and Analysis

This section reports the empirical results. We start by discussing the results from an IV specification. We estimate two sets of regressions: an IV probit equation characterising the school attendance and an IV Tobit equation characterising GAGE. The findings from a non-parametric method are then reported.

4.7.1 Results of the Instrumental Variables Specification

The potential endogeneity of the number of working hours is verified through a Hausman test. The chi-square test rejects the joint exogeneity of hours worked and its square term in almost all specifications that we estimated (see Tables 4.7 and 4.8, bottom). To solve the endogeneity problem, we relied on instruments such as a set of industry dummies (agriculture, manufacturing, wholesale and retail and service) where the child works. We choose a sector of employment because the amount of hours worked that is available for a child depends on the job to which he or she is assigned in a particular sector. There is no evidence that the industry dummies considered in this essay have any direct impact on schooling. We assume that the only influence that the sector of employment has on schooling must come through its effect on the number of hours worked, and not through any other channels. According to NCLS 2002, children employed in manufacturing and wholesale and retail work approximately 50 hours or more per week compared to other sectors in Bangladesh. In these sectors, children may be offered to work more hours, or they may decide to work long hours due to a higher likelihood of extra payment. For example, it is well known that in the garments industries in Bangladesh children are offered to work more than the minimum number of hours available (i.e. eight hours a day), which might partly explain why child labour hours and schooling are not compatible. In the case of the agriculture or service sectors, it can be assumed that only a small amount of flexibility in working hours is available. The harvest season, for example, is unlikely to be compatible with public school schedules. Thus, it is reasonable to assume that jobs in these sectors do not provide time that can be used for studying.

Since we lack a strong theoretical justification for the exogeneity of these instruments, we rely on empirical tests. The relevant test lends strong credence to our use of industry dummies as instruments for both 'child labour hours' and 'child labour hours (squared)' equations (with a *p*-value of 0.000) (see Tables C1 and C2 in Appendix C, bottom).[62] Over-identification tests are also conducted to evaluate whether the proposed instruments can sensibly be excluded from both 'school attendance' and GAGE equations. The Hansen *J*-statistic concludes that all instruments can legitimately be excluded in the estimation of both 'school attendance' and GAGE equations (Tables 4.7 and 4.8, bottom).

School Attendance

Table 4.7 presents the results of school attendance using an IV probit for all children. We also present the marginal effects in Table 4.7 as they are more easily interpreted. The marginal effects of each variable on the probability of attending school show the following.

Initially, we include individual covariates, such as the child's age, its square term and the dummy variable for a female child, but exclude education levels of the mother and the father and other covariates as explanatory variables in the regression. We find that there is a significant age effect: age is positive and statistically significant (Column 2). This implies that the probability of school attendance increases with the age of the child. The square of the age of the child shows that there is a non-linearity in the age effect (the square of the age term becomes negative and statistically significant). As the child gets older, the likelihood that he or she will drop out of school and engage in market-oriented work increases. This is consistent with findings for various countries reported in the literature; see the review of Dar et al. (2002). One possible explanation of this result is that older children either have completed their studies or have failed to continue them. It may also be the case that as children grow up, they acquire more experience and more human capital, which creates a prospect for higher wages that induce them to leave school. Being female reduces the chance of being in school. This is a common scenario in Bangladesh and in other South Asian countries (Ray 2000a; Rosati and Rossi 2003) where there is a significant gender differential in the probability

[62] We perform an F-test that the coefficients on the industry dummies are jointly zero. The F-statistics ranging from 137 to 274 depending on different specifications for school attendance, indicating that the instruments add significantly to the prediction of the working hours (Table C1 of Appendix C, bottom). The corresponding F-statistics for GAGE specifications varied between 136 and 272 (Table C2 of Appendix C, bottom). Tables C1 and C2 also report the adjusted R-squared for school attendance and GAGE regressions, which varied between 0.20 and 0.30.

of school attendance and this bias is generally in favour of male children. The marginal effects (evaluated at the mean of the independent variables) show that being a female child is associated with a decrease of approximately 18 percentage points in the probability of school attendance.

The coefficient estimates for ‘child labour hours’ and ‘child labour hours (squared)’ are statistically significant, but are of opposite signs at conventional levels. The negative magnitude of the estimated coefficients of the ‘child labour hours’ variable support the proposition that working hours adversely affects the probability of the child attending school from the very first hour of work. However, the estimated positive coefficients of ‘child labour hours (squared)’ suggest that the adverse marginal impact of child labour hours on the schooling variable weakens when working hours increase.[63]

Next we turn to an analysis of how parents’ education influences the nature of trade-off between working hours and school attendance. We begin with the first comprehensive indicator of the mother’s level of education: whether or not the mother is more educated than the father. In this specification, we do not include the father’s level of education to avoid possible correlation with the mother’s education, if there is assortative mating. In general, the results confirm that if the mother’s education is higher than that of the father’s, it promotes child welfare by reducing children’s hours of work and by increasing school attainment (see Column 4). The effect, however, is not stronger for a female child, as indicated by the interaction of the female dummy with the mother’s education.

The other indicator of the mother’s education, such as the level of education of the mother (as measured by the highest grade attained), is included to verify the robustness of our results outlined above. The mother’s level of education shows a greater positive effect on the schooling of male and female children but has a differentially higher effect on the female child, as shown by the positive coefficient for the interaction

[63] What this implies is that hours worked by a child has a U-shaped relationship with schooling outcomes. The result is consistent with the findings of Ray and Lancaster (2005). We do not have sufficient evidence to explain the reasons for this pattern of relationship between child labour hours and school attendance. However, it is undeniable that long working hours, beyond a certain threshold level, have a negative effect on children’s education. To explore this issue, we include the third and fourth degrees of polynomials in hours and find that the relationship between child labour hours and the likelihood of school attendance is not always confined to a U-shaped relationship.

of the female dummy with the mother's level of education (see Column 6). Shafiq (2007), using the HIES 2000, found similar results for Bangladesh. This result further corroborates the findings that children's school attainment is linked to the educational attainment of the parent of the same sex as the child (see, for example, Emerson and Souza 2007).

Next we include the father's level of education (as measured by the highest grade attained) to account for the trade-off between child labour and schooling in the effect of both the mother's and father's level of education. Altogether, the results suggest that the father's education has a greater positive impact on his daughter's school attendance than on a son's. This is indicated by the interaction of the female dummy with the father's education and, hence, shifts the trade-off towards a daughter's schooling (Column 8). This is not a new result in the context of Bangladesh. Using the HIES 2000 Shafiq (2007) finds that a female child has a seven percentage point higher probability of being enrolled in school by having an educated father. One possible explanation is that father's education may be more important because fathers are often more educated than mothers in developing countries. Alternatively, it may be that fathers play a more active role in certain kinds of decisions. On the other hand, though the effect of the mother's level of education increases the probability of school attendance for all children; the preference for educating a female child virtually disappeared when controlled with the father's education. We verify the robustness of these results by including additional covariates, namely household compositions, household assets and community characteristics (see Column 10).

Importantly, in Column 10, the same pattern emerges as in Column 8. The results are summarised as follows: as far the as the effect of child labour hours is concerned, we find a similar association between the number of hours worked and the probability of school attendance. The coefficient for a female child is lower but still significant. We note that the presence of school children between the ages of 5 and 17 reduces the probability of school attendance. This finding corroborates past evidence from Bangladesh (Amin et al. 2006). These findings may shed light in favour of the quality-quantity trade-off and the effects of sibling competition. Furthermore, it is argued that large numbers of school-aged children demand more resources to be put into their education, which in turn forces them to be employed in case of parental resource constraints, to make schooling possible for themselves and for their siblings. This may have a negative impact on their schooling outcome. On the other hand, there is strong evidence that the presence of adult female members in the household increases the probability of child school attendance. This effect might reflect the large decision-making power of females among adults in the household. As expected, the probability of child school attendance increases in urban areas than in rural areas. We also note that the father's occupation reduces the probability of attending school if the father is involved in agricultural activities (the reference category is non-agricultural activities). There are three possible routes whereby a father's agricultural activities may affect his child's schooling. First, a person working in the agricultural sector may be resource constrained compared to those involved in non-agricultural activities and may decide that it is not worthwhile to send his children to school at all. Second, agricultural activities in Bangladesh are characterised by seasonal variation, and therefore, it is not uncommon for families to engage in non-agricultural activities to supplement household income. Involvement in non-agricultural activities by adult members increases the demand for child labour in activities where child and adult labour are substitutes. Third, a household whose primary livelihood is from agricultural activities may see a lack of relevance in formal education. With respect to parents' education, we find that both the mother's and father's education has a positive effect on the probability of attending school, though none is statistically significant, possibly picking up the wealth effect within the household. The positive and significant coefficient of interaction term between the child's gender and the father's education might suggest the trade-off towards a daughter's schooling. These results perhaps suggest that parents' education, particularly the father's education, may not be a good proxy for the permanent income effect and, hence, increases the likelihood that a male child will work.

GAGE

Further evidence on the adverse impact of child labour on child schooling is shown in Table 4.8, which presents IV Tobit estimates of GAGE. To better understand the results, we limit our discussion to marginal effects presented in Table 4.8. In Column 2, the marginal effects of IV Tobit estimates further confirm the adverse impact of working hours on the child's grade attainment. These findings are supported by Ray and Lancaster (2005) who find similar results measured by SAGE. As expected, being a female child is associated with a lower grade attainment.

However, we notice that if the mother's level of education is higher than that of fathers, there is a significant reduction of grade attainment for both male and female children. It appears that children are more likely to be exploited as child labour if the mother is more educated than the father, in which case mothers perhaps cannot participate in household decision-making on equal terms (see Column 4). On the other hand, the mother's level of education (as measured by the highest grade attained) shows a positive and significant effect on the grade attainment of her children (Column 6). In addition, this effect is stronger for educating a female child. Interestingly, the inclusion of the father's level of education in the model shows that there is no significant change on the coefficients for the mother's level of education, though the magnitude of the coefficient is now smaller (Column 8).[64] Also, the magnitude of the impact of the father's level of education is significantly smaller than that of the mother's. The father's education, however, does not reveal any bias towards a female child.

The final robustness check includes all covariates. We again find that being a female child is associated with lower grade attainments, but this is no longer statistically significant (see Column 10). This perhaps suggests that grade attainments for female children are determined more by institutional factors or other issues than by household decisions. Another point is noteworthy. The coefficients for child labour hours suggests that the importance of working hours for grade attainment is much weaker, even after controlling for individual, parental and household characteristics. These results imply that other factors are more crucial in determining a child's grade attainment. The results regarding parents' education remain in most cases.

4.7.2 Results of a Non-Parametric Approach

In the previous section, we have shown that a small increase in child labour hours is detrimental to child schooling, considering the endogeneity problem of child labour hours. To make our findings more robust, we further examine the relationship between child labour hours and schooling with the kernel (weighted) regression approach. The relationship between working hours and the probability of school attendance is presented in Figure 4.2. Figure 4.3 reports the case of GAGE. A closer examination of Figure 4.2

[64] However, these findings must be treated with some caution due to the effect of the ability bias or the effect of assortative mating on the intergenerational transmission of schooling.

shows that the probability of attending school declines continually with the increase of hours spent at work in economic activities and, hence, provides evidence that work and schooling are competing activities. This finding is consistent with Guarcello et al. (2006) who find a similar association in four countries (Bolivia, Mali, Cambodia and Senegal) between hours worked (spent in domestic chores and economic activity) and not attending school. Similarly, we find that there is a decline in grade attainment if working hours increase. These results highlight the importance of employing a non-parametric estimation method which is free of restrictive assumptions on the functional form.

4.8 Controlling for Sample Selection Bias

As our estimates refer only to the subsample of children working in economic activity, this could generate a selection bias in the estimates. Children who participate may have unobserved characteristics that are correlated with the unobservables in the outcome equation (in this case, school).[65] We address this selection bias by using the sample selection model. The selection (first-stage regression) should be on hours worked, as children working different hours might share different characteristics. The selection equation should then be based on a Tobit model. However, the selection variable would be either 0 for non-workers and positive for the other observations. One simple way to estimate the model is to reformulate the Tobit model as a probit model, selecting on a variable defined as 1 for working children and 0 for non-working children. This is the approach followed in this essay.[66] Hence, the selection equation is:

$$I_i = \begin{cases} 1 & if\ (L_i = \beta x_i + \nu_i) > 0 \\ 0 & otherwise \end{cases} \qquad (4.5)$$

where I_i and L_i are the latent and observed working hours of child $i(i = 1, \ldots, K)$, respectively. We observe child work participation as $I_i = 1$ if $L_i = \beta x_i + \nu_i > 0$. The vector x_i denotes the aggregate form of all the explanatory variables (except for the number of adult males and females over 17 years) included in Equation (4.3) and

[65] Working children, for example, who are in school are often found to have lower test scores than non-working children. However, it may be wrong to say that working leads to poor test outcomes, based simply on that information. It may be that children with less scholastic aptitude 'select' into the state of working, while more intellectually capable children remain in school.

[66] Theoretically, this approach may sacrifice some efficiency by discarding information on the dependent variable. However, this is not necessarily true in a finite sample (Green 1997).

$\nu_i \sim IIDN(0,1)$. The selection correction term or the inverse Mill's ratio (λ) obtained from Equation (4.5) is included in Equation (4.6):

$$S_i = \alpha_0 + \alpha_1 L_i + \alpha_2 L_i^2 + \alpha_3 \hat{\upsilon}_{L_i} + \alpha_4 \hat{\upsilon}_{L_i}{}^2 + \alpha_5 G_i + \alpha_6 R_m + \alpha_7 (G_i * R_m) + \alpha_8 E_m + \alpha_9 (G_i * E_m) + \alpha_{10} E_f + \alpha_{11} (G_i * E_f) + \alpha_{12} Z + \alpha_{13} \lambda_i + \mu_i \tag{4.6}$$

In fact, Equation (4.6) is equivalent to Equation (4.4), except for the addition of a selection correction term, which is included to adjust for the non-random sample. Equation (4.6) is estimated by probit if the variable of interest is school attendance. This technique is known as the Heckman probit model where both the selection equation and the outcome equation are binary choices (Van de Ven and Van Praag 1981).[67] In the case of GAGE, Equation (4.6) cannot be estimated by Tobit in a sample selection framework. In general, the Heckman model assumes that the errors of the participation and schooling outcomes equations are correlated and the participation decision dominates the schooling outcomes. Dominance implies that observed zero values for GAGE are the result of participation decision only and that once the first hurdle (that is, participation) is passed, censoring is no longer relevant. This implies that only individuals with positive values for GAGE are included in the GAGE equation. This is typical of Heckman's generalised sample selection model (Jones 1989). In our case, we simplify the Heckman sample selection model by assuming that the participation and schooling outcomes equations are independent. In this case, the model reduces a probit for participation (Equation 4.5) and OLS using the subsample when GAGE is greater than zero.[68] In this case, both equations are estimated separately.

As is well known, the sample selection model requires an exclusion restriction, in the form of one or more variables that appear in the participation equation but not in the schooling outcomes' equation.[69] We include the sex of the household head and the number of adults aged over 17 years in the household as the exclusion restrictions.

[67] Examples of the application of the Heckman probit model are available in other fields (see, for example Trejo 1993).

[68] A similar method has been used by Aristei et al. (2008) in the study of alcohol consumption.

[69] In principle, instruments are not needed. The sample selection model can be identified by the non-linearity of the inverse Mill's ratio. However, estimates of the selection model often lead to substantial collinearity between the predicted inverse Mill's ratio term and the remaining covariates in the outcome (schooling) equation when a common set of covariates was entered into selection and outcome (schooling) equations.

How Important is the Selection Effect?

The results from the second stage of a selection model of school attendance and GAGE are shown in Tables 4.9 and 4.10, respectively. The marginal effects of the corresponding estimates for school attendance are also reported. The signs of the estimated coefficient on the inverse Mill's ratio are of interest here. The results suggest that the exclusion of the non-working children generates the selection bias. This is confirmed by a likelihood ratio (LR) test of ρ in four out of five school attendance equations (see Table 4.9, bottom).

In general, the sample selection estimates are largely similar to IV estimates (Table 4.7). We find that with the inclusion of the inverse Mill's ratio, the magnitude of coefficients changes (as well as their standard errors), but the overall patterns are not dramatically different. The pattern of significant coefficients is largely unaffected.

We next turn to the marginal effects of the probability of school attendance adjusted for sample selection bias (Table 4.9). We again find that when working hours increase the probability of school attendance declines, but at a decreasing rate.[70] However, a higher level of schooling of the mother compared to father now exerts a negative and significant effect on school attendance for all children (Column 4 of Table 4.9). It is difficult to give an explanation for this result. All we can say is that the result is quite sensitive to the specification used. The magnitude of the impact of the mother's education (as measured by the highest grade attained) on her children remain positive and statistically significant (Column 6). We also note that the mother's level of education shows a significant preference for educating a female child, but the effect is now less strong. As before, the impact of the mother's education continued to be significant when we included the father's level of education (as measured by the highest grade attained) in the regression (Column 8). The father's education still shows a significant preference for educating a female child. Finally, the model is estimated with all covariates to ensure that the results are not being biased by the potential sample selection (Column 10). When we compare these results with those presented in Column 10 of Table 4.7, the results are qualitatively similar, however, the magnitude of the impact of the father's education on school attendance of a female child is now smaller.

[70]The relationship between hours worked and school attendance is examined further with different degrees of polynomials in hours and we find (not shown) that the relationship between child labour hours and school attendance is not confined to a U-shaped relationship.

Turning to the OLS estimates of grade attainment corrected for sample selection bias (Table 4.10), we reveal a similar association between child schooling and child labour hours. However, there are several interesting and important differences from IV Tobit estimates presented in Table 4.8. First, the mother's education (as measured by the highest grade attained) does not have any impact on her child's education. However, it now shows a significantly larger impact on educating a female child (Column 6). Second, the impact of the father's education (as measured by the highest grade attained) becomes negative and statistically insignificant when we include the mother's education (Column 8). When the model is estimated with all covariates the effect of the father's education becomes positive, but continues to be statistically insignificant (Column 10). Thus, the sample selection correction does lead to substantive differences in inference.

4.8.1 Comparison with a Double-Hurdle Model

In this sub-section, we compare the sample selection model for GAGE with a double-hurdle model. In the sample selection model, we estimate the GAGE equation by using only positive observations, but this could lead to biased results. To solve this problem, we need to analyse the underlying reasons for observed zero values. One potential reason might be that many of the respondents are not attending school during the particular survey period (i.e. their highest grade should be zero).We therefore consider the possibility of a corner solution. This implies that zero values are the result of the participation decision, and the survey respondents may have zero values for GAGE. This reason motivates us to use a different approach to estimate the GAGE equation. The approach can be analysed by a double-hurdle framework. Thus, in order to observe positive values for GAGE one needs to pass two hurdles: the participation hurdle and the schooling outcome hurdle (i.e. GAGE). In estimating a double-hurdle model, it is common to assume that error terms from Equations (4.5) and (4.6) are independent (Cragg 1971). The model is essentially then a two-step procedure with a probit for probability of participation in the first stage and a truncated normal regression in the second stage. In contrast to the sample selection model, a double-hurdle model does not require exclusion restrictions. However, a specification issue in double-hurdle models concerns the choice of the regressors to be included in the participation and schooling outcome equations. Indeed, Aristei et al. (2008) argued that the inclusion of the same set of regressors in each hurdle model makes it difficult to identify the parameters of the model correctly, and exclusion restrictions must be imposed. The exclusion restrictions are the same as those used in the sample selection model.

Generally, we identify substantive differences in the results between the sample selection model and a double-hurdle model, as presented in Table 4.11.[71] First, the relationship between working hours and grade attainment is much weaker (the coefficients for working hours are not significant at conventional levels) (see Columns 6, 8 and 10). Second, the effect of the mother's education (as measured by the highest grade attained) becomes insignificant when controlled with the father's education (as measured by the highest grade attained); whereas the father's education now shows a positive effect on his child's grade attainment but continues to be statistically insignificant (Column 8). However, the effect of the father's education now becomes statistically significant after controlling for individual, household and communal characteristics (Column 10).

4.9 Additional Robustness Checks

This sub-section discusses the robustness of our main findings. First, we examine what happens if we restrict ourselves to the sample of wage (paid) workers. Paid work involves longer hours than other sorts of work (see Figure 4.4), and virtually rules out school attendance. In contrast, other forms of child labour may be more compatible with schooling. Isolating paid work, therefore, allows us to control for the confounding influence of working hours on children's schooling. Second, we also consider how the results change if the analysis is conducted separately for urban and rural areas. We continued to restrict our sample to children aged 7-17 for GAGE specifications. Note that we again limit our discussion to marginal effects in all regression models.

4.9.1 Isolating Wage Employees

In the results presented in Table 4.12, we find that working hours significantly reduces school attendance but at a decreasing rate. On the other hand, a different pattern emerges with the Tobit regressions for GAGE with the number of hours worked, as we now

[71] It is likely that methodological differences contribute to these dissimilar results.

obtain a statistically insignificant effect of working hours on a child's grade attainment (Table 4.13). When we look at the results for parental characteristics, we find that the effect of the mother's education (as measured by the highest grade attained) continues to be significant with and without controlling for the father's education (as measured by the highest grade attained) in both school attendance and GAGE regressions. While the effect of the father's education is significant in GAGE equations, it becomes statistically insignificant in school attendance equations if we take into account the mother's education. Interestingly, the mother's education does not show any preference for educating a female child in all specifications under consideration. While the father's education though shows a strong preference for educating a female child in the school attendance equation, this effect virtually disappears when controlled only with the mother's education and with all covariates in GAGE equations (see Columns 8 and 10 of Table 4.13).

4.9.2 Results by Urban and Rural Areas

There are several points to be noted when looking at urban and rural areas separately (Tables 4.14 to 4.17). Looking first at the results for school attendance (Tables 4.14 and 4.15), we find that when working hours increase children are less likely to attend school, but at a decreasing rate in both areas. Results are reversed with the GAGE specifications (Tables 4.16 and 4.17), as the adverse effect of working hours on grade attainment is, to some extent, significant only among the rural subsample (see Table 4.17). This may be due to the pattern of activities available to rural children. Turning next to parents' education, we reveal that both the mother's and father's education (as measured by the highest grade attained) show a positive and significant effect on school attendance and the grade attainment of male and female children in rural areas (Tables 4.15 and 4.17). The effect of parents' education, however, is weaker for school attendance specifications after controlling for individual, household and communal characteristics (see Column 10 of Table 4.15). A similar effect of parents' education is not found in urban areas. This is especially the case for child's school attendance. It is important to note that the mother's education does not show any bias towards a female child's school attendance in both urban and rural areas (see Tables 4.14 and 4.15) and in the case of the grade attainment, particularly in urban areas (Table 4.16), while for the father's education, this is true for school attendance in rural areas and for the grade attainment in both urban and rural areas (Tables 4.16 and 4.17).

4.10 Summary and Conclusion

This essay has investigated the trade-off between child labour hours and child schooling outcomes in Bangladesh using the individual-level unit record data from NCLS 2002. Controlling for a large number of covariates and correcting for all sources of endogeneity bias by instrumenting child working hours with a set of industry dummies, this study shows that working hours significantly reduces child schooling. Additionally, a non-parametric analysis shows that the relationship between working hours and schooling is not always confined to a U-shaped relationship rather children's schooling declines when working hours increase.

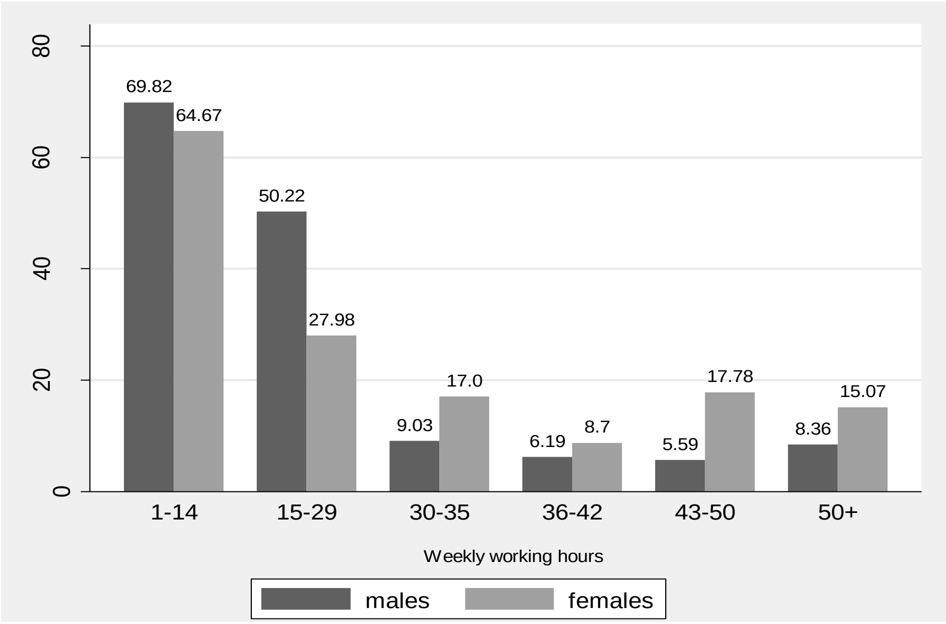
We find that parental characteristics, especially their level of education, affect the work-schooling trade-off across genders. The mother's education (as measured by the highest grade attained) show a significant preference in investing in female child's education, while a more educated mother (compared to the father) does not reveal any preferences towards a female child. A similar impact is found for age-adjusted grade attainment. However, the effect of the mother's education virtually disappears after controlling for all covariates. This is especially the case for a child's school attendance. We find that the father's education (as measured by the highest grade attained) affects daughter's school attendance relative to sons, even after controlling for the mother's education. The interpretation of the results is unchanged when the model is estimated with the full set of covariates. The magnitude of the impact of the father's education is larger in age-adjusted grade attainment and does not show any significant preference for educating a female child with and without the full set of covariates. After correcting for potential sources of selection bias, the qualitative results remain for school attendance but moderately change for age-adjusted grade attainment equations.

These results are in most cases very similar when we restrict our analysis to wage employees among children or split the sample by urban and rural areas. The most relevant issue is that the trade-off between working hours and child schooling observed previously still holds, except for Tobit regressions for GAGE. This is especially the case for wage employees among children and when we use the sample of urban child workers aged between 7-17 years.

These results have strong policy implications. While we find that the mother's and father's education has a positive impact on schooling of their children, the marginal impact of the mother's education on a child's schooling is considerably larger, irrespective of gender of a child. These findings, therefore, are consistent with our previous proposition that improvement in women's schooling and consequent increases in bargaining power have a greater beneficial impact on children compared with increases in men's schooling. Therefore, policies which improve education levels, especially the education levels of women, are more likely to reduce poverty and the incidence of child labour.

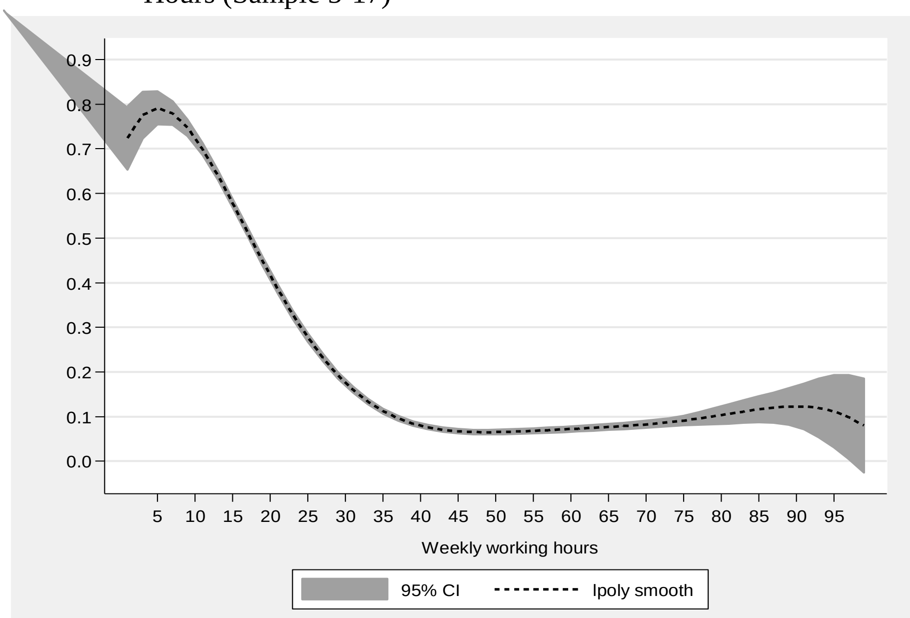
One other result is worth noting for its policy implication. Parent's educational levels are more likely to be correlated with the probability that a male child will work rather than go to school. Although the Bangladesh government adopts a variety of initiatives to ease the problem of child labour, our results strongly suggest that there is a limited relevance of these policies to supplement household income, which ultimately results in a higher incidence of child labour, particularly among young males. While education-related policy measures appear to be ineffective in the context of Bangladesh, we suggest the need for a combination of policies, such as stronger enforcement of compulsory schooling for children and an employment generation scheme for adults, to reduce child labour.

Figure 4.1: Working Hours and School Attendance of Children Aged 5-17, by Gender



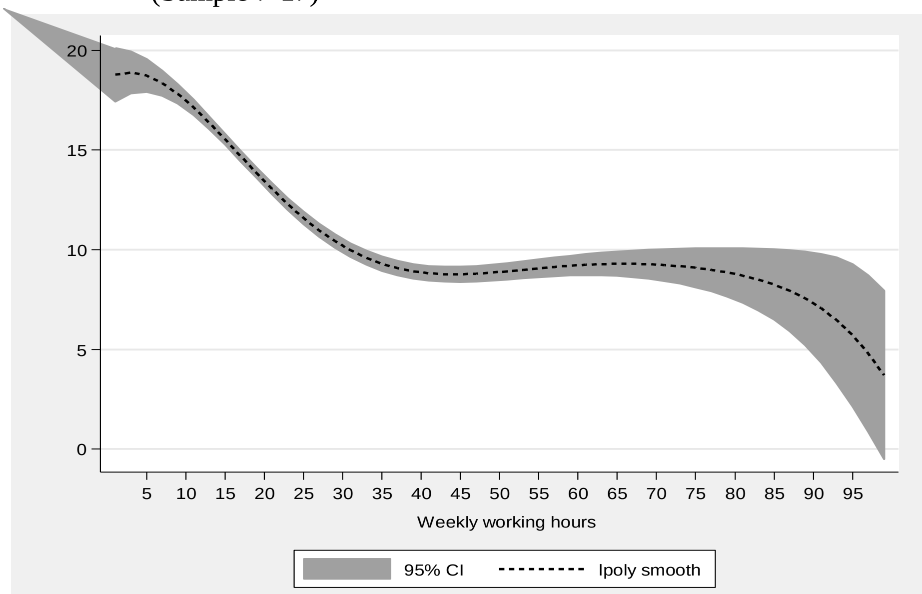
Source: Data are from NCLS 2002.

Figure 4.2: Kernel (weighted) Regression, School Attendance versus Working Hours (Sample 5-17)



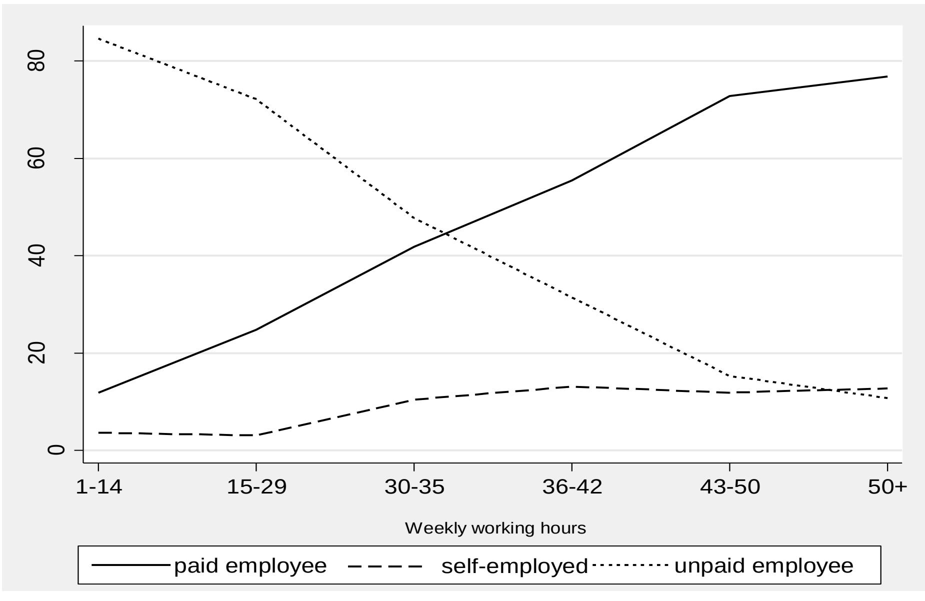
Source: Data are from NCLS 2002.

Figure 4.3: Kernel (weighted) Regression, GAGE versus Working Hours (Sample 7-17)



Source: Data are from NCLS 2002.

Figure 4.4: Distribution of Working Hours of Children Aged 5-17, by Work Status



Source: Data are from NCLS 2002.

Table 4.1: Estimates of Economically Active Children Aged 5-14 in the South Asia Region, by Gender

	Number of children (millions)	Males (%)	Females (%)
India	13.6	60	40
Bangladesh	5.0	71	29
Pakistan	3.3	73	27
Nepal	2.6	46	54
Sri Lanka[a]	0.5	62	38

Source: Ray (2004). [a]Sri Lankan Child Activity Status (SIMPOC 1999).

Table 4.2: Descriptive Statistics, by Child Work Status

	Workers			Non-workers				
Variables	N	Mean	Std. Dev.	N	Mean	Std. Dev.	*t*-test	
Child Covariates								
Age (years)	8,900	13.62	(2.1687)	5,162	8.3673	(3.8664)	103.12	***
Female (1= female)	8,900	0.2416	(0.4281)	5,162	0.4859	(0.4998)	-30.64	***
GAGE (Grade-for-Age)	8,900	12.7334	(13.6052)	5,162	9.1230	(13.8778)	12.01	***
Currently attending school (1= yes)	8,900	0.3415	(0.4742)	5,162	0.0651	(0.2467)	38.93	***
Working hours (hour/week)	8,900	28.52	(16.6865)					
Household Covariates								
Number of children 0-4	8,900	0.4853	(0.7528)	5,162	0.6885	(0.7925)	-15.21	***
Number of school children 5-17	8,900	2.8327	(1.1968)	5,162	2.6025	(1.2295)	10.88	***
Number of adult males over 17 years	8,900	1.4161	(0.8122)	5,162	1.2420	(0.7258)	12.73	***
Number of adult females over 17 years	8,900	1.3316	(0.6377)	5,162	1.2454	(0.5942)	7.91	***
Sex of household head (1= male)	8,900	0.9418	(0.2341)	5,162	0.9380	(0.2412)	0.91	
Household has a television (1=yes)	8,900	0.1422	(0.3493)	5,162	0.1197	(0.3247)	3.78	***
Household has a radio (1=yes)	8,900	0.2701	(0.4440)	5,162	0.2044	(0.4033)	8.75	***
Household has a bicycle (1=yes)	8,900	0.2108	(0.4079)	5,162	0.1523	(0.3593)	8.56	***
Own marginal land, less than 0.5 acre	8,900	0.5978	(0.4904)	5,162	0.7088	(0.4543)	-13.3	***
Own small land, between 0.5 and 2 acre	8,900	0.2639	(0.4408)	5,162	0.2121	(0.4089)	6.90	***
Own large land, greater than 2 acre	8,900	0.1383	(0.3452)	5,162	0.0790	(0.2698)	10.60	***
Parent's Covariates								
Occupation of father (1= agriculture, 0 = non-agriculture)	8,900	1.4870	(0.4999)	5,162	1.6025	(0.4894)	-13.31	***
Mother's education (highest grade)	8,900	0.2961	(0.7805)	5,162	0.3741	(0.9695)	-5.22	***
Father's education (highest grade)	8,900	0.6090	(1.2116)	5,162	0.6323	(1.3582)	-1.05	
Community Covariates								
Formal school (1= Government school, 0 = informal school)	8,900	0.5917	(0.4915)	5,162	0.2096	(0.4071)	47.23	***
NGO school (1= Non-Government school, 0 = informal school)	8,900	0.0098	(0.0984)	5,162	0.0037	(0.0606)	4.03	***
Household uses piped water (1=yes)	8,900	0.0276	(0.1639)	5,162	0.0347	(0.1830)	-2.35	**
Residence								
Urban (1 = urban, 0 = rural)	8,900	1.7182	(0.4499)	5,162	1.6645	(0.4722)	6.70	***

Notes: Data are from NCLS 2002. Informal school: informal education activities (for example, family education and others). Std. Dev. is standard deviation. *t*-test for difference (working-non-working children).*** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.3: Male and Female Children's Activity, by Urban and Rural Areas

	Urban		Rural	
	Males (%)	Females (%)	Males (%)	Females (%)
Both school and work	20.60	13.34	26.20	16.78
Work	47.66	28.08	47.07	31.50
School	2.14	3.47	2.23	2.45
Idle[a]	29.60	55.11	24.50	49.27
Total	100	100	100	100

Notes: Data are from NCLS 2002. [a]Reportedly involved neither in schooling nor in child labour.

Table 4.4: Work Participation and Attendance Rates of Children Aged 5-17 in Urban and Rural Areas, by Gender

Age	School Attendance				Work Participation			
	Urban		Rural		Urban		Rural	
	Males (%)	Females (%)	Males (%)	Females (%)	Males (%)	Females (%)	Males (%)	Females (%)
5	0.00	0.81	0.53	1.72	1.03	0.81	1.42	2.41
6	0.52	1.60	2.85	3.00	2.09	1.60	3.2	4.29
7	7.84	6.60	5.37	10.95	10.78	3.77	7.32	6.47
8	10.26	15.71	22.28	13.68	20.51	11.43	30.05	13.68
9	14.04	22.50	18.66	25.00	42.11	27.50	48.51	29.17
10	12.03	23.19	17.43	16.23	75.94	30.43	75.00	33.12
11	22.89	29.03	29.79	8.24	79.52	61.29	77.13	65.88
12	49.01	41.26	59.08	49.27	91.09	70.63	91.1	82.04
13	49.82	32.16	54.08	45.79	95.20	89.13	95.62	89.90
14	38.18	36.31	44.44	32.4	91.74	82.12	93.95	88.27
15	11.54	12.41	11.48	3.65	90.38	56.93	89.63	64.96
16	13.24	8.06	13.14	8.06	85.37	64.52	87.74	65.40
17	15.00	6.49	13.69	9.09	88.33	53.25	87.03	60.00
Total	22.74	16.82	28.43	19.23	68.26	41.42	73.27	48.28

Notes: Data are from NCLS 2002. School attendance rate refers to the number of 5-17 years old children attending school expressed as a percentage of total children in this age group.

Table 4.5: The Employment Status of Working Children Aged 5-17 in Urban and Rural Areas, by Gender

	Urban		Rural	
	Males (%)	Females (%)	Males (%)	Females (%)
Paid employee	48.01	27.68	40.24	13.90
Self-employed	8.58	4.53	7.50	2.51
Unpaid employee	43.41	67.79	52.25	83.59
Total	100	100	100	100

Source: Data are from NCLS 2002.

Table 4.6: Employment of Children Aged 5-17 in Urban and Rural Areas, by Gender and Industry

Industry	Urban		Rural	
	Males (%)	Females (%)	Males (%)	Females (%)
Agriculture	21.50	29.03	63.23	70.72
Manufacturing	23.33	42.28	10.11	15.25
Wholesale and Retail	5.39	5.03	3.10	1.67
Construction	44.35	18.12	21.02	8.75
Service	5.44	5.54	2.54	3.60
Total	100	100	100	100

Source: Data are from NCLS 2002.

Table 4.7: IV Probit Estimates of School Attendance (Sample 5-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME
Child's age	0.5874***	0.2010***	0.5855***	0.2003***	0.5763***	0.1973***	0.5750***	0.1967***	-0.1667**	-0.0367**
	(0.0574)	(0.0196)	(0.0574)	(0.0196)	(0.0577)	(0.0198)	(0.0579)	(0.0198)	(0.0827)	(0.0178)
Child's age (squared)	-0.0202***	-0.0069***	-0.0199***	-0.0068***	-0.0199***	-0.0068***	-0.0201***	-0.0069***	0.0024	0.0005
	(0.0027)	(0.0009)	(0.0027)	(0.0009)	(0.0026)	(0.0009)	(0.0026)	(0.0009)	(0.0029)	(0.0006)
Female	-0.5796***	-0.1792***	-0.5957***	-0.1835***	-0.6262***	-0.1919***	-0.6549***	-0.1994***	-0.5854***	-0.1088***
	(0.0629)	(0.0172)	(0.0640)	(0.0173)	(0.0726)	(0.0194)	(0.0795)	(0.0209)	(0.1075)	(0.0170)
Child labour hours	-0.3512***	-0.1202***	-0.3602***	-0.1232***	-0.3479***	-0.1191***	-0.3402***	-0.1164***	-0.3688***	-0.0812***
	(0.0681)	(0.0233)	(0.0685)	(0.0234)	(0.0728)	(0.0249)	(0.0745)	(0.0254)	(0.0906)	(0.0201)
Child labour hours (squared)	0.0040***	0.0014***	0.0041***	0.0014***	0.0039***	0.0013***	0.0038***	0.0013***	0.0043***	0.0009***
	(0.0009)	(0.0003)	(0.0009)	(0.0003)	(0.0009)	(0.0003)	(0.0010)	(0.0003)	(0.0012)	(0.0003)
Residual_child labour hours	0.2423***	0.0829***	0.2514***	0.0860***	0.2433***	0.0833***	0.2368***	0.0810***	0.2689***	0.0592***
	(0.0682)	(0.0233)	(0.0686)	(0.0234)	(0.0728)	(0.0249)	(0.0746)	(0.0255)	(0.0906)	(0.0200)
Residual_child labour hours (squared)	-0.0030***	-0.0010***	-0.0031***	-0.0011***	-0.0030***	-0.0010***	-0.0029***	-0.0010***	-0.0034***	-0.0008***
	(0.0009)	(0.0003)	(0.0009)	(0.0003)	(0.0009)	(0.0003)	(0.0010)	(0.0003)	(0.0012)	(0.0003)
Number of children 0-4									-0.0100	-0.0022
									(0.0283)	(0.0062)
Number of school children 5-17									-0.0412**	-0.0091**
									(0.0182)	(0.0040)
Number of adult males over 17 years									-0.0267	-0.0059
									(0.0261)	(0.0058)
Number of adult females over 17 years									0.0900***	0.0198***
									(0.0335)	(0.0074)
Occupation of father									-0.1764***	-0.0388***
									(0.0508)	(0.0112)
Household uses piped water									-0.0668	-0.0142
									(0.1378)	(0.0283)
Household has a television									-0.0693	-0.0148
									(0.0802)	(0.0167)
Household has a radio									-0.0156	-0.0034
									(0.0469)	(0.0102)

Continued on next page

Table 4.7 (continued): IV Probit Estimates of School Attendance (Sample 5-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME
Household has a bicycle									0.1835***	0.0427***
									(0.0491)	(0.0121)
Formal school									2.5905***	0.4825***
									(0.1329)	(0.0197)
NGO school									2.1600***	0.7196***
									(0.1813)	(0.0415)
Own marginal land, less than 0.5 acre									0.0552	0.0121
									(0.0692)	(0.0150)
Own large land, greater than 2 acre									-0.0024	-0.0005
									(0.0685)	(0.0150)
Urban									0.1341**	0.0295**
									(0.0666)	(0.0147)
Mother more educated than father			0.1970**	0.0704**						
			(0.0893)	(0.0332)						
Mother more educated than father x Female			0.1806	0.0645						
			(0.1709)	(0.0635)						
Mother's education (highest grade)					0.1585***	0.0542***	0.1007**	0.0344**	0.0445	0.0098
					(0.0501)	(0.0172)	(0.0401)	(0.0138)	(0.0393)	(0.0087)
Mother's education x Female					0.1337**	0.0458**	0.0526	0.0180	0.0350	0.0077
					(0.0558)	(0.0191)	(0.0594)	(0.0203)	(0.0697)	(0.0153)
Father's education (highest grade)							0.0818***	0.0280***	0.0180	0.0040
							(0.0271)	(0.0093)	(0.0243)	(0.0054)
Father's education x Female							0.0842**	0.0288**	0.1056**	0.0232**
							(0.0386)	(0.0132)	(0.0443)	(0.0098)
Constant	1.1254		1.2318		1.1624		1.0794		5.1391***	
	(0.8389)		(0.8412)		(0.8575)		(0.8728)		(0.8396)	
Hausman test of endogeneity: $\chi2(2)$	15.55		17.03		19.85		20.73		8.95	
P>$\chi2$	(p = 0.0004)		(p = 0.0002)		(p = 0.0000)		(p = 0.0000)		(p = 0.0114)	
Hansen *J*- statistic (Test of overidentifying restrictions)	7.46		6.85		7.30		4.41		7.60	
	(p = 0.2574)		(p = 0.2465)		(p = 0.2730)		(p = 0.1641)		(p = 0.0223)	
N	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900

Notes: Data are from NCLS 2002. ME is marginal effects of school attendance. Standard errors in parenthesis. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.8: IV Tobit Estimates of GAGE (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME
Child's age	11.4441***	5.0780***	11.2345***	4.9886***	9.6943***	4.3369***	9.1763***	4.1165***	-12.2796***	-4.1712***
	(1.2047)	(0.5326)	(1.1998)	(0.5308)	(1.1539)	(0.5146)	(1.1389)	(0.5094)	(0.8880)	(0.2833)
Child's age (squared)	-0.3693***	-0.1639***	-0.3596***	-0.1597***	-0.3112***	-0.1392***	-0.2976***	-0.1335***	0.4243***	0.1441***
	(0.0502)	(0.0222)	(0.0499)	(0.0221)	(0.0470)	(0.0210)	(0.0463)	(0.0207)	(0.0323)	(0.0104)
Female	-3.3279***	-1.4407***	3.5359	1.6125	-3.7608***	-1.6351***	-3.6223***	-1.5804***	-0.7856	-0.2638
	(0.8301)	(0.3505)	(5.0473)	(2.3634)	(0.9098)	(0.3843)	(0.9695)	(0.4113)	(0.7722)	(0.2564)
Child labour hours	-2.1900***	-0.9718***	-2.2513***	-0.9997***	-1.7102**	-0.7651**	-1.4995*	-0.6727*	-0.6778	-0.2302
	(0.7886)	(0.3499)	(0.7866)	(0.3493)	(0.8036)	(0.3595)	(0.8102)	(0.3635)	(0.5913)	(0.2010)
Child labour hours (squared)	0.0270***	0.0120***	0.0276***	0.0123***	0.0204**	0.0091**	0.0176*	0.0079*	0.0077	0.0026
	(0.0101)	(0.0045)	(0.0101)	(0.0045)	(0.0103)	(0.0046)	(0.0104)	(0.0047)	(0.0076)	(0.0026)
Residual_child labour hours	1.1846	0.5256	1.2568	0.5581	0.8425	0.3769	0.6763	0.3034	0.6944	0.2359
	(0.7905)	(0.3508)	(0.7885)	(0.3501)	(0.8055)	(0.3604)	(0.8121)	(0.3643)	(0.5924)	(0.2014)
Residual_child labour hours (squared)	-0.0181*	-0.0080*	-0.0188*	-0.0084*	-0.0126	-0.0057	-0.0103	-0.0046	-0.0083	-0.0028
	(0.0101)	(0.0045)	(0.0101)	(0.0045)	(0.0104)	(0.0046)	(0.0104)	(0.0047)	(0.0076)	(0.0026)
Number of children 0-4									0.1745	0.0593
									(0.2162)	(0.0734)
Number of school children 5-17									-0.2812**	-0.0955**
									(0.1341)	(0.0456)
Number of adult males over 17 years									0.3414*	0.1160*
									(0.1939)	(0.0658)
Number of adult females over 17 years									0.6704***	0.2277***
									(0.2474)	(0.0841)
Occupation of father									0.0949	0.0323
									(0.3850)	(0.1308)
Household uses piped water									0.6257	0.2161
									(0.9774)	(0.3432)
Household has a television									0.9787*	0.3391*
									(0.5482)	(0.1939)
Household has a radio									0.2211	0.0753
									(0.3402)	(0.1163)

Continued on next page

Table 4.8 (continued): IV Tobit Estimates of GAGE (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME
Household has a bicycle									-0.1033	-0.0350
									(0.3578)	(0.1211)
Formal school									53.5726***	18.2491***
									(1.4866)	(0.3129)
NGO school									-46.4376	-6.0572***
									(0.0000)	(0.1599)
Own marginal land, less than 0.5 acre									-0.6875	-0.2345
									(0.4922)	(0.1686)
Own large land, greater than 2 acre									0.1965	0.0670
									(0.4896)	(0.1677)
Urban									0.4223	0.1434
									(0.4730)	(0.1608)
Mother more educated than father			-6.7313***	-2.9890***						
			(1.2260)	(0.5443)						
Mother more educated than father x Female			-3.5419	-1.5728						
			(2.5703)	(1.1413)						
Mother's education (highest grade)					4.8883***	2.1869***	2.5047***	1.1236***	0.6587**	0.2237**
					(0.5793)	(0.2592)	(0.4844)	(0.2173)	(0.2663)	(0.0907)
Mother's education x Female					2.2199***	0.9931***	2.7614***	1.2387***	0.2999*	0.1019*
					(0.7276)	(0.3255)	(0.3312)	(0.1485)	(0.1673)	(0.0569)
Father's education (highest grade)							2.0024**	0.8983**	1.0690**	0.3631**
							(0.8074)	(0.3622)	(0.4982)	(0.1695)
Father's education x Female							0.0337	0.0151	0.4195	0.1425
							(0.5125)	(0.2299)	(0.3145)	(0.1069)
Constant	-45.1607***		-29.9509**		-40.2175***		-39.7461***		63.4142***	
	(12.4829)		(12.0995)		(11.9742)		(11.8992)		(7.1690)	
Sigma	20.0735***		20.0064***		19.5085***		19.2959***		10.7219***	
	(0.2160)		(0.2152)		(0.2095)		(0.2071)		(0.1046)	
Hausman test of endogeneity: F(2, 8839)	11.70		10.54		5.00		3.87		1.17	
P>F	(P = 0.0000)		(P = 0.0000)		(P = 0.0000)		(P = 0.0000)		(P = 0.0000)	
Hansen *J* -statistic (Test of overidentifying restrictions): $\chi 2(2)$	18.25		17.81		13.95		12.31		9.12	
P>$\chi 2$	(P = 0.1089)		(P = 0.1354)		(P = 0.9330)		(P = 0.2121)		(P = 0.1044)	
N	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848

Notes: Data are from NCLS 2002. ME is marginal effects of GAGE. Standard errors in parenthesis. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.9: Heckman Probit Estimates of School Attendance (Sample 5-17)

Variables	(1) Child work participation	(2) ME	(3) Child work participation	(4) ME	(5) Child work participation	(6) ME	(7) Child work participation	(8) ME	(9) Child work participation	(10) ME
Child's age	1.4084***	0.2476***	1.4086***	0.2479***	1.4066***	0.2463***	1.4059***	0.2461***	1.4057***	-0.0224
	(0.0331)	(0.0222)	(0.0331)	(0.0222)	(0.0331)	(0.0227)	(0.0331)	(0.0229)	(0.0340)	(0.0218)
Child's age (squared)	-0.0505***	-0.0086***	-0.0505***	-0.0085***	-0.0504***	-0.0086***	-0.0503***	-0.0087***	-0.0501***	0.0000
	(0.0014)	(0.0010)	(0.0014)	(0.0010)	(0.0014)	(0.0010)	(0.0014)	(0.0010)	(0.0014)	(0.0008)
Female	-0.6154***	-0.1798***	-0.7744***	-0.0852	-0.6078***	-0.1919***	-0.6019***	-0.1990***	-0.5924***	-0.1092***
	(0.0295)	(0.0170)	(0.2819)	(0.1043)	(0.0314)	(0.0192)	(0.0329)	(0.0207)	(0.0332)	(0.0170)
Child labour hours		-0.1200***		-0.1229***		-0.1188***		-0.1163***		-0.0811***
		(0.0231)		(0.0232)		(0.0247)		(0.0253)		(0.0200)
Child labour hours (squared)		0.0014***		0.0014***		0.0013***		0.0013***		0.0009***
		(0.0003)		(0.0003)		(0.0003)		(0.0003)		(0.0003)
Residual_child labour hours		0.0832***		0.0863***		0.0835***		0.0814***		0.0591***
		(0.0232)		(0.0233)		(0.0247)		(0.0253)		(0.0199)
Residual_child labour hours (squared)		-0.0010***		-0.0011***		-0.0010***		-0.0010		-0.0008***
		(0.0003)		(0.0003)		(0.0003)		(0.0003)		(0.0003)
Number of children 0-4									-0.0031	-0.0021
									(0.0192)	(0.0062)
Number of school children 5-17									0.0328***	-0.0089**
									(0.0121)	(0.0040)
Occupation of father									-0.1826***	-0.0389***
									(0.0339)	(0.0112)
Household uses piped water									0.1003	-0.0139
									(0.0890)	(0.0282)
Household has a television									0.0138	-0.0148
									(0.0464)	(0.0167)
Household has a radio									0.0639*	-0.0034
									(0.0349)	(0.0102)
Household has a bicycle									-0.0190	0.0425***
									(0.0385)	(0.0121)
Formal school									-0.1120***	0.4797***
									(0.0338)	(0.0199)

Continued on next page

Table 4.9 (continued): Heckman Probit Estimates of School Attendance (Sample 5-17)

Variables	(1) Child work participation	(2) ME	(3) Child work participation	(4) ME	(5) Child work participation	(6) ME	(7) Child work participation	(8) ME	(9) Child work participation	(10) ME
NGO school									0.0670	0.7177***
									(0.1555)	(0.0420)
Own marginal land, less than 0.5 acre									-0.0605	0.0119
									(0.0373)	(0.0150)
Own large land, greater than 2 acre									0.0692	-0.0007
									(0.0523)	(0.0150)
Urban									0.0087	0.0296**
									(0.0356)	(0.0146)
Mother more educated than father			0.0290	-0.0680**						
			(0.0842)	(0.0303)						
Mother more educated than father x Female			0.0811	-0.0566						
			(0.1434)	(0.0574)						
Mother's education (highest grade)					-0.0162	0.0531***	0.0022	0.0338**	0.0151	0.0097
					(0.0220)	(0.0171)	(0.0274)	(0.0136)	(0.0285)	(0.0086)
Mother's education x Female					-0.0279	0.0431**	-0.0101	0.0168	-0.0170	0.0077
					(0.0376)	(0.0188)	(0.0456)	(0.0199)	(0.0462)	(0.0153)
Father's education (highest grade)							-0.0229	0.0275***	-0.0179	0.0039
							(0.0182)	(0.0092)	(0.0188)	(0.0053)
Father's education x Female							-0.0170	0.0270**	-0.0216	0.0230**
							(0.0280)	(0.0130)	(0.0284)	(0.0097)
Constant	-8.5328***		-8.5928***		-8.5371***		-8.5505***		-8.3104***	
	(0.1979)		(0.2581)		(0.1980)		(0.1982)		(0.2252)	
Exclusions restrictions										
Sex of household head	0.2702***		0.2737***		0.2856***		0.2999***		0.2768***	
	(0.0566)		(0.0566)		(0.0567)		(0.0569)		(0.0580)	
Number of adults over 17 years	-0.0397***		-0.0400***		-0.0388***		-0.0363***		-0.0441***	
	(0.0120)		(0.0120)		(0.0120)		(0.0120)		(0.0127)	
inverse Mill's ratio		0.3983***		0.4067***		0.4107***		0.4103***		0.1523
		(0.0972)		(0.0982)		(0.1017)		(0.1029)		(0.1239)
LR test of ρ = 0	17.49		17.95		17.06		16.63		1.61	
p>χ2(1)	(p = 0.0000)		(p = 0.0000)		(p = 0.0000)		(p = 0.0000)		(p = 0.0000)	
N	14,062	8,900	14,062	8,900	14,062	8,900	14,062	8,900	14,062	8,900

Notes: Data are from NCLS 2002. ME is marginal effects of school attendance. Standard errors in parenthesis. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.10: Selectivity Adjusted Estimates of GAGE (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE
Child's age	1.7035***	15.8681***	1.7040***	15.8921***	1.7060***	15.2708***	1.7081***	14.9260***	1.7377***	10.5176***
	(0.0462)	(2.6792)	(0.0462)	(2.6763)	(0.0463)	(2.6459)	(0.0463)	(2.6409)	(0.0478)	(2.7560)
Child's age (squared)	-0.0616***	-0.5766***	-0.0617***	-0.5768***	-0.0617***	-0.5575***	-0.0618***	-0.5456***	-0.0625***	-0.3884***
	(0.0018)	(0.0972)	(0.0018)	(0.0971)	(0.0018)	(0.0958)	(0.0018)	(0.0955)	(0.0019)	(0.0981)
Female	-0.6627***	-9.9633***	-0.8691***	-12.7396***	-0.6571***	-10.3028***	-0.6525***	-10.2306***	-0.6486***	-8.4159***
	(0.0306)	(0.9830)	(0.2936)	(3.1479)	(0.0326)	(1.0065)	(0.0341)	(1.0252)	(0.0345)	(1.1177)
Child labour hours		-1.1476**		-1.1650**		-0.9612*		-0.9286*		-0.8882
		(0.5172)		(0.5172)		(0.5367)		(0.5463)		(0.5804)
Child labour hours (squared)		0.0136**		0.0138**		0.0110		0.0106		0.0105
		(0.0066)		(0.0066)		(0.0069)		(0.0070)		(0.0075)
Residual_child labour hours		1.0305**		1.0494**		0.8823		0.8575		0.8599
		(0.5184)		(0.5184)		(0.5378)		(0.5473)		(0.5814)
Residual_child labour hours (squared)		-0.0130*		-0.0132**		-0.0107		-0.0104		-0.0107
		(0.0066)		(0.0066)		(0.0069)		(0.0070)		(0.0075)
Number of children 0-4									0.0023	0.2256
									(0.0201)	(0.2138)
Number of school children 5-17									0.0194	-0.1893
									(0.0126)	(0.1350)
Occupation of father									-0.1711***	-1.5442***
									(0.0351)	(0.4231)
Household uses piped water									0.1277	1.6128*
									(0.0931)	(0.9685)
Household has a television									0.0312	1.3047**
									(0.0477)	(0.5376)
Household has a radio									0.0398	0.7282**
									(0.0360)	(0.3360)
Household has a bicycle									-0.0265	-0.2176
									(0.0397)	(0.3514)

Continued on next page

Table 4.10 (continued): Selectivity Adjusted Estimates of GAGE (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE
Formal school									-0.1640***	-13.2833***
									(0.0346)	(4.1266)
NGO school									0.0261	
									(0.1574)	
Own marginal land, less than 0.5 acre									-0.0603	-1.2793***
									(0.0386)	(0.4869)
Own large land, greater than 2 acre									0.0362	0.7408
									(0.0537)	(0.4811)
Urban									0.0029	0.6601
									(0.0369)	(0.4650)
Mother more educated than father			0.0114	-0.7659						
			(0.0887)	(0.7230)						
Mother more educated than father x Female			0.1054	1.3947						
			(0.1493)	(1.5094)						
Mother's education (highest grade)					-0.0269	0.4819	0.0036	0.5601*	0.0155	0.6968***
					(0.0236)	(0.3736)	(0.0297)	(0.2992)	(0.0303)	(0.2610)
Mother's education x Female					-0.0213	1.0563**	-0.0115	1.0476**	-0.0133	1.0674**
					(0.0395)	(0.4430)	(0.0481)	(0.4796)	(0.0483)	(0.4893)
Father's education (highest grade)							-0.0337*	-0.0533	-0.0243	0.1343
							(0.0191)	(0.2129)	(0.0196)	(0.1653)
Father's education x Female							-0.0104	0.0144	-0.0142	0.1257
							(0.0291)	(0.3138)	(0.0292)	(0.3099)
Constant	-10.4171***	-113.9558***	-10.4441***	-112.5017***	-10.4389***	-111.6896***	-10.4642***	-109.2735***	-10.3535***	-55.8434**
	(0.2858)	(22.8025)	(0.3365)	(22.6787)	(0.2862)	(22.4931)	(0.2865)	(22.4470)	(0.3120)	(22.7719)
Exclusion restrictions										
Sex of household Head	0.2833***		0.2851***		0.2929***		0.3052***		0.2775***	
	(0.0578)		(0.0579)		(0.0581)		(0.0583)		(0.0593)	
Number of adults over 17 years	-0.0356***		-0.0356***		-0.0351***		-0.0332***		-0.0426***	
	(0.0122)		(0.0122)		(0.0122)		(0.0123)		(0.0130)	
inverse Mill's ratio		140.5323***		140.6213***		138.7977***		137.3378***		108.2542***
		(11.8816)		(11.8624)		(11.7061)		(11.6661)		(11.5912)
N	11,547	5,268	11,547	5,268	11,547	5,268	11,547	5,268	11,547	5,268

Notes: Data are from NCLS 2002. NGO school in Column 10 dropped due to collinearity. Robust standard errors are in parenthesis. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.11: Double-Hurdle Estimates of GAGE (Sample 7-17)

Variables	(1) Child work participation	(2) GAGE	(3) Child work participation	(4) GAGE	(5) Child work participation	(6) GAGE	(7) Child work participation	(8) GAGE	(9) Child work participation	(10) GAGE
Child's age	1.7035***	-15.4236***	1.7040***	-15.4339***	1.7060***	-15.7261***	1.7081***	-15.8054***	1.7377***	-15.0215***
	(0.0462)	(1.1001)	(0.0462)	(1.0984)	(0.0463)	(1.0714)	(0.0463)	(1.0667)	(0.0478)	(1.1041)
Child's age (squared)	-0.0616***	0.5464***	-0.0617***	0.5474***	-0.0617***	0.5542***	-0.0618***	0.5556***	-0.0625***	0.5198***
	(0.0018)	(0.0449)	(0.0018)	(0.0447)	(0.0018)	(0.0427)	(0.0018)	(0.0424)	(0.0019)	(0.0401)
Female	-0.6627***	-0.4820	-0.8691***	0.8929	-0.6571***	-1.1342	-0.6525***	-1.3578	-0.6486***	-0.9949
	(0.0306)	(0.6944)	(0.2936)	(3.5549)	(0.0326)	(0.7801)	(0.0341)	(0.8436)	(0.0345)	(0.9541)
Child labour hours		-1.3851**		-1.4023**		-1.1436		-1.0946		-0.9108
		(0.6796)		(0.6789)		(0.6983)		(0.7098)		(0.7444)
Child labour hours (squared)		0.0164*		0.0166*		0.0131		0.0125		0.0106
		(0.0087)		(0.0087)		(0.0090)		(0.0091)		(0.0096)
Residual_child labour hours		1.3001*		1.3189*		1.1047		1.0651		0.9205
		(0.6811)		(0.6804)		(0.6997)		(0.7112)		(0.7457)
Residual_child labour hours (squared)		-0.0163*		-0.0165*		-0.0134		-0.0129		-0.0114
		(0.0087)		(0.0087)		(0.0090)		(0.0091)		(0.0096)
Number of children 0-4									0.0023	0.2159
									(0.0201)	(0.2616)
Number of school children 5-17									0.0194	-0.5352***
									(0.0126)	(0.1639)
Occupation of father									-0.1711***	0.2625
									(0.0351)	(0.4687)
Household uses piped water									0.1277	0.5615
									(0.0931)	(1.1762)
Household has a television									0.0312	1.3036*
									(0.0477)	(0.6674)
Household has a radio									0.0398	0.4733
									(0.0360)	(0.4098)
Household has a bicycle									-0.0265	-0.0434
									(0.0397)	(0.4285)

Continued on next page

Table 4.11 (continued): Double-Hurdle Estimates of GAGE (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE	Child work participation	GAGE
Formal school									-0.1640***	-14.4585***
									(0.0346)	(4.6635)
NGO school									0.0261	
									(0.1574)	
Own marginal land, less than 0.5 acre									-0.0603	-0.8734
									(0.0386)	(0.6040)
Own large land, greater than 2 acre									0.0362	0.6146
									(0.0537)	(0.5846)
Urban									0.0029	0.6120
									(0.0369)	(0.5804)
Mother more educated than father			0.0114	-0.7936						
			(0.0887)	(0.8971)						
Mother more educated than father x Female			0.1054	-0.7188						
			(0.1493)	(1.8211)						
Mother's education (highest grade)					-0.0269	0.8548*	0.0036	0.6000	0.0155	0.6688**
					(0.0236)	(0.4750)	(0.0297)	(0.3724)	(0.0303)	(0.3165)
Mother's education x Female					-0.0213	1.7433***	-0.0115	1.1979**	-0.0133	1.1354**
					(0.0395)	(0.5287)	(0.0481)	(0.5658)	(0.0483)	(0.5752)
Father's education (highest grade)							-0.0337*	0.3242	-0.0243	0.3745*
							(0.0191)	(0.2638)	(0.0196)	(0.1987)
Father's education x Female							-0.0104	0.5794	-0.0142	0.4987
							(0.0291)	(0.3749)	(0.0292)	(0.3679)
Constant	-10.4171***	148.6462***	-10.4441***	150.4115***	-10.4389***	147.5952***	-10.4642***	147.4900***	-10.3535***	153.8724***
	(0.2858)	(11.1280)	(0.3365)	(10.7728)	(0.2862)	(10.7941)	(0.2865)	(10.8124)	(0.3120)	(10.1527)
Exclusion restrictions										
Sex of household head	0.2833***		0.2851***		0.2929***		0.3052***		0.2775***	
	(0.0578)		(0.0579)		(0.0581)		(0.0583)		(0.0593)	
Number of adults over 17 years	-0.0356***		-0.0356***		-0.0351***		-0.0332***		-0.0426***	
	(0.0122)		(0.0122)		(0.0122)		(0.0123)		(0.0130)	
Sigma		11.9150***		11.9092***		11.7560***		11.7223***		11.6247***
		(0.1557)		(0.1556)		(0.1524)		(0.1517)		(0.1497)
N	11,547	5,268	11,547	5,268	11,547	5,268	11,547	5,268	11,547	5,268

Notes: Data are from NCLS 2002. NGO school in Column 10 dropped due to collinearity. Robust standard errors are in parenthesis. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.12: IV Probit Estimates of School Attendance for Child Wage (Paid) Employee (Sample 5-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME
Child labour hours	-0.2504***	-0.0568***	-0.2552***	-0.0579***	-0.2515***	-0.0568***	-0.2490***	-0.0560***	-0.3436***	-0.0283***
	(0.0765)	(0.0173)	(0.0766)	(0.0173)	(0.0766)	(0.0172)	(0.0763)	(0.0171)	(0.0825)	(0.0074)
Child labour hours (squared)	0.0027***	0.0006***	0.0028***	0.0006***	0.0027***	0.0006***	0.0027***	0.0006***	0.0037***	0.0003***
	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0010)	(0.0001)
Residual_child labour hours	0.1931**	0.0439**	0.1983***	0.0450***	0.1948**	0.0440**	0.1918**	0.0431**	0.2834***	0.0234***
	(0.0767)	(0.0173)	(0.0767)	(0.0174)	(0.0767)	(0.0173)	(0.0764)	(0.0171)	(0.0828)	(0.0073)
Residual_child labour hours (squared)	-0.0023***	-0.0005***	-0.0024***	-0.0005***	-0.0023***	-0.0005***	-0.0023**	-0.0005***	-0.0033***	-0.0003***
	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0009)	(0.0002)	(0.0010)	(0.0001)
Mother more educated than father			0.1256	0.0303						
			(0.1558)	(0.0397)						
Mother more educated than father x Female			-0.2141	-0.0432						
			(0.5054)	(0.0894)						
Mother's education (highest grade)					0.1368**	0.0309**	0.1088*	0.0245*	0.0001	0.0000
					(0.0539)	(0.0122)	(0.0567)	(0.0128)	(0.0664)	(0.0055)
Mother's education x Female					0.0865	0.0195	-0.1303	-0.0293	-0.1473	-0.0121
					(0.1305)	(0.0295)	(0.1503)	(0.0338)	(0.1707)	(0.0141)
Father's education (highest grade)							0.0556	0.0125	-0.0251	-0.0021
							(0.0338)	(0.0076)	(0.0418)	(0.0035)
Father's education x Female							0.2888***	0.0650***	0.4216***	0.0347***
							(0.1040)	(0.0234)	(0.1173)	(0.0104)
Constant	3.2971***		3.3710***		3.3259***		3.1264**		7.8897***	
	(1.2736)		(1.2751)		(1.2695)		(1.2539)		(1.3612)	
N	3,246	3,246	3,246	3,246	3,246	3,246	3,246	3,246	3,246	3,246

Notes: Data are from NCLS 2002. ME is marginal effects of school attendance. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for urban areas, occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.13: IV Tobit Estimates of GAGE for Child Wage (Paid) Employee (Sample 7-17)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME
Child labour hours	-0.1401	-0.0805	-0.1600	-0.0920	-0.1268	-0.0731	-0.1336	-0.0773	-0.3126	-0.2249
	(0.5232)	(0.3007)	(0.5212)	(0.2998)	(0.5176)	(0.2986)	(0.5118)	(0.2963)	(0.2980)	(0.2143)
Child labour hours (squared)	0.0037	0.0021	0.0039	0.0022	0.0033	0.0019	0.0032	0.0018	0.0037	0.0027
	(0.0060)	(0.0035)	(0.0060)	(0.0035)	(0.0060)	(0.0035)	(0.0059)	(0.0034)	(0.0035)	(0.0025)
Residual_child labour hours	-0.0327	-0.0188	-0.0131	-0.0075	-0.0328	-0.0189	-0.0251	-0.0145	0.3232	0.2325
	(0.5256)	(0.3021)	(0.5237)	(0.3012)	(0.5200)	(0.3000)	(0.5142)	(0.2977)	(0.2996)	(0.2156)
Residual_child labour hours (squared)	-0.0027	-0.0015	-0.0028	-0.0016	-0.0024	-0.0014	-0.0022	-0.0013	-0.0039	-0.0028
	(0.0061)	(0.0035)	(0.0060)	(0.0035)	(0.0060)	(0.0035)	(0.0059)	(0.0034)	(0.0035)	(0.0025)
Mother more educated than father			-4.0385***	-2.3226***						
			(1.2117)	(0.6973)						
Mother more educated than father x Female			1.8633	1.0716						
			(4.2266)	(2.4308)						
Mother's education (highest grade)					3.1252***	1.8027***	2.0667***	1.1966***	0.1889	0.1359
					(0.4577)	(0.2646)	(0.4736)	(0.2745)	(0.2970)	(0.2136)
Mother's education x Female					0.7876	0.4543	-0.5140	-0.2976	0.7027	0.5055
					(1.1225)	(0.6475)	(1.2621)	(0.7308)	(0.8032)	(0.5778)
Father's education (highest grade)							1.8328***	1.0612***	0.2346	0.1688
							(0.2645)	(0.1536)	(0.1703)	(0.1225)
Father's education x Female							-0.5140	-0.2976	0.7027	0.5055
							(1.2621)	(0.7308)	(0.8032)	(0.5778)
Constant	19.4124**		28.0148***		20.4353**		19.7529**		36.9943***	
	(9.8539)		(10.1947)		(9.7192)		(9.5846)		(5.5386)	
Sigma	12.1936***		12.1762***		12.0886***		11.9810***		7.5887***	
	(0.1518)		(0.1515)		(0.1504)		(0.1491)		(0.0944)	
N	3,236	3,236	3,236	3,236	3,236	3,236	3,236	3,236	3,236	3,236

Notes: Data are from NCLS 2002. ME is marginal effects of GAGE. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for urban areas, occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.14: IV Probit Estimates of School Attendance, by Urban Location (Sample 5-17)

	Urban Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME
Child labour hours	-0.2472***	-0.0785***	-0.2513***	-0.0798***	-0.2504***	-0.0796***	-0.2654***	-0.0845***	-0.2584**	-0.0530**
	(0.0845)	(0.0268)	(0.0847)	(0.0269)	(0.0863)	(0.0275)	(0.0876)	(0.0279)	(0.1048)	(0.0217)
Child labour hours (squared)	0.0026**	0.0008**	0.0026**	0.0008**	0.0026**	0.0008**	0.0028**	0.0009**	0.0027**	0.0006**
	(0.0011)	(0.0003)	(0.0011)	(0.0003)	(0.0011)	(0.0003)	(0.0011)	(0.0004)	(0.0013)	(0.0003)
Residual_child labour hours	0.1397*	0.0444*	0.1436*	0.0456*	0.1491*	0.0474*	0.1659*	0.0528*	0.1574	0.0323
	(0.0846)	(0.0269)	(0.0848)	(0.0269)	(0.0865)	(0.0275)	(0.0878)	(0.0279)	(0.1050)	(0.0216)
Residual_child labour hours (squared)	-0.0017	-0.0005	-0.0018	-0.0006	-0.0018*	-0.0006*	-0.0020*	-0.0006*	-0.0018	-0.0004
	(0.0011)	(0.0003)	(0.0011)	(0.0003)	(0.0011)	(0.0003)	(0.0011)	(0.0004)	(0.0013)	(0.0003)
Mother more educated than father			0.2471*	0.0839						
			(0.1502)	(0.0540)						
Mother more educated than father x Female			-0.0393	-0.0123						
			(0.3169)	(0.0980)						
Mother's education (highest grade)					0.2270***	0.0722***	0.1799***	0.0573***	0.0729	0.0149
					(0.0602)	(0.0192)	(0.0533)	(0.0171)	(0.0604)	(0.0125)
Mother's education x Female					0.0079	0.0025	-0.1156	-0.0368	-0.1476	-0.0303
					(0.0786)	(0.0250)	(0.0996)	(0.0317)	(0.1207)	(0.0248)
Father's education (highest grade)							0.0487	0.0155	0.0128	0.0026
							(0.0423)	(0.0135)	(0.0462)	(0.0095)
Father's education x Female							0.1570**	0.0500**	0.2132**	0.0437**
							(0.0667)	(0.0212)	(0.0837)	(0.0173)
Constant	0.5736		0.6499		0.8820		0.9920		4.9081***	
	(1.2967)		(1.3015)		(1.2681)		(1.2962)		(1.4446)	
N	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508	2,508

Notes: Data are from NCLS 2002. ME is marginal effects of school attendance. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.15: IV Probit Estimates of School Attendance, by Rural Location (Sample 5-17)

	Rural Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME	IV probit	ME
Child labour hours	-0.2289***	-0.0804***	-0.2351***	-0.0825***	-0.2008**	-0.0706**	-0.1722**	-0.0605**	-0.2633***	-0.0594***
	(0.0804)	(0.0282)	(0.0812)	(0.0285)	(0.0837)	(0.0294)	(0.0847)	(0.0297)	(0.0971)	(0.0220)
Child labour hours (squared)	0.0024**	0.0008**	0.0025**	0.0009**	0.0020*	0.0007*	0.0016	0.0006	0.0030**	0.0007**
	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0013)	(0.0003)
Residual_child labour hours	0.1170	0.0411	0.1238	0.0435	0.0929	0.0326	0.0650	0.0228	0.1622*	0.0366*
	(0.0805)	(0.0283)	(0.0813)	(0.0285)	(0.0839)	(0.0295)	(0.0848)	(0.0298)	(0.0974)	(0.0220)
Residual_child labour hours (squared)	-0.0014	-0.0005	-0.0015	-0.0005	-0.0011	-0.0004	-0.0007	-0.0002	-0.0021*	-0.0005*
	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0011)	(0.0004)	(0.0013)	(0.0003)
Mother more educated than father			0.2617**	0.0967**						
			(0.1262)	(0.0485)						
Mother more educated than father x Female			0.2141	0.0787						
			(0.2071)	(0.0791)						
Mother's education (highest grade)					0.2548***	0.0895***	0.1582***	0.0555***	0.0481	0.0109
					(0.0617)	(0.0217)	(0.0533)	(0.0187)	(0.0554)	(0.0125)
Mother's education x Female					0.1292*	0.0454*	0.0960	0.0337	0.0628	0.0142
					(0.0730)	(0.0257)	(0.0759)	(0.0266)	(0.0896)	(0.0202)
Father's education (highest grade)							0.1548***	0.0543***	0.0305	0.0069
							(0.0300)	(0.0106)	(0.0282)	(0.0064)
Father's education x Female							-0.0024	-0.0008	0.0520	0.0117
							(0.0473)	(0.0166)	(0.0547)	(0.0123)
Constant	-0.4223		-0.3639		-0.6913		-0.9444		4.3595***	
	(0.8997)		(0.9051)		(0.9176)		(0.9229)		(0.9429)	
N	6,392	6,392	6,392	6,392	6,392	6,392	6,392	6,392	6,392	6,392

Notes: Data are from NCLS 2002. ME is marginal effects of school attendance. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.16: IV Tobit Estimates of GAGE, by Urban Location (Sample 7-17)

	Urban Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME
Child labour hours	-0.2356	-0.1492	-0.2508	-0.1591	-0.0695	-0.0448	-0.6503***	-0.4198***	0.2638	0.2141
	(0.6529)	(0.4134)	(0.6496)	(0.4120)	(0.6349)	(0.4097)	(0.1097)	(0.0710)	(0.4082)	(0.3313)
Child labour hours (squared)	0.0013	0.0008	0.0015	0.0010	-0.0005	-0.0003	0.0055***	0.0036***	-0.0040	-0.0032
	(0.0082)	(0.0052)	(0.0082)	(0.0052)	(0.0080)	(0.0051)	(0.0011)	(0.0007)	(0.0051)	(0.0041)
Residual_child labour hours	-0.3166	-0.2005	-0.3012	-0.1910	-0.3626	-0.2340	0.2330**	0.1504**	-0.3538	-0.2872
	(0.6552)	(0.4149)	(0.6519)	(0.4134)	(0.6371)	(0.4111)	(0.1013)	(0.0654)	(0.4097)	(0.3325)
Residual_child labour hours (squared)	0.0031	0.0020	0.0030	0.0019	0.0040	0.0026	-0.0022**	-0.0014**	0.0046	0.0037
	(0.0082)	(0.0052)	(0.0082)	(0.0052)	(0.0080)	(0.0052)	(0.0009)	(0.0006)	(0.0051)	(0.0042)
Mother more educated than father			-4.7377***	-3.0047***						
			(1.2574)	(0.7982)						
Mother more educated than father x Female			0.1270	0.0806						
			(2.9221)	(1.8532)						
Mother's education (highest grade)					4.0527***	2.6152***	2.9098***	1.8785***	1.0440***	0.8473***
					(0.4664)	(0.3024)	(0.4635)	(0.3000)	(0.2668)	(0.2167)
Mother's education x Female					-0.0147	-0.0095	0.9971	0.6437	-0.2290	-0.1858
					(0.6385)	(0.4120)	(0.7073)	(0.4567)	(0.5093)	(0.4134)
Father's education (highest grade)							1.2540***	0.8121***	0.3055	0.2479
							(0.3264)	(0.2116)	(0.1898)	(0.1541)
Father's education x Female							0.7232	0.4684	-0.2290	-0.1858
							(0.5051)	(0.3271)	(0.5093)	(0.4134)
Constant	9.1464		19.3825		15.1199		29.3658***		32.1402***	
	(12.4000)		(12.9322)		(11.5024)		(8.4492)		(6.9660)	
Sigma	12.7672***		12.7235***		12.2710***		12.2590***		7.8150***	
	(0.1809)		(0.1803)		(0.1739)		(0.1737)		(0.1108)	
N	2,497	2,497	2,497	2,497	2,497	2,497	2,497	2,497	2,497	2,497

Notes: Data are from NCLS 2002. ME is marginal effects of GAGE. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 4.17: IV Tobit Estimates of GAGE, by Rural Location (Sample 7-17)

	Rural Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME	IV Tobit	ME
Child labour hours	-1.5214**	-0.9358**	-1.5200**	-0.9361**	-1.1775*	-0.7316*	-0.9227	-0.5762	-0.4600	-0.3504
	(0.6053)	(0.3724)	(0.6056)	(0.3730)	(0.6123)	(0.3804)	(0.6142)	(0.3836)	(0.4120)	(0.3138)
Child labour hours (squared)	0.0191**	0.0117**	0.0190**	0.0117**	0.0144*	0.0089*	0.0110	0.0069	0.0055	0.0042
	(0.0081)	(0.0050)	(0.0081)	(0.0050)	(0.0082)	(0.0051)	(0.0082)	(0.0051)	(0.0054)	(0.0041)
Residual_child labour hours	0.9350	0.5751	0.9438	0.5813	0.6628	0.4118	0.4311	0.2692	0.4731	0.3604
	(0.6069)	(0.3733)	(0.6071)	(0.3739)	(0.6137)	(0.3813)	(0.6156)	(0.3844)	(0.4130)	(0.3146)
Residual_child labour hours (squared)	-0.0136*	-0.0083*	-0.0136*	-0.0084*	-0.0096	-0.0059	-0.0064	-0.0040	-0.0058	-0.0044
	(0.0081)	(0.0050)	(0.0081)	(0.0050)	(0.0082)	(0.0051)	(0.0082)	(0.0051)	(0.0054)	(0.0041)
Mother more educated than father			-3.3205***	-2.0450***						
			(1.0836)	(0.6674)						
Mother more educated than father x Female			-3.5575*	-2.1910*						
			(2.0069)	(1.2362)						
Mother's education (highest grade)					3.1174***	1.9368***	1.5459***	0.9654***	0.4640*	0.3535*
					(0.4735)	(0.2943)	(0.4219)	(0.2635)	(0.2622)	(0.1997)
Mother's education x Female					2.6997***	1.6773***	2.5522***	1.5939***	1.2405***	0.9449***
					(0.6370)	(0.3960)	(0.6865)	(0.4289)	(0.4705)	(0.3585)
Father's education (highest grade)							2.1113***	1.3185***	0.2726*	0.2077*
							(0.2452)	(0.1532)	(0.1426)	(0.1086)
Father's education x Female							-0.1762	-0.1100	0.2584	0.1969
							(0.4187)	(0.2615)	(0.2818)	(0.2147)
Constant	-4.7776		2.1957		-5.8063		-6.7477		27.3481***	
	(8.3301)		(7.7267)		(8.0815)		(8.0030)		(4.6688)	
Sigma	13.3200***		13.2830***		13.0323***		12.8888***		8.6740***	
	(0.1182)		(0.0555)		(0.0530)		(0.0516)		(0.0247)	
N	6,351	6,351	6,351	6,351	6,351	6,351	6,351	6,351	6,351	6,351

Notes: Data are from NCLS 2002. ME is marginal effects of GAGE. Robust standard errors are in parenthesis. In Columns 1-2, the variables included but not reported are child's age (in years) and its square term, sex of child, in Columns 9-10, the variables included but not reported are child's age (in years) and its square term, sex of child, number of children 0-4, number of school children 5-17, number of adult males and females over 17 years, dummies for occupation of the father and dwelling characteristics and facilities enjoyed by the household, dummies for household ownership of land (for example, own marginal land, less than 0.5 acre; own large land, greater than 2 acre) and communal characteristics (for example, presence of formal and NGO schools).*** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Chapter 5

Health Consequences of Child Labour in Bangladesh

5.1 Introduction

While increased attention is being paid to school performance of child workers, the effects of their work activities on their health have not received the same attention. Identifying the health effects of child labour is indispensable because children's health is directly related to their future economic prospects and their welfare in their adult life. It is also important for a policy perspective to identify the hazardous types of child labour, in which the majority of working children are engaged (for example, commercial and small-scale agriculture, construction, deep-sea fishing, domestic chores, mining and quarrying).[72] Child labour in hazardous jobs is subject to acute physical injuries and illnesses, and this figure is not insignificant. In 2000, the International Labour Organisation (ILO) estimated that 170 million of total 350 million working children around the world were working in hazardous jobs that had adverse effects on their safety and health, as well as on their moral development (Huebler 2006). This dismal picture is remarkably significant in developing countries where children working under hazardous conditions account for up to 10 percent of all work-related injuries (Ashagrie 1997). To date, existing evidence on health injuries or illnesses to working children in developing countries is fairly limited and the results are mixed, but it supports the hypothesis that child labour is strongly associated with poor health (Guarcello et al. 2004; Wolff and

[72]Hazardous work by children is any activity or occupation that by its nature or type has, or leads to, adverse effects on the child's safety, health (physical or mental) and moral development. Hazards can also results from excessive workloads, the physical conditions of work and/or work intensity in terms of the duration or hours of work even where the activity or occupation is known to be non-hazardous or safe.

Maliki 2008). However, work-related injuries and fatalities to children are not confined to less-developed countries. For example, there is evidence that children working on farms in the United States often experience agricultural-related injuries (see Fassa 2003 for more details).

A number of studies also examine the effect of child labour on health using objective measures of child's health that are known to be determined early in an individual's life, such as weight-for-age (O'Donnell et al. 2005), height-for-age (Kana et al. 2010; O'Donnell et al. 2005), body-mass index (BMI)[73](Beegle et al. 2009; Kana et al. 2010) and height growth (Beegle et al. 2009; O'Donnell et al. 2005).[74] All of these studies, however, find either little or no correlation between child labour and anthropometric indicators. Kana et al. (2010), for example, report that the impact of child work on the health status of children need not always be negative if they work below the threshold level (i.e. 35 or 45 hours a week). As a case in point, they refer to children working in rural Cambodia. While O'Donnell et al. (2005) find little evidence of negative correlation between child labour and child health when the analysis is restricted to a contemporaneous relationship between child work and anthropometric indicators, the results are opposite when considering the longer-term consequences to the health of the child. O'Donnell et al. (2005) find that past child work has a significant positive effect on current illness after controlling for endogeneity of child labour.

Additionally, empirical literature presents some evidence of the positive impact of child labour on the living standards of families and, hence, on the health of the child (Smith 1999; Steckel 1995). This is consistent with the literature that suggests that a disproportionate share of total household income will be allocated to maintain the strength and health of the most productive members, whether the household is modelled as a single decision making unit or as a collection of bargaining agents (Pitt et al. 1990). In addition, any negative impact of child labour on the individual's health may be obscured by selection of the healthiest individuals into work. O'Donnell et al. (2005), for example, claim this effect when investigating the effect of child labour on the weight-for-age of children in rural Vietnam.

[73] The body-mass index is equal to weight in kilograms, divided by height in metres squared.
[74] Weight-for-age and height-for-age are commonly expressed in the form of Z-scores, which compare a child's weight or height with the weight and height of a similar child from a referenced healthy population. More precisely, a weight-for-age Z score is a child's weight expressed as a number of standard deviations from the mean weight of children of the same age and gender in a well-nourished population. The reference population used is the (WHO recommended) US population.

In this essay, we focus on subjective health assessments by the child, or by a parent on behalf of a child, as we seek to estimate the contemporaneous effect of child labour on a child's self-reported injury or illness.[75] Though self-reports of health are subject to considerable over-, under- and mis-reporting, depending on various circumstances, there is evidence that self-reported health is closely correlated with underlying morbidity, and that such self-reporting is a good predictor of future mortality (Idler and Benyamini 1997; Kaplan and Camacho 1983). Moreover, self-reports of health in general have their own distinct scientific value. For instance, it has been shown that such reports contain information on health status even after conditioning on objective measures of health (Idler and Benyamini 1997). Thus, results from 'subjective' measures should not be viewed as some lower order of evidence. Furthermore, the use of such a measure of one's health can lead us to identify the direct effect of work on child health.

The central issue in estimating the causal effect of child labour on health outcomes is that child labour is endogenous in the health outcome equation arising both from reverse causality and unobserved heterogeneity (O'Donnell et al. 2005; Wolff and Maliki 2008). While the existing literature has addressed this endogeneity issue using different methods that rely on a different set of instruments (for example, adult employment rate, the number of primary buildings, household land holdings, indicators of the commune economy and the local labour market and a proxy for school quality), these identifying variables are frequently hard to justify based on economic theory.

In this essay, we examine the effects of child labour on child injury or illness using the bivariate probit approach, because child labour and health outcomes may be determined simultaneously, following O'Donnell et al. (2005) and Wolff and Maliki (2008). In addition, we adopt the semi-parametric approach to understand the association between working hours and subjective child health. We implement these methods by exploiting the ILO-sponsored Child Labour Survey (NCLS) conducted in Bangladesh in 2002-2003, which was specifically designed to provide data on the relationship between children's work and health outcomes.

[75] Data limitations prevent us from incorporating anthropometric indicators. However, though anthropometric indicators have the advantage of objectivity, they also have certain limitations. A particular problem with the use of anthropometric indicators in the context of child labour is that they are better measures of nutrition and health experience at younger ages when child labour is not prevalent, and the anthropometric indicators such as height growth is more appropriate for studying the longer-term health consequences of child labour.

Research on health outcomes of child labour in Bangladesh is severely limited, and most of the existing studies on child labour mainly investigate whether child work is a deterrent or a complement to school attendance and/or enrolments (see, for example, Amin et al. 2004; Khanam 2008; Ravallion and Wodon 2000; Shafiq 2007). The exceptions include Guarcello et al. (2004), who, using the same dataset, found that the number of working hours had a significant effect on the probability of injury. It is worth stressing, however, that their results are limited in two important respects. First, they do not scrutinise the possible endogeneity of child labour hours. In a model of child health, both child working hours and health outcomes may be determined simultaneously. If so, treating child labour hours as exogenous could result in biased estimates. Second, the authors do not include illnesses that would have occurred through child work.

This essay differs from the Guarcello et al. (2004) study in the following ways. First, by acknowledging the multidimensional nature of injury or illness, we examine different types of work-related injury or illness. We apply the bivariate probit approach to explore the effect of work on subjective child health, considering the endogeneity problem of child labour. Second, we investigate the relationship between working hours and injury or illness. An indicator of work participation masks the effect of different degrees of work intensity. Although working hours are only an indirect measure of work intensity, long working hours undoubtedly pose health risks and therefore also merit consideration in examining the effect of child labour hours on health status. We use Robinson's (1988) semi-parametric regression estimator, treating child working hours as endogenous. Third, in a further analysis we study the effect of child work on subjective health in rural areas and across age groups. Fourth, we investigate whether a relationship exists between work heterogeneity of child work and health status. In doing so, we examine the effect of hours on health in different sectors by using the semi-parametric specification. Finally, following Guarcello et al. (2004), we extend our analysis to study the severity of injury or illness by using a proxy measure, that is, we utilise information on whether children receive any medical treatment. In doing so, we again tested the endogeneity of child labour hours which Guarcello et al. (2004) did not consider. Here, we follow Kana et al. (2010) and apply a method proposed by Ravallion and Wodon (2000).

Our empirical analysis reaches three major conclusions. First, we find evidence that child work has a negative impact on subjective child health when we correct for potential sources of endogeneity bias in the bivariate probit model. These conclusions persist even when we consider child labour hours, restrict our analysis to rural children, and split the sample by sectors of employment. Second, we find strong evidence for poor health among younger children, while some evidence for health disadvantages among relatively older children has also been documented. Third, our results show that the severity of injury or illness should also be considered when examining the effect of child labour on health status, as the intensity of injury or illness is significantly higher in construction and manufacturing than in other sectors.

5.2 Features of Child Labour in Bangladesh

In spite of legislation and cultural norms, children are relatively less protected in Bangladesh. Though child work in Bangladesh has deep cultural roots and is often considered to have positive effects on the development of the child, the country has enacted a series of legislation that restricts the conditions under which a child can work. At present, there are 25 special laws and ordinances to protect and improve the status of children in Bangladesh (Khanam 2006). Some believe, however, that there is a lack of harmony among laws that uniformly prohibit the employment of children or set a minimum age for employment. Under the current law, the legal minimum age for employment is between 12 and 16, depending on the sector. However, Bangladesh Export Processing Zones Authority (BEPZA) has restricted the minimum age to 14 for employment in EPZs. Furthermore, since 1990, primary school education has become compulsory in Bangladesh, and the country has adopted school subsidy provision to improve schooling and thereby attract and retain children. However, previous literature has shown that participation in the child labour force may not be responsive to education-related policy measures.[76]

The National Child Labour Survey (NCLS) 2002-2003 found that 7.9 million children between the ages of 5 and 17 are working and that 8 percent of the working children between the ages of 5 and 17 are hurt or become sick due to work.[77] These child workers often are found to work long hours in a variety of hazardous occupations and

[76] See Ravallion and Wodon (2000) for more details.
[77] All these statistics are based on the Report on the National Child Labour Survey 2002-2003 (NCLS 2002), Bangladesh Bureau of Statistics, Dhaka, Bangladesh.

sectors that have the potential to seriously damage their health (for example, in bidis,[78] manufacturing, construction, tanneries, and the seafood and the garment industries). Children also work in informal sectors and small-scale firms, which are, by nature, difficult to regulate. Most children who work in these environments are not given protective clothing or equipment, or the clothing provided has generally been designed for adults and is therefore useless for children.

5.3 Data and Descriptive Statistics

This study uses individual-level data for 2002-2003 from the second National Child Labour Survey (henceforth, NCLS 2002) conducted by the Bangladesh Bureau of Statistics (BBS) within the framework of an Integrated Multipurpose Sample Design (IMPS). The NCLS 2002 included a child population between the ages of 5 and 17 from 40,000 households. However, the NCLS 2002 excluded children living on the streets or in institutions such as prisons, orphanages or welfare centres. The dataset contains information on a range of individual (age, gender, marital status, educational attainment, employment status, hours worked and wages earned) and household-level attributes (household size and composition, land holding, location, asset ownership). In addition, the NCLS 2002 includes information on health and safety for every child of the household engaged in economic activities[79] that are relevant to the present study. The health assessment data includes self-reported illness and injuries for those children who are engaged in economic activities. Specifically, the question used to define a work-related injury or illness in NCLS 2002 is 'Has the child ever experienced any injury or illness due to work?' The survey, however, did not clearly define the reference period for the self-reported injury or illness. That is, it is unclear if the reference period for injury or illness was last year, last week, or indeed at any time in the past. Nine health categories are included in the survey questionnaire, including eye/ear infection, skin infection, stiff neck or backache, problems of stomach or lung disease, tiredness/exhaustion, burns (any type), body injuries, loss of limbs and others. The

[78] A bidi is a type of small, hand-rolled cigarette.

[79] Economic activity contains all market production and certain types of non-market production including production and processing of primary products for own consumption and production of fixed assets for own use. It excludes unpaid activities, such as, unpaid domestic activities and voluntary community services.

respondents were explicitly asked whether they had experienced each one of these nine injuries or illnesses.

We focus on child workers between the ages of 5 and 17 who work as paid employees (paid in cash or in kind), who are self-employed or who work as unpaid employees (for example, who work on the family farm or in the family business) related to the household head. Following Beegle et al. (2009), we include children who are enrolled at school to avoid the issue that child labour can affect contemporaneous schooling decisions.[80] However, we cannot include children performing domestic chores as the NCLS dataset does not collect any information on injury or illness directly related to domestic chores. In addition, children with missing ages, work and/or health variables are excluded. Therefore, the analysis is based on 16,010 children. In this sample, 77 percent (12,363) are male children and 23 percent (3,647) are female children. Of this sample of 16,010 children, nearly 90 percent (14,437) are economically active.

We examine two health indicators as dependent variables for this analysis. The first indicator is whether a child reports any work-related injury or illness. The second indicator is whether a child reports any work-related symptoms of injury or illness. For both measures, we generate a binary variable, which takes a value of 1 if a child reports any injury or illness or symptoms of injury or illness, and 0 otherwise. The health complaints or symptoms of injury or illness used in our setting are divided into four categories: tiredness/exhaustion, backache, body injury (including 'loss of limbs') and other health problems (for example, infection, burns and lung diseases).[81] Correlations between different forms of injury or illnesses that are used in this essay are presented in Table 5.1.

We consider two different measures of child labour. The first measure is a dummy variable indicating whether the child is simultaneously employed and enrolled in school. The second measure is the number of hours worked by the child in the reference week during which the child was employed. We include a rich set of covariates that are intended to control for both individual and household characteristics that may affect health outcomes and child labour choices. Individual characteristics include the child's

[80] In doing so, we may identify a 'pure' child labour effect among the sample of children who work. At this point, it should be noted that the selection of only children enrolled in school may induce a selection bias. *A priori*, this selection bias is expected to attenuate our findings.
[81] Infection includes 'eye/ear' and 'skin' infections.

age and a quadratic of the child's age (Guarcello et al. 2004; Kana et al. 2010),[82] the child's gender, the child's vaccination status and the child's sector of employment. Sectors of employment may capture the type of hazards to which the child worker is exposed. We consider in our analysis the main sectors of employment, i.e. agriculture, manufacturing, wholesale and retail and construction. With respect to health outcomes, construction work appears to be the most hazardous form of child labour because of the use of dangerous tools and machinery and exposure to falling objects (see Guarcello et al. 2004 for more details). As it is likely that gender bias, if any, may change with age (as older girls may have to care for younger siblings), we use the interaction between the female dummy variable and age. At the household level, parental age and education, household composition, dwelling characteristics and facilities enjoyed by the household are included. The remaining measure includes a dummy variable indicating urban residence to control for differential labour markets of children and their parents. Definitions and descriptive statistics for all regressors are given in Table D1 in Appendix D based on child work status (i.e. working and non-working children).

In Table 5.2 we illustrate health conditions by gender and by work status. The results in Table 5.2 show that the intensity of health complaints varies by gender and labour participation. Working children tend to have more health complaints than non-working children. The activities of working children are, therefore, more likely to be disrupted due to their health problems. The difference is statistically significant at the 1 percent level.[83] In addition, working male children tend to have more complaints than working female children and the difference is generally statistically significant at conventional levels of significance. Approximately 21 percent of working male children experience an injury or illness due to work; the corresponding number for female children is only 6 percent.

In Figure 5.1, we demonstrate the link between the number of working hours and poor health in Bangladesh. The reported number of health complaints declines when moving from 1-14 hours to 15-29 hours, but it then rises for each subsequent set of hours. For both male and female children, there is a significant increase in reported

[82] In the health equation, the child's age is included to capture the notion that some health conditions may be age related, while in the work equation age will determine the opportunity cost of the child's time. The child's age squared is included to capture a non-linearity in the age effect.
[83] We computed this result using a standard *t*-test.

health complaints when children move from the 15-29 hours per week range to the 43-50 hours per week range, and male children report more injuries or illnesses than do female counterparts. Moreover, it is worth noting that because we are not controlling for the characteristics of the individual, we are therefore unable to draw a more complete conclusion concerning the effect of the number of working hours on poor health.

One of the primary objectives of this essay is to determine the health consequences of child labour across the employment sector. In a developing country such as Bangladesh, there is a lack of homogeneity in child labour across a sector of employment. Thus, one presumes that its health consequences similarly lack homogeneity. Furthermore, given the legislative framework in Bangladesh, one would expect that there would be children of different ages across the sector. This is evident in the data. Overall, approximately 61 percent of the working children (aged 5-17) are in agriculture (see Table 5.3). This is not surprising given the economic activities represented in agricultural sector (livestock, fishery, daily work for poor wages, and unpaid family businesses). Work in wholesale and retail is the second most common form of child work, with 21 percent of working children engaged in this sector, while relatively few children work in construction (3 percent). The mean age of children employed in agriculture, manufacturing and wholesale and retail is 13 years, while the mean age is 14 years for those in construction and service, respectively (see Table 5.3). The sample statistics further show that approximately 45 percent of the youngest children (aged 5-9) is likely to be in agriculture. This proportion drops to approximately 27 percent in wholesale and retail and 22 percent in manufacturing. At the same time, the proportion of the oldest children (aged 14-17) is also high in agriculture at approximately 54 percent. The corresponding proportions for the oldest children are 25 percent in wholesale and retail and 11 percent in manufacturing.

Table 5.3 also shows that the proportion of children reporting any injury or illness is highest in agriculture (49 percent) followed by manufacturing (23 percent). The reason might be related to the fact that children in agricultural activities in developing countries are often involved in applying pesticides and/or operating machinery, and Bangladesh is not an exception. With respect to symptoms of injury or illness, approximately 61 percent of children experienced tiredness/exhaustion in agriculture. The corresponding numbers in manufacturing and wholesale and retail are approximately 18 percent and 12 percent, respectively. While approximately 30 percent of children report body injuries in manufacturing, the corresponding number in agriculture is approximately 20 percent. These results demonstrate that heterogeneity of child work over different sectors have different impacts on child health.

5.4 Estimation Framework

5.4.1 Model of Work-Health Relationship

We first explore the effect of child work participation on health outcomes. The health status equation and the labour market outcome can be expressed as follows:

$$\mathcal{H}_n^{\ *} = \pi_0 + \pi_1 \mathcal{Q}_n + \pi_2 \ell_n + \epsilon_n; \quad n = 1, \ldots, \mathcal{N} \qquad (5.1)$$

$$\ell_n^{\ *} = \beta_0 + \beta_1 \mathcal{Q}_n + \varepsilon_n; \qquad n = 1, \ldots, \mathcal{N} \qquad (5.2)$$

where $\mathcal{H}_n^{\ *}$ and $\ell_n^{\ *}$are latent values of health status and labour choice of child n. In all the estimates, $\mathcal{Q}_n$ is a vector of observable characteristics for child n, which are assumed to be predetermined to health outcomes and child labour choice. These include the following characteristics: the child's age and a quadratic of the child's age, the child's gender, the interaction between the female dummy variable and the child's age, parental age and education, household composition, dwelling characteristics and facilities enjoyed by the household. Additionally, in the health status equation, we control for the child's vaccination status, the child's sector of employment, the child's protection at the workplace, and housing/hygiene condition (safe drinking water and satisfactory sanitation). The coefficient π_2 represents the contemporaneous association between work and health outcomes and ϵ_n and ε_n are random factors.

In practice, however, we do not observe the latent health status $\mathcal{H}_n^{\ *}$(which is a self-reported illness or injury or occurrence of symptoms of injury or illness), but the data provides information on its observed counterpart, which we denote by $\mathcal{H}_n$. As we are only aware of the occurrence of injury or illness, we have $\mathcal{H}_n = 1$ when the child says he or she is injured or ill or has any symptoms of injury or illness ($\mathcal{H}_n^{\ *} > 0$) and $\mathcal{H}_n = 0$ otherwise ($\mathcal{H}_n^{\ *} < 0$). On the other hand, it is important to note that the child labour choice is the observed one and not its latent counterpart in the child health equation. Therefore, we have $\ell_n = 1$ if $\ell_n^{\ *} > 0$ and $\ell_n = 0$ otherwise if $\ell_n^{\ *} < 0$.

Thus, the estimating equations are:

$$\mathcal{H}_n = \pi_0 + \pi_1 \mathcal{Q}_n + \pi_2 \ell_n + \epsilon_n; \quad n = 1, \dots, \mathcal{N} \tag{5.3}$$

$$\ell_n = \beta_0 + \beta_1 \mathcal{Q}_n + \varepsilon_n; \quad n = 1, \dots, \mathcal{N} \tag{5.4}$$

One potential problem in estimating Equation (5.3) is that when including only individual health conditions, there may be omitted variable bias if there is co-morbidity. In this case, omitted health conditions are correlated with the included health condition and with self-reported health. For this reason, we also adopt a specification summarising any injury or illness. As outlined above, we use a dummy variable for having any injury or illness.

In addition, despite the inclusion of $\mathcal{Q}_n$, there is a strong reason to remain concerned about the potential endogeneity of the child labour variable in the health outcome of Equation (5.3), as it is not reasonable to assume that corr$(\epsilon, \varepsilon | \mathcal{Q}_n = 0)$. First, if child labour and health outcomes are determined simultaneously, a reverse causal pathway is possible. Some recent evidence for this reverse causality is O'Donnell et al. (2005), who argue that a health shock may derive from a workplace accident or be the accumulated effect of past work experience. Second, child work could be correlated with unobserved factors (such as unobserved personal traits or parental preferences) that are related to health outcomes, which are undetermined *a priori* (O'Donnell et al. 2005). In $\mathcal{Q}_n$, we include control for factors that may affect health outcomes directly and also may affect current work status through parental preferences. We have not been able to completely account for unobserved effects; and thus, these factors remain in the error terms of Equations (5.3) and (5.4). Third, a child's current health status depends on the child's initial endowment of health, and gross investment (and thus inputs used to produce investments) in all previous periods (Grossman 1972). In $\mathcal{Q}_n$ we include control for factors that may affect current health status through prior health investment, such as the child's gender (Burgess et al. 2004). However, it is possible that this factor may not completely account for such effects; therefore, these factors remain in the error terms of Equations (5.3) and (5.4).

In the case of a binary labour market outcome, we address the simultaneity bias by estimating Equations (5.3) and (5.4′), with the recursive bivariate probit model.[84] This approach models the corr$(\epsilon, \varepsilon | \mathcal{Q}_n)$ explicitly by using a full information maximum

[84] For recent studies using the bivariate probit model, see O'Donnell et al. (2005) and Wolff and Maliki (2008).

likelihood strategy. The bivariate probit model assumes that the error terms ϵ_n and ε_n in Equations (5.3) and (5.4′) are jointly distributed as bivariate normal with means zero, variance one and correlation ρ and the equations are estimated simultaneously using the maximum likelihood method:

$$\mathcal{H}_n = \pi_0 + \pi_1 \mathcal{Q}_n + \pi_2 \ell_n + \epsilon_n; \;\; n = 1, \ldots, \mathcal{N} \qquad (5.3)$$

$$\ell_n = \beta_0 + \beta_1 \mathcal{Q}_n + \beta_2 \mathcal{V}_n + \varepsilon_n; \quad n = 1, \ldots, \mathcal{N} \qquad (5.4')$$

$$\begin{bmatrix} \epsilon \\ \varepsilon \end{bmatrix} \sim \mathcal{N} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right] \qquad (5.5)$$

Although the bivariate probit model can be identified without an exclusion restriction, this strategy is not considered credible in the empirical literature.[85] Thus, our approach is to follow prior research and include a set of variables $(\mathcal{V}_n)$ in the child labour equation but exclude them from the health status equation. We use the migration status of the household, an interaction between the migration status and the location (rural or urban areas) of the household and a proxy for school quality as identifying variables. These instruments are justified in Section 5.4.3.1.

5.4.2 Model of the Hour-Health Relationship

In this sub-section, we extend our analysis to the case of hours worked. Representing child work activity through a simple participation dummy variable may obscure any variation in the work effect with the duration of work. Most of the studies on child labour and child health used specifications in which health outcomes are linear in terms of hours worked (Guarcello et al. 2004; Wolff and Maliki 2008). Recent evidence, however, shows that the effect of hours worked is not linear for different health outcomes. See, for example, Kana et al. (2010). An alternative to the linearity assumption is to find the correct specification using a fully non-parametric approach. A fully non-parametric approach has the advantage that misspecification is, by definition, avoided. This approach, however, becomes infeasible in practice if the dimension of $\mathcal{Q}_n$

[85] In general, the bivariate probit model does not require the exclusion restrictions. The linearity and normality assumptions of the model are sufficient for the identification of the model. However, Altonji et al. (2005) have shown that bivariate probit models without exclusion restrictions can sometime produce misleading results that are consistent with a powerful instrumental variable.

is large and the number of observations is limited. Non-parametric estimators then suffer from the curse of dimensionality due to the slow rate of convergence of the estimator.

In this essay, the size of our sample is 16,010 observations while the dimension of $\mathcal{Q}_n$ is rather large, which makes a completely non-parametric estimator infeasible. In this study, we use Robinson's (1988) semi-parametric estimator to allow for a flexible functional form relationship between hours worked and health outcomes and, at the same time, to avoid the curse of dimensionality. More specifically, the health status equation has the following form:

$$\mathcal{H}_n = \pi_0 + \pi_1\mathcal{Q}_n + \mathcal{F}(\ell_n) + \epsilon_n; \;\; n = 1, \dots, \mathcal{N} \qquad (5.6)$$

where ℓ_n is now the number of hours worked during the reference week (one week before the survey) that enters the equation non-linearly according to a non-binding function $\mathcal{F}$. To control for confounding effects, we include the (log) of weekly hours worked. The health status equation includes all the controls ($\mathcal{Q}_n$) that were used in the bivariate probit specification. We estimate Equation (5.6) using Robinson's partial linear regression models.[86]

The partial linear regression model is estimated using the stepwise procedure of Robinson (1988). Taking conditional expectations given ℓ_n in Equation (5.6) and subtracting these on both sides of the equation gives us:

$$\mathcal{H}_n - E(\mathcal{H}_n|\ell_n) = \big(\mathcal{Q}_n - E(\mathcal{Q}_n|\ell_n)\big)\pi_1 + \epsilon_n; \;\; n = 1, \dots, \mathcal{N} \quad (5.7)$$

In the first step, the conditional means $E(\mathcal{H}_n|\ell_n)$ and $E(\mathcal{Q}_n|\ell_n)$ are estimated non-parametrically, using univariate kernel regressions. In the second step, π_1 is estimated by OLS in Equation (5.7), after replacing the conditional means by their estimates. Robinson shows that the resulting estimator for π_1 is $\sqrt{n}$-consistent and asymptotically normal.

There is some concern, however, that ℓ_n is endogenous in the health status equation (see, for example, Kana et al. 2010). If $E(\epsilon|\ell_n, \mathcal{Q}_n) \neq 0$, the above estimators

[86] It is common to use linear probability models where we treat a binary outcome variable as a continuous one. See, for example, Reinhold and Jürges (2011), who employed linear probability models in their study of parental income and child health in Germany.

will not be consistent. To take the potential endogeneity of ℓ_n into account, we use the augmented regression technique proposed by Holly and Sargan (1982). Assume that:

$$\ell_n = \beta_0 + \beta_1 \mathcal{Q}_n + \beta_2 \mathcal{V}_n + \varepsilon_n; \;\; n = 1, \ldots, \mathcal{N} \quad (5.8)$$

$$\text{with} \quad E(\varepsilon | \mathcal{V}_n, \mathcal{Q}_n) = 0 \quad (5.9)$$

$$\text{and} \;\; E(\epsilon | \ell_n, \mathcal{V}_n, \mathcal{Q}_n) = z\varepsilon \quad (5.10)$$

Then the health status equation (Equation 5.6) can be rewritten as:

$$\mathcal{H}_n = \pi_0 + \pi_1 \mathcal{Q}_n + \mathcal{F}(\ell_n) + z\varepsilon + \tilde{\epsilon}_n; \;\; n = 1, \ldots, \mathcal{N} \quad (5.11)$$

$$\text{with} \quad E(\tilde{\epsilon} | \ell_n, \mathcal{V}_n, \mathcal{Q}) = 0 \quad (5.12)$$

Because ε_n is not observed, we estimate Equation (5.8) by OLS and obtain the residual $\hat{\varepsilon}_n$, which is the consistent estimate of ε_n. Note that in this model $(\mathcal{Q}_n)$ includes the following set of controls: the child's age (in years) and its square term, sex of a child, the interaction between the child's age and sex of a child, the child's vaccination status, dummy variables for sector of employment and urban areas, age of parents, household composition, dummy variables for parental education and for household lighting. The instruments $(\mathcal{V}_n)$ in Equation (5.8) are the same as those used for the bivariate probit specification. Equation (5.11) will now be applied with ε_n replaced by $\hat{\varepsilon}_n$. Estimation of Equations (5.6) to (5.12) uses data on 14,437 individuals, who report positive working hours. We dropped the observations for zero working hours because the logarithm of zero is undefined. However, doing this may lead to sample selection bias, but we address this estimation bias in Section 5.6.1.

5.4.3 Instruments

The challenge inherent in implementing either the bivariate probit or the semi-parametric methods requires the existence of at least one exogenous variable that is significant with the determinants of child labour, but that is not directly related to the probability of being injured or ill. Some examples of identifying variables used in prior work are household land holdings, indicators of the commune economy and the local labour market, such as the rice price and migrant ratio, or a proxy for school quality

(O'Donnell et al. 2005); local adult employment rate and the number of school buildings (Wolff and Maliki 2008); dependency ratio of the household, household possessions such as agricultural land and cattle, and the number of cattle per herd (Kana et al. 2010), commune-level rice price and natural disasters (Beegle et al. 2009). These identifying variables cannot be used in this essay for both conceptual and empirical reasons. For example, in the case of agricultural land holdings and the number of cattle, it is not clear how household land holdings and numbers of cattle are valid instruments in the present context as we do not confine our analysis only to rural working children. Ideally, possession of productive assets, such as agricultural land, livestock and other farm animals, is an important determinant of child labour in rural areas of developing countries where child labour increases the returns of these assets relatively inexpensively (Cockburn and Dostie 2007). While previous studies have found that agricultural land holdings show a strong correlation with child labour, the analysis presented in this study indicates that this variable may not be exogenous.[87]Among the remaining instruments, we use a dependency ratio constructed on the basis of household-level information. Interestingly, we do not find any relevance of this imputed variable to the determination of child work.

We experimented with several possible instruments that meet the relevance and exogeneity conditions and that were available in the data. After several attempts, three instruments were used in the model. Table 5.4 documents lists and definitions of instruments that are used in this essay. The first is a dummy variable indicating the migration status of the household if the household leaves the usual place of residence to find work. The migration status of the household has often been used as an instrument for child work based on the argument that living standards and child work will be influenced by the condition of the economy and the labour market where the household lives. It is, therefore, necessary to construct an interaction term between the migration status and the location of the household. This is a second instrument. The other instrument is a proxy for school quality. The quality of schooling is a potentially important determinant of child labour. For the school quality measure, we generate a binary variable which is equal to 1 if the child reports that his source of education is an

[87]We estimate bivariate probit models (not shown) for different health conditions and include total household land holdings in both the health status and labour market equations. Controlling for all regressors, total household land holdings appear to have significant influence on child health outcomes except for other health problems (z = -1.19, p = 0.233).

informal school and 0 otherwise.[88] In the case of school education in an informal school, it is reasonable to assume that it may not directly affect the intensity of injury or illness. This informal schooling could be used as a good predictor of child labour, as it is well known in Bangladesh that this kind of education is of lower quality compared to formal schools. The relevance of these instruments is verified in the following section.

5.4.3.1 Checking the Validity of the Instruments

We consider several specification tests that examine the statistical performance of the instruments for the work equation in the bivariate probit specification. As with the bivariate probit model, a likelihood ratio (LR) test was run, and the results indicate that the models with and without the identifying restrictions are significantly different from each other. In the first health indicator (any injury/illness), the value of $\chi^2(3)$= 9.14 with a p-value of 0.0275. In the case of different health conditions (symptoms of injury/illness), the corresponding values are $\chi^2(3)$ = 5.52, with a p-value of 0.0634 (tiredness/exhaustion); $\chi^2(3)$= 11.58 with a p-value of 0.0090 (body injuries); $\chi^2(3)$= 11.77 with a p-value of 0.0082 (backache) and $\chi^2(3)$ = 11.09 with a p-value of 0.0113 (other health problems). Over-identification is checked, somewhat informally, for the bivariate probit specification. At first, we run bivariate probit models for the health outcome that include three identifying variables (the migration status of the household, an interaction term between the migration status and the household location and the school quality) in both the health status and labour market equations. Interestingly, all three variables were statistically significant predictors of health outcomes at the 5 percent level, which reduces confidence in our identification strategy in all health models. However, the exclusion restriction is not rejected if we use only the school quality variable to identify the model and include the migration status and an interaction term between the migration status and the household location in the health outcome equation (except for reporting any injury or illness, body injuries and backache). The estimates for the work coefficient are fairly robust to variations on the identification strategy.

[88] NCLS 2002 collects information on types of schools, such as the formal public schools, the NGO schools and an informal school.

In partial linear regression models, we estimate treating working hours as endogenous and include the migration status of the household and the school quality in the instrument set, but we drop an interaction term between the migration status and the household location, as these are not significant determinants of working hours. The relevance of the remaining instruments is verified with empirical tests. The relevant test lends strong credence to our use of two identifying variables.[89] In addition, the Hansen test for over-identification indicates that the instruments are valid in the sense that their influence works only through the endogenous variable but not for all of the health conditions that we considered.[90] Instead, we focus on the partial linear model estimates for the main results of this essay and provide specification test results for the parametric against the partially linear model as a reference (see Footnote 96).

Having established that the instruments have power in the first stage (see Column 1 of Table D2 in Appendix D for the first-stage estimates), we next consider the possibility that instruments may be correlated with an omitted variable. For example, the migration status of the household could drive changes in household income or the quality of schooling. Although it is not possible to test the validity of the instrument with respect to all of the potentially excluded variables, we can examine their correlation with a range of relevant variables that are observed. In Table 5.5, we consider whether household migration influences household income. If migration significantly predicts household income, this would suggest that migration is associated with some structural feature of the location (such as economic activities) and thereby potentially violate the exclusion restriction. We do not find any significant relationship of this. This suggests that correlation between migration and household income should not explain the effect of child labour on child's health. Furthermore, there is no evidence that migration is correlated with youth development training and the quality of schooling. Finally, we confirm that migration is not correlated with the incidence of injury or illness among children. Overall, these results support our use of the migration status of the household as a valid instrument for child labour.

---

[89] We perform an *F*-test such that the coefficients on the instruments are jointly zero. The first stage F-statistics is 4.53 with a negligible p-value of 0.0108. The value of R-squared is 0.27, indicating that the instruments add significantly to the prediction of the (log) of the number of working hours.

[90]The Hansen test for over-identifying restrictions gives a $\chi^2(2)$ test statistic of 5.49 (p-value = 0.0191) for reporting any injury or illness; 1.08 (p-value = 0.2983) for tiredness/exhaustion; 0.3009 (p-value = 0.5833) for body injuries; 0.1039 (p-value of 0.7472) for backache; 4.48 (p-value of 0.0394) for other health problems.

5.5 Empirical Results

Table 5.6 presents the results of the recursive bivariate probit model. As a reference, we have also provided the estimates gained from the univariate probit model. It is clear that the exogeneity of child work is rejected in the univariate probit model at any reasonable levels of significance in all health conditions except for body injuries and other health problems. This suggests that there is no advantage of the univariate probit model over the bivariate probit model in this analysis. This is confirmed by a Smith-Blundell test in the univariate probit model.

The univariate probit estimates indicate a positive and significant relationship between current injury or illness and child work. This relationship indicates that labour force participation is associated with poor health. The result persists when we turn to different symptoms of injury or illness. For example, for children who work, the probability of experiencing tiredness/exhaustion is approximately 71 percent, while the probability of suffering from other health problems is approximately 26 percent. This relationship increases substantially in magnitude when moving to the bivariate probit model, with the exception of backache. This suggests a more robust effect of child labour on health (Table D3 in Appendix D for full estimates).[91] The Wald specification test of the correlation coefficient of errors suggests that child work is endogenous in all health conditions except for tiredness/exhaustion and backache (see Table 5.6, bottom). In addition, the coefficient of correlation between the residuals of the health outcomes and the child work equation is always significantly negative in three out of five health conditions, implying that considering child work as exogenous leads to biased estimates.[92]

[91] We further investigate our analysis by including regional dummies (for example, Chittagong, Rajshahi, Khulna, Barisal, Sylhet and Noakhali. The reference category is Dhaka) in our baseline model to capture the unobserved factors (for example, climate, hospital facilities and public hygiene) that may affect the causal relationship between health and labour supply. If regional dummies are good proxies for unobserved factors affecting work and child health, then we can be confident that we are indeed measuring a causal effect of child work. Of course, there are still other unobserved factors driving the correlation between child work and subjective child health. In general, we find (not shown) a strong positive effect of work on the probability of reporting injury or illness, which reiterates our findings from Table D3. These results suggest that the effect of work on health seems to be mediated through regional dummies and, hence, these factors are perhaps important determinants.

[92] O'Donnell et al. (2005, p.454) obtained a similar negative value of the correlation coefficient of errors in rural Vietnam and interpreted this result as 'selection into work on the basis of unobserved health determinants.'

Turning next to the effects of other covariates on the probability of reporting injury or illness provides some interesting results (see Table D3 in Appendix D). Consistent with our descriptive analysis, girls are less likely to report injury or illness, which suggests that the nature of work undertaken by girls may be less onerous.[93] Interestingly, protection (use of working dress) at the workplace does not reduce injury or illness except for tiredness/exhaustion and body injuries.[94,95] These findings are similar to those reported by Guarcello et al. (2004) for Cambodia. Along with Guarcello et al. (2004), our results indicate that the use of protective clothing is not sufficient to fully compensate for the additional risks related to work. The number of adults living in the household significantly reduces the probability of injury or illness (except for body injuries and other health problems). This finding is consistent with the findings of O'Donnell et al. (2005). One explanation could be that the presence of adult persons is likely to relax the liquidity constraints of the household and, therefore, potentially decreases child labour participation and the probability of poor health conditions. As expected, children are more likely to report backache if they work in agriculture, though the effect is not statistically different from zero at conventional levels of significance. Clearly, construction and manufacturing jobs appear to endanger child health as the coefficients for poor health conditions are greater in magnitude than they are in other sectors, although the estimated coefficients for tiredness/exhaustion, backache and other health problems in the construction sector and tiredness/exhaustion in manufacturing

[93] The findings may be under-reported because NCLS 2002 does not report injury or illness attributed to domestic work, and this is the type of work that female children most often do. Thus, some caution should be given to this result. An alternative explanation is that girls may be less forthcoming in reporting injury or illness in Bangladesh due to social and cultural factors.

[94] At this point it should be noted that these strange results do not disappear when controlling for the interaction between protection and sectors of employment and regressing health outcomes on protection, sectors of employment and an interaction between protection and sectors of employment at the same time. However, we do find the expected sign for the coefficient on the interaction between protection and sectors of employment. This indicates that safety levels reduce the risk of injury or illness across all sectors of employment.

[95] It is important to note that protection at the workplace may be a potentially endogenous variable due to the possibility of reverse causality. A greater protection can be adopted in more hazardous jobs. We test the exogeneity of protection at the workplace by a Smith-Blundell test in the univariate probit model. The instruments are as defined for the bivariate probit. Exogeneity of this variable is not rejected at any reasonable levels of significance in all health conditions with the exception of backache ($\chi^2(1)$= 6.36, p =0.0117). Furthermore, given it is the work effect that is of central interest, we simply verify whether the estimate of this parameter appears to be contaminated by any endogeneity of protection at the workplace variable. As we treated child labour as endogenous, we excluded the variable protection at the workplace and re-estimated the bivariate probit model for all health conditions. The estimates generated from these models are very similar to those presented in Table D3. In particular, the bivariate probit work coefficient is robust to dropping to protection at the workplace variable, varying between 0.6803 and 1.7201 and remaining significant at the 1 percent level. These sensitivity tests suggest that the estimated parameters including the child work variable are not contaminated by endogeneity bias, deriving from protection at the workplace.

sector are not statistically significant. This result supports the global consensus that construction jobs are more hazardous in nature and thus raise health risks for children. When turning to parental characteristics, we find that the father's age significantly reduces the reporting of any injury or illness, backache and other health problems, while the father's higher education (secondary education) has the reverse effect on health conditions, such as body injuries. One possible explanation could be that child labour does not necessarily substitute for adult labour income and, hence, yields negative effects on health. However, a mother's higher education (secondary education) relates negatively with all health outcomes. A similar result was found by O'Donnell et al. (2005) for rural Vietnam. These results most likely suggest that highly educated women may be more aware of the adverse impact of child labour through access to information (i.e. exposure to media) and, consequently, adopt necessary steps (for example, to use preventive and curative medicines to treat illness) to improve child health. Safe drinking water, satisfactory sanitation and the number of rooms in the household also significantly reduce the probability of injury or illness. These results suggest that the quality of living conditions within the household is important.

Next, we turn to the results of partially linear models (Table 5.7) when children's working hours are taken into account and when controlling for similar sets of covariates as in the bivariate probit model.[96] The estimate of residual is significant for all health conditions (except for other health problems), implying that exogeneity of hours worked is rejected in a partially linear regression model at conventional levels of significance. Regarding the effect of the (log) of the number of hours worked, the significance test of the hour variable indicates that the number of hours worked significantly influences the probability of injury or illness (in every case, the p-value is 0.000). To show how occurrence of injury or illness varies with working hours, we show the non-parametrically estimated relationship between the (log) of the number of hours worked and health conditions in Figure 5.2. Reporting any injury or illness clearly decreases with the number of working hours, as do other health problems, but it then increases with the number of working hours after a certain threshold. The nonlinearity that we find may be attributed to the fact that a certain number of working hours is associated with a

[96] The bottom panel of Table 5.7 presents a one-sided specification test result for the parametric against the partially linear model. For the different health outcomes, both the linear model (i.e. the health outcomes depend linearly on the log of the number of hours worked) and quadratic specifications are rejected.

particular age and gender composition or other characteristics (for example, the task performed), which increase the likelihood of injury or illness after a certain threshold. While body injury and backache are generally constant with the number of hours worked, tiredness/exhaustion steadily increases with the number of hours worked.

The results of the parametric aspect (see Table D4 in Appendix D for full estimates) suggest that partial linear estimates are qualitatively similar to the bivariate probit specifications, although the magnitude of the impact of covariates is considerably smaller than that of the bivariate probit estimates. As previously discussed, protection at the workplace will not necessarily reduce injury or illness. This conclusion, however, emerges from only three health conditions: reporting any injury or illness, tiredness/exhaustion (though insignificant) and backache. Interestingly, jobs in agriculture and in wholesale and retail are found to be detrimental to child health. For example, children are more likely to report any injury or illness or backache when they work in agriculture and wholesale and retail. This implies that work participation itself does not appear to endanger health for those children who work in these sectors (as we see in the bivariate probit specification in Table D3). However, the risk of poor health conditions increases the longer the children are exposed to health hazards in these sectors. As regards parental characteristics, the results are similar to those obtained with the bivariate probit specification but not for all health conditions. For example, while the father's age significantly reduces tiredness/exhaustion in a partially linear regression model, the reverse effect is observed in the bivariate probit specification.

5.6 Further Analyses

5.6.1 Controlling for Sample Selection Bias

It is possible that when the (log) of the number of hours worked is censored at zero for a significant fraction of the observations, the conventional regressions method fails to account for the qualitative difference between limit (zero) observations (i.e. persons who have never participated in the labour force) and non-limit observations (i.e. persons who have entered the labour force and for whom the number of hours worked is greater than zero). A method that uses only non-limit observations suffers from sample selection bias, since persons for whom the number of hours worked is positive may not randomly drawn from the population, but a self-selected group. With self-selected samples, the mean value of the error term in the outcome equation may not equal zero, violating the basic assumption of the classical OLS model. More particularly, the error term may be correlated with the included variables, leading to estimation bias. We adopt the Heckman (1979) two-step approach to take account of the selection issue into the sample of children with positive working hours and use a similar set of covariates (except the variable of the number of children for each child in the household) that are used in the main analysis.[97]

As is well known, the sample selection model requires an exclusion restriction, in the form of one or more variables that appear in the participation equation but not in the outcome equation (the log of the number of hours worked). Given the lack of a credible exclusion restriction, we followed two alternative approaches to achieve identification of the selectivity term (the inverse Mill's ratio) though neither may be ideal. The first approach is identification through functional form[98] and the second is using variables that are significant in the participation equation (selection equation) but insignificant in the outcome equation.[99]

The selectivity-corrected equations of the (log) of the number of hours worked, conditional on participation, are presented in Table 5.8, using both methods of identification of the inverse Mill's ratio. Both show that selectivity into participation is unimportant. The sign of the inverse Mill's ratio (though insignificant) is as expected, that is, those who are likely to participate in the labour force are those who work more hours of work than do children in general. One possible explanation is that children who participate must be those with higher ambition and/or motivation. Given the imperfect selectivity-correction strategy and, more importantly, given the inverse Mill's ratio is not statistically significant, we suggest that the censoring effect appears to be trivial in our analysis.

[97] The Tobit procedure has been used in the literature to model censored dependent variables but it is a restrictive solution, simultaneously modelling the decision to participate in the labour force and the decision of the number of hours worked. A more appropriate approach is to treat the decision to participate in the labour force as essentially separate from decisions of the number of hours worked.
[98] The sample selection model can be identified by the non-linearity of the inverse Mill's ratio.
[99] Using a similar procedure, Kingdon (2002) corrected sample selection bias due to selection of individuals with positive years of schooling.

5.6.2 Isolating the Rural Sample

In this sub-section, we examine the robustness of our results when we restrict ourselves to the sample of rural child workers aged between 5 and 17 years.[100] Child labour is much more prevalent in rural areas and is significantly more common in agriculture (75 percent) (see Table 5.9). However, urban children work more hours than do their rural counterparts. Consequently, the proportions of reporting injury or illness or reporting any symptoms of injury or illness are higher in urban areas than in rural areas. This findings support the hypothesis that working hours have a negative impact on health status.

Focusing on the impact of child work participation on child health outcomes, it is noted that bivariate probit estimates for rural areas (Table 5.10) are quite similar to those for the full sample (Table 5.6). The one notable change is that the work coefficient for backache becomes statistically significant; it rises in magnitude but remains negative. These results have obtained by using only the migration status of the household and the school quality variables as instruments.[101] We find that migration status has significantly reduced the probability of child labour at conventional levels of significance, while schooling in an informal school is associated with a significant increase in the probability of child labour at the 1 percent level (see Table 5.11). Additionally, the relevance of these instruments is checked by running the bivariate probit models with and without these instruments. The LR test results suggest that adding these instruments to the model significantly improves the fit of the model compared to a model without these instruments.[102]

Turning finally to the impact of child working hours, partial linear estimates show an effect very similar to that of the full sample (Table 5.12). See Column 2 of Table D2 for the first-stage estimates and Table D5 for the full estimates of the partial

[100] Unfortunately, the analysis is not conducted in urban areas as the available instruments used in the main analysis did not work for child labour in urban areas.

[101] In the rural sample, in the estimated bivariate model, we experimented with total household land holdings as a possible determinant of child labour. While the significance of this instrument is confirmed in the work equation, the exclusion condition appears to be rejected in all health conditions. It is not clear what is driving this result.

[102] In the first health indicator (any injury/illness), the $\chi^2(2)$= 4.65 with a p-value of 0.0977. In the case of different health conditions (symptoms of injury/illness), the corresponding values are $\chi^2(2)$ = 5.52, with a p-value of 0.0634 (tiredness/exhaustion); $\chi^2(2)$= 6.42 with a p-value of 0.0403 (body injuries); $\chi^2(2)$= 25.04 with a p-value of 0.000 (backache) and $\chi^2(2)$ = 5.89 with a p-value of 0.0526 (other health problems).

linear regression model. Again, most estimates regarding the residual are statistically significant, suggesting that working hours are endogenous. Analysing the child's working hours, we find that the hour effect is significantly different from zero. This is confirmed by a significance test on hours. The instruments are the same as those used for the bivariate probit model for the rural sample. These instruments perform better with respect to the over-identification test and are now even stronger. The Hansen test for over-identifying restrictions yields a $\chi^2(2)$ test statistic of 9.80 (p-value = 0.0017) for reporting any injury or illness; 0.9268 (p-value = 0.3357) for tiredness/exhaustion; 0.5788 (p-value = 0.4467) for body injuries; 0.1232 (p-value of 0.7256) for backache; 5.20 (p-value of 0.0226) for other health problems. As in the full sample, we find a non-linear relationship between the (log) of the number of working hours and health outcomes.

5.6.3 Age Groups

In this sub-section, we present evidence for the relationship between child labour and subjective child health according to age. In the NCLS data, Guarcello et al. (2004) find that work-related injury or illness increases with age, though they did not offer any consistent explanation for this. These findings could be interpreted as support for the notion that older children work more hours than younger children, and hence their health worsens. Therefore, the health outcomes for different age groups are not essentially parallel. There is, however, no evidence of different health outcomes related to work between younger and older children.

We consider three age groups (10-13, 14-17 and 10-17) and estimate bivariate probit models for each age group in order to gauge the association between subjective health and child labour using similar sets of covariates and instruments that were used in the main analysis. We find some evidence that the probability of reporting injury or illness is somewhat larger in the oldest age group (aged 14-17) (Table 5.13). This holds particularly true for the case of tiredness/exhaustion. One possible explanation could be that older children are most likely chosen for physically demanding activities that cause him/her to become tired/exhausted. The point estimates for tiredness/exhaustion are 1.0068 (z = 4.20) for age 10-13 and 2.0315 (z = 24.15) for age 14-17. For the other health outcomes, the results are mixed across age groups. For example, we find weak evidence for reporting any injury or illness (except for age group 10-17). Furthermore, we find evidence that work increases the likelihood of backache and other health problems but does so much more strongly for younger children than for older children.[103] These results may be associated with the view that some health conditions are age related. Another possible interpretation is that younger children are involved in activities that could be too physically demanding for their age, which may weaken their immune system and leave them vulnerable to illness.

5.6.4 Heterogeneity of Work Effect on Injury or Illness

In this sub-section, we analyse the heterogeneity of the work effect on subjective child health. Heterogeneity can take place among child workers who work in different sectors. We also need to know how the number of working hours affects the health of the child across different sectors. The effect of working hours on health by sector is important as it is likely to shed light on whether it is more appropriate to target activity by sector or by a combination of both sector and working hours to identify the overall risk of suffering from injury or illness due to work. To explore the association between the number of working hours and health conditions in different sectors, we re-estimated the partially linear model discussed in Section 5.4.2, taking into account the endogeneity of child labour hours in health status equations. This analysis relies on three instruments, that is, the migration status of the household, an interaction term between the migration status and the household location and the school quality.[104]

In Table 5.14, we present non-parametric estimates of the relationship between working hours and health conditions in selected sectors. These sectors are agriculture, manufacturing, wholesale and retail and construction. The estimates of the residuals in all health conditions across a sector of employment suggest that the exogeneity of hours worked is rejected, though not for all health conditions that we considered. As before, there is evidence of the effect of number of hours worked on the probability of injury or illness across a sector of employment (in every case, the p-value is 0.000). This result is confirmed by specification tests on the number of working hours for all health conditions.

[103]The points estimates for backache and other health problems are 1.7503 (z = 3.31) and 1.0855 (z = 6.22) for age 10-13; the corresponding values for age 14-17 are 0.7048 (z = 1.52) and 0.5250 (z = 1.71), respectively.

[104]All these instrumental variables have strong explanatory power in that they have a high F-statistic. Over-identification is not rejected at the 5 percent level.

In Figure 5.3, we show how the occurrence of any injury or illness varies with the (log) of the number of hours worked in selected sectors in Bangladesh. In agriculture, injury or illness increases steadily with the number of hours worked after a certain threshold. A more or less similar pattern is obtained for manufacturing with different thresholds. Furthermore, the semi-parametric estimates of reporting any injury or illness in wholesale and retail declines before it becomes almost constant with the number of hours worked. The construction sector seems to have a different pattern, showing a sharp increase in injury or illness with the number of hours worked. These results may be attributed to the characteristics of the different sectors.

5.6.5 Severity of Injury or Illness

Before we conclude this essay, an important issue to emphasise is the severity of injury or illness. To take this into account, we need reliable indications of the seriousness of the injury or illness. While the NCLS 2002 does not collect direct information on whether a child is seriously injured or ill, the survey collects information on whether children received any medical treatment or consulted with the doctor following an injury or illness. Though the type of treatment received is far from being a perfect measure for the severity of injury or illness, we use the information available on the treatment as a proxy for the intensity of the injury. We have determined that three possible events follow the occurrence of an injury or illness: (i) the injury or illness did not require medical treatment; (ii) the injury or illness did require medical treatment; (iii) the injury or illness did require other treatments, such as hospitalisation. 'The injury or illness did not require medical treatment' is the reference category. Given the nature of the dependent variable, we have estimated the model using an ordered probit model. The analysis was restricted to children between the ages of 5 and 17 and focused on the impact of the number of hours worked. We also use the quadratic term for working hours to capture the non-linear effects of the hours worked. The potential endogeneity of the hour variable is confirmed through a Durbin-Wu-Hausman test. The chi-square test rejects the joint exogeneity of hours worked and its square term ($\chi^2(2) = 6.13$, $p = 0.0467$). Failure to reject the endogeneity of the hour variable in the ordered probit model suggests that we need to instrument hours worked and its square term.[105] The instruments are the same as

[105] We follow the procedure proposed by Ravallion and Wodon (2000). That is, in the first stage we estimate child labour hours and its square term by a Tobit model and obtain the residuals. The second

that used in the main analysis. Their relevance to the determination of the number of hours worked is confirmed by significant rejection of the exclusion restrictions on the respective reduced-form regressions. The assumed exogeneity of instruments is tested and not rejected.[106]

Without instrumentation, the number of hours worked is positively and significantly associated with the seriousness of the health episode (i.e. $\pi_{hour} = 0.0638$; $z = 19.12$) (see Table 5.15). This finding is consistent with the finding of Guarcello et al. (2004) in the case of Cambodia. However, the impact of hours weakens as the labour hours increase (i.e. $\pi_{hoursq} = -0.0004$; z = -11.43). If child working hours are instrumented, the effect becomes negative but remains statistically significant (i.e. $\pi_{hour} = -0.2457$; z = -1.65). The negative magnitude of the estimated coefficients of the hour variable suggests that the number of working hours do not influence intensity of injury or illness from the very first hour of work. However, the severity of injury or illness increases as the labour hours increase but is no longer statistically significant (i.e. $\pi_{hoursq} = 0.0035$; z = 1.58). These results indicate that if children work more than the threshold level (i.e. 35 hours a week), the intensity of injury or illness will eventually increase (see also Figure 5.4).

With respect to the effect of other covariates, we find that female and young children are less likely to experience severe accidents. As expected, the mother's higher education levels significantly reduce the probability of getting a serious injury or illness. Manufacturing and construction are the two sectors where the intensity of injury or illness is considerably larger compared to other sectors. For example, the estimated coefficient for agriculture is 2.385 (z = 2.02), and for wholesale and retail it is 2.076 (z = 1.88); however the corresponding values for manufacture and construction are 2.863 (z = 2.44) and 2.99 (z = 2.36), respectively.[107] The evidence suggests that severity of injury or illness should also be considered when attempting to identify the work effect on health because the consideration of the severity of poor health changes the relative risk

---

stage is estimated by an ordered probit model wherein the predicted residuals from the first-stage regressions are included as additional regressors to obtain the consistent estimates of each parameter.

[106] Following Kana et al. (2010), we apply the Wald test for instrumental variables. The null hypothesis is that the coefficients for instruments are simultaneously equal to zero. We cannot reject this, and instruments are exogenous for the health outcome ($\chi^2(3) = 3.56$, p = 0.3125).

[107] However, conclusions from this analysis should be taken with care as reporting and treatment can be influenced by individual and household characteristics, as well as by employment sector.

levels of work in different sectors. Hence, this is an important issue, and one that should be addressed by policymakers.

5.7 Summary and Conclusion

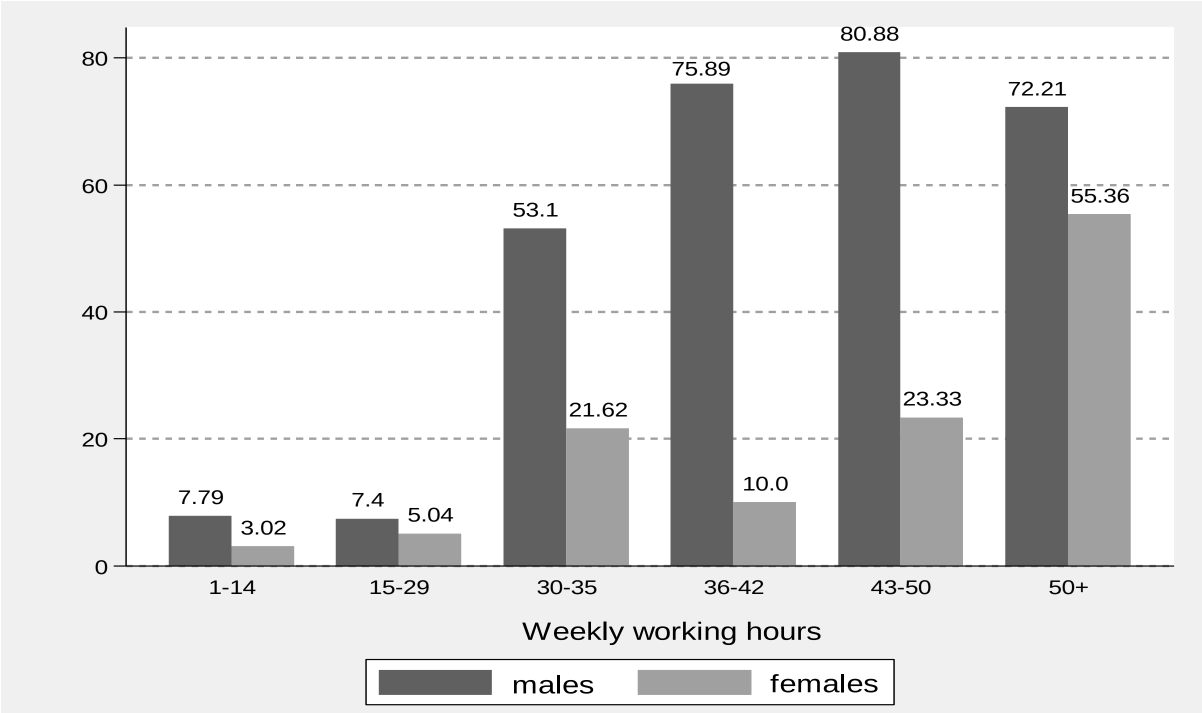
In this essay, we find that once we allow for potential endogeneity in the bivariate probit framework, there is a statistically significant positive association between child labour and the probability of reporting any injury or illness, tiredness/exhaustion, body injury and other health problems (infection, burns and lung diseases). The adverse impact of child labour on subjective child health in Bangladesh found in this study is consistent with the previous literature (Guarcello et al. 2004, Brazil and Cambodia study; Wolff and Maliki 2008, Indonesia study). This result appears to be reasonably robust when we restrict our analysis to rural children. We also find similar results when the analysis is extended to the relationship between the number of hours worked and the probability of reporting injury and illness, applying the semi-parametric approach. Our semi-parametric estimates suggest that the relationship between the number of hours worked and health status is non-linear, particularly in the case of reporting any injury or illness and other health problems.

Conducting further analyses, we studied the effect of child labour by age and found that younger children were more likely to suffer from backache and other health problems than older children, while the probability of reporting tiredness/exhaustion was greater in the oldest age group. We also investigated the effect of the number of working hours by sector on subjective child health. The semi-parametric estimates show that reporting any injury or illness increases with the number of hours worked but that they vary significantly across employment sectors. We have attributed this result to the characteristics of different sectors. Furthermore, we find evidence that the intensity of injury or illness increases with the number of hours worked after taking into account the endogeneity of child labour hours. This result holds true more in the construction and manufacturing sectors than for other sectors.

Given that we have shown that child labour leads to substantial increases in the probability of injury or illness, it is hoped that the results presented in this study will be useful for policymakers when implementing laws directed towards minimising or eliminating child labour. Because it may be extremely difficult to reduce or eliminate child labour in a developing country such as Bangladesh, policies are needed to improve the safety of child labour in those sectors that are most damaging to health, especially construction and manufacturing. Moreover, the sample statistics show that the ages of working children vary significantly in these two sectors. Overall, younger children are more likely to be employed in the manufacturing sector than in the construction sector. This strongly suggests that, while Bangladesh labour codes implement a minimum age of 18 years for hazardous work, there is a considerable lack of enforcement of this legislation. Thus, emphasis should be placed on a more effective implementation of legislation, including adequate monitoring.

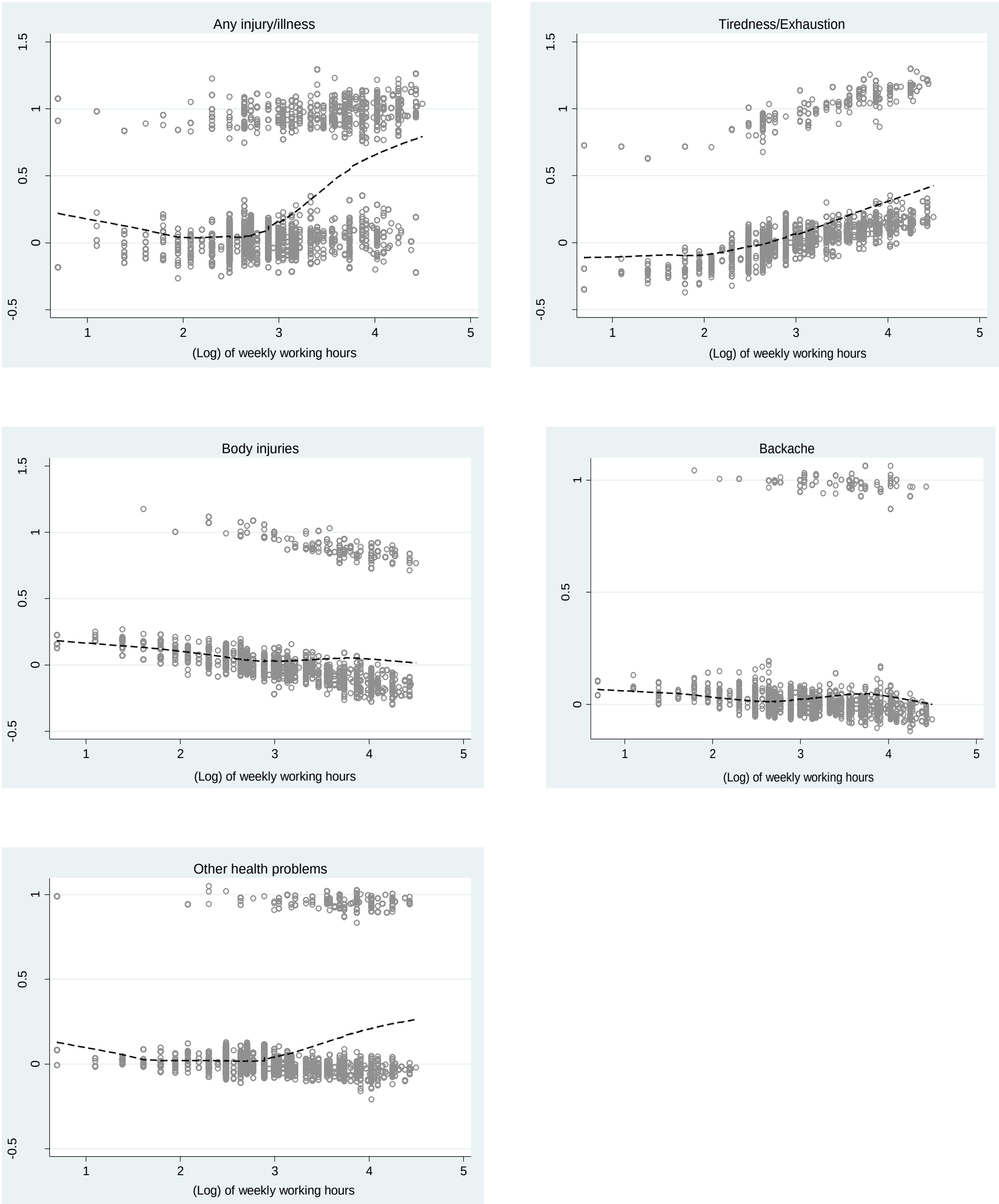
However, one clear limitation of this study is that the value of self-assessment alone is often not clear from a policy perspective. It would be difficult to evaluate the benefits of a public policy that may improve subjective health but leave more objective measures of health unchanged (for example, weight-for-age). Thus, more detailed data are required to analyse the issues of child labour and both the subjective and objective measures of child health. Panel data may also be useful for a further analysis of the long-term effects of child labour.

Figure 5.1: Working Hours and Health Injury/Illness of Children Aged 5-17 Years, by Gender



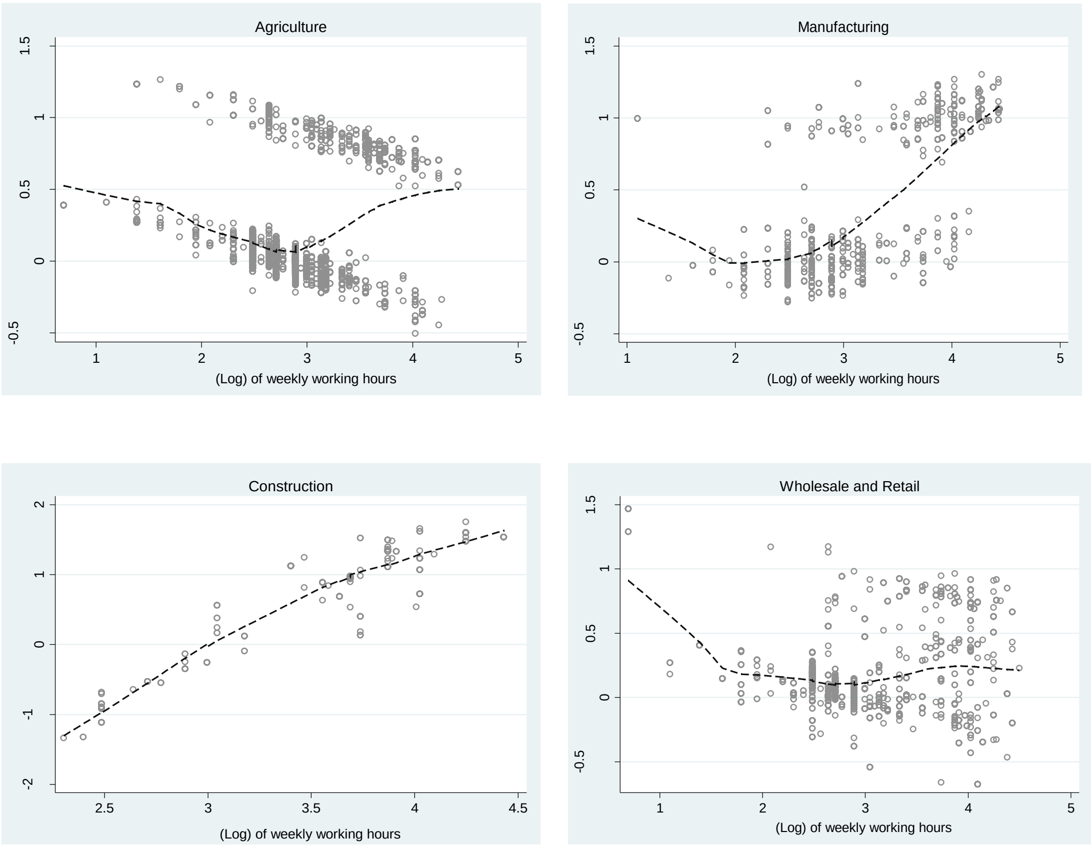
Source: Data are from NCLS 2002.

Figure 5.2: Non-linear Relationship between Working Hours (in Logs) and Health Outcomes



Source: Data are from NCLS 2002.

Figure 5.3: Non-linear Relationship between Working Hours (in Logs) and Reporting Any Injury/Illness, by Sector



Source: Data are from NCLS 2002.

Figure 5.4: Lowess Plot of Intensity of Injury/Illness, by Weekly Hours Worked

Source: Data are from NCLS 2002.

Table 5.1: Correlation between Different Forms of Injury/Illness

N = 16,010	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Injury/Illness	1				
Tiredness/Exhaustion	0.5289***	1			
Body injuries	0.4655***	-0.0509***	1		
Backache	0.3204***	-0.0351***	-0.0309***	1	
Other health problems	0.5202***	-0.0569***	-0.0501***	-0.0345***	1

Notes: Data are from NCLS 2002. 'Other health problems' include infection, burns and lung diseases. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.2: The Percentage of Health Conditions, by Gender and Work Status

	Workers			Non-workers			
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	*t* -test
By work status							
Injury/Illness	14,437	0.1814	0.3854	1,573	0.0801	0.2715	10.15 ***
Tiredness/Exhaustion	14,437	0.0580	0.2337	1,573	0.0248	0.1555	5.50 ***
Body injuries	14,437	0.0454	0.2083	1,573	0.0197	0.1390	4.78 ***
Backache	14,437	0.0221	0.1470	1,573	0.0089	0.0939	3.48 ***
Other health problems	14,437	0.0559	0.2297	1,573	0.0267	0.1613	4.91 ***
	Males			Females			
By gender	N	Mean	Std. Dev.	N	Mean	Std. Dev.	*t* -test
Injury/Illness	11,401	0.2143	0.4104	3036	0.0575	0.2331	20.19 ***
Tiredness/Exhaustion	11,401	0.0650	0.2465	3036	0.0316	0.1750	7.00 ***
Body injuries	11,401	0.0552	0.2285	3036	0.0086	0.0922	11.02 ***
Backache	11,401	0.0253	0.1572	3036	0.0099	0.0989	5.16 ***
Other health problems	11,401	0.0686	0.2531	3036	0.0076	0.0867	13.12 ***

Notes: Data are from NCLS 2002. Std. Dev. is standard deviation. *t*-test for difference (working-Non-working children) and (Males-Females). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.3: Age and Health Conditions of Working Children, by Sectors of Employment

	Mean	5-9 years	10-13 years	14-17 years	5-17 years
By age					
Agriculture	13.04	45.35	66.94	54.23	61.40
Manufacturing	12.98	22.25	12.66	10.96	12.23
Construction	14.02	1.13	1.29	4.60	2.59
Wholesale and Retail	13.42	26.76	17.87	24.53	20.72
Service	14.29	4.51	1.24	5.67	3.07
N		355	8,388	5,694	14,437
	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
By health conditions					
Agriculture	48.84	60.93	20.1	47.02	58.74
Manufacturing	22.87	18.04	29.73	30.09	19.45
Construction	8.21	5.5	18.75	3.45	4.34
Wholesale and Retail	17.07	12.43	26.07	19.44	13.63
Service	3.02	3.11	3.35	0.00	3.84
N	2,619	837	656	317	807

Notes: Data are from NCLS 2002. ‘Other health problems’ include infection, burns and lung diseases.

Table 5.4: List of Instruments and their Definitions

Instruments	Definitions
Migration status	A dummy variable which equals 1 if the household leaves the usual place of residence to find work.
Migration status x urban	An interaction term between the migration status and the location of the household. The reference category for the location of the household is rural area.
School quality	A dummy variable which equals 1 if a child reports that his/her source of education is informal schooling. The reference category is the formal public schools and/or the NGO schools.

Table 5.5: Robustness of Instruments

Dependent variable	Household average monthly income (Tk.)	Youth development training	Informal school	Injury/Illness
Migration status	0.0000	-3.5829	-3.3357	-0.0818
	(-0.0067)	(0.1650)	(0.1625)	(-0.1858)
N	16,010	16,010	16,010	16,010

Notes: Data are from NCLS 2002. Probit estimates. Variables included but not reported for different specifications are child’s age (in years) and its square term, sex of child, the interaction between child’s age and sex, dummies for sector of employment, urban areas, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household.

Table 5.6: Effect of Child Work on Injury/Illness, for Various Specifications

	Symptoms of Injury/Illness				
	Injury/Illness (child work)[a]	Tiredness/Exhaustion (child work)[a]	Body injuries (child work)[a]	Backache (child work)[a]	Other health problems (child work)[a]
Univariate Probit	0.7195***	0.7065***	0.6056***	0.2116**	0.2597*
	(0.0948)	(0.1202)	(0.1472)	(0.1043)	(0.1332)
Smith-Blundell Test of exogeneity: $\chi2(1)$	2.45	30.14	0.1341	12.69	0.3719
Prob.>C2 =	(p = 0.1172)	(p = 0.0000)	(p = 0.7143)	(p = 0.0000)	(p = 0.5420)
Log-pseudolikelihood	-5248.56	-2799.98	-2093.72	-1330.72	-2524.69
Pseudo-R^2	0.28	0.18	0.26	0.18	0.24
Bivariate Probit	1.3265***	1.9037***	0.7836***	-0.0697	0.6724***
	(0.1308)	(0.4078)	(0.1534)	(0.5943)	(0.1645)
Correlation of errors (ρ)	-0.3644***	-1.0899	-0.0947**	0.1478	-0.2358***
	(0.0551)	(0.8570)	(0.0415)	(0.2871)	(0.0664)
Wald test of ρ = 0	43.75	1.62	5.21	0.27	12.61
	(p = 0.0000)	(p = 0.2034)	(p = 0.0224)	(p = 0.6066)	(p = 0.0004)
N	16,010	16,010	16,010	16,010	16,010

Notes: Data are from NCLS 2002.[a]'Child work' is a binary variable. Standard errors in parentheses and are computed robustly to account for heteroskedasticity. 'Body injury' includes 'loss of limbs'. 'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, urban areas, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.7: Effect of Working Hours on Injury/Illness - Partial Linear Model Estimates

	Symptoms of Injury/Illness				
	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Semi-parametric model					
Residual	0.7940***	0.3491*	0.8677***	-0.4932***	0.0703
	(0.2256)	(0.1791)	(0.1586)	(0.1269)	(0.1627)
Significance test on hour	671.46	550.89	526.03	441.80	600.03
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Specific Tests:	Against semi-parametric models				
Linear model	670.38	550.80	525.52	441.09	598.84
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Quadratic model	611.63	537.4	521.86	441.09	574.51
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
N	14,436	14,436	14,436	14,436	14,436

Notes: Data are from NCLS 2002. 'Hour' is (log) of the number of hours worked by the child. Standard errors in parentheses. Body injury' includes 'loss of limbs'. 'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, urban areas, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.8: Heckman Sample Selection Model Estimates

Variables	Probit model of participation	Identification of inverse Mill's ratio by functional form: (Log) of the number of hours worked[a]	Identification of inverse Mill's ratio based on empirically justifiable exclusion restriction: (Log) of the number of hours worked[a]
Child's age	0.8897***	-0.1830***	-0.1818***
	(0.0601)	(0.0271)	(0.0271)
Child's age (squared)	-0.0336***	0.0112***	0.0111***
	(0.0024)	(0.0010)	(0.0010)
Female	0.4467*	0.3904***	0.3920***
	(0.2532)	(0.1044)	(0.1044)
Age*female	-0.0374*	-0.0432***	-0.0433***
	(0.0201)	(0.0081)	(0.0080)
Agriculture	2.9849***	-0.0818	-0.0693
	(0.0546)	(0.1135)	(0.1126)
Manufacturing	2.7580***	0.2333**	0.2450**
	(0.0773)	(0.1134)	(0.1126)
Construction	2.5595***	0.4232***	0.4345***
	(0.1245)	(0.1135)	(0.1127)
Wholesale and Retail	2.7907***	-0.0048	0.0067
	(0.0738)	(0.1120)	(0.1113)
Number of children 0-4	-0.2633***	0.0144**	0.0133**
	(0.0296)	(0.0062)	(0.0060)
Number of school children 5-17	-0.0020	0.0227***	0.0230***
	(0.0177)	(0.0036)	(0.0036)
Number of adults over 17 years	-0.0083	-0.0277***	-0.0261***
	(0.0209)	(0.0040)	(0.0034)
Father's age	-0.0146***	0.0010	
	(0.0034)	(0.0007)	
Father has primary education	-0.0348	-0.0387***	-0.0386***
	(0.0577)	(0.0112)	(0.0112)
Father has secondary education	0.4364***	0.0383***	0.0390***
	(0.0715)	(0.0106)	(0.0106)
Mother's age	0.0156***	-0.0006	
	(0.0046)	(0.0010)	
Mother has primary education	0.3964***	-0.0942***	-0.0957***
	(0.0773)	(0.0118)	(0.0119)
Mother has secondary education	-0.0319	-0.2743***	-0.2758***
	(0.0776)	(0.0107)	(0.0106)
Migration status	5.6314***	0.4446	
	(0.3527)	(0.6267)	
Migration status x urban	-2.9390***	-0.1585	
	(0.2448)	(0.3195)	
Electricity	0.2536***	-0.0817***	-0.0821***
	(0.0507)	(0.0095)	(0.0095)
Urban	-0.4246***	-0.0459***	-0.0468***
	(0.0545)	(0.0104)	(0.0104)
inverse Mill's ratio		0.2116	0.2457
		(0.3799)	(0.3774)
Constant	-5.6356***	3.4779***	3.4659***
	(0.3842)	(0.3448)	(0.3440)
N	16,010	14,437	14,437

Notes: Data are from NCLS 2002. Standard errors in parentheses. [a]OLS estimates. The exclusion restrictions are as follows: parental age, the migration status of the household, an interaction term between the migration status and the location of the household.*** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.9: Sample Means and Proportions of Key Variables, by Area and Child Work Status

	Urban					Rural				
	Workers		Non-workers			Workers		Non-workers		
Variables	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test	Mean	Std. Dev.	Mean	Std. Dev.	*t* -test
Age (in years)	13.2955	1.9081	12.7375	3.1668	5.23 ***	13.1297	1.7624	12.2600	3.2975	14.24 ***
Female (1= female)	0.1988	0.3991	0.5250	0.5000	-15.29 ***	0.2153	0.4111	0.3419	0.4745	-9.81 ***
Child labour hours (weekly hours of work)	24.4200	18.1113				20.4533	12.7119			
Injury/Illness (1 = yes)	0.2038	0.4029	0.0725	0.2596	6.40 ***	0.1717	0.3771	0.0827	0.2755	7.84 ***
Tiredness/Exhaustion (1 = yes)	0.0634	0.2436	0.0050	0.0706	4.77 ***	0.0556	0.2292	0.0315	0.1749	3.48 ***
Body injuries (1 = yes)	0.0631	0.2432	0.0600	0.2378	0.25	0.0378	0.1906	0.0060	0.0771	5.66 ***
Backache (1 = yes)	0.0366	0.1878	0.0025	0.0500	3.62 ***	0.0158	0.1247	0.0111	0.1047	1.24
Other health problems (1 = yes)	0.0407	0.1976	0.0050	0.0706	3.59 ***	0.0625	0.2421	0.0341	0.1816	3.89 ***
Agriculture (1= if child's sector of employment is agriculture)	0.3118	0.4633	0.0075	0.0864	13.11 ***	0.7453	0.4357	0.1449	0.3522	45.49 ***
Manufacturing (1= if child's sector of employment is manufacturing)	0.2349	0.4240	0.0550	0.2283	8.38 ***	0.0733	0.2607	0.0171	0.1295	7.29 ***
Construction (1 = if child's sector of employment is construction)	0.0535	0.2251	0.0000	0.0000	4.76 ***	0.0139	0.1171	0.0136	0.1160	0.07
Wholesale and Retail (1= if child's sector of employment is wholesale and retail)	0.3495	0.4769	0.0725	0.2596	11.46 ***	0.1454	0.3525	0.0205	0.1416	12.02 ***
N	4,372		400			10,065		1,173		

Notes: Data are from NCLS 2002. Std. Dev. is standard deviation. *t*-test for difference (Working-Non-working children). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.10: Effect of Child Work on Injury/Illness, for Various Specifications: Rural Sample

	Symptoms of Injury/Illness				
	Injury/Illness (child work)[a]	Tiredness/Exhaustion (child work)[a]	Body injuries (child work)[a]	Backache (child work)[a]	Other health problems (child work)[a]
Univariate Probit	0.9717***	0.8132***	1.2636***	0.1100	0.3754**
	(0.1333)	(0.1595)	(0.2681)	(0.1038)	(0.1630)
Smith-Blundell Test of exogeneity: $\chi2(1)$	0.0018	0.0062	3.51	7.21	0.1857
Prob.>$c2$ =	(p = 0.9661)	(p = 0.9368)	(p = 0.0610)	(p = 0.0073)	(p = 0.6666)
Log-pseudolikelihood	-3580.70	-1939.20	-1244.96	-730.90	-1932.38
Pseudo-R^2	0.28	0.17	0.26	0.18	0.24
Bivariate Probit	1.6165***	1.4123***	1.9305***	-2.4510***	0.7542***
	(0.1620)	(0.2409)	(0.3210)	(0.1380)	(0.1840)
Correlation of errors (ρ)	-0.4108***	-0.3874***	-0.4747***	14.0098***	-0.2156***
	(0.0731)	(0.1302)	(0.1382)	(1.7991)	(0.0598)
Wald test of ρ = 0	31.57	8.86	11.80	60.64	12.99
	(p = 0.0000)	(p = 0.0029)	(p = 0.0006)	(p = 0.0000)	(p = 0.0003)
N	11,238	11,238	11,238	11,238	11,238

Notes: Data are from NCLS 2002.[a]'Child work' is a binary variable. Standard errors in parentheses and are computed robustly to account for heteroskedasticity. 'Body injury' includes 'loss of limbs'. 'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.11: The Power of the Instrumental Variables in Determining Child Work: Estimates from Bivariate Probit models of Injury/Illness, Rural sample

	Symptoms of Injury/Illness				
Instruments	Injury/Illness (child work)[a]	Tiredness/Exhaustion (child work)[a]	Body injuries (child work)[a]	Backache (child work)[a]	Other health problems (child work)[a]
Migration status	-0.4810**	-0.5725**	-0.5741**	-1.0059***	-0.5256**
	(0.2101)	(0.2251)	(0.2242)	(0.1882)	(0.2376)
Informal school	3.9972***	4.1762***	4.2930***	6.3395***	4.2459***
	(0.1336)	(0.1371)	(0.1672)	(0.0471)	(0.0718)
N	11,238	11,238	11,238	11,238	11,238

Notes: Data are from NCLS 2002. [a]'Child work' is a binary variable. Standard errors in parentheses. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.12: Effect of Working Hours on Injury/Illness, Partial Linear Model Estimates: Rural Sample

	Symptoms of Injury/Illness				
	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Semi-parametric model					
Residual	0.5105*	0.2972	0.3968**	-0.7179***	0.5344**
	(0.2811)	(0.2239)	(0.1893)	(0.1536)	(0.2091)
Significance test on hour	547.54	434.75	392.84	258.57	503.99
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Specific Tests:	Against semi-parametric models				
Linear model	500.0	422.22	416.67	275.0	450.0
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Quadratic model	507.14	433.33	433.33	300.0	462.5
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
N	10,064	10,064	10,064	10,064	10,064

Notes: Data are from NCLS 2002. 'Hour' is (log) of the number of hours worked by the child. Standard errors in parentheses. 'Body injury' includes 'loss of limbs'. 'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.13: Effect of Child Work on Injury/Illness across Age Groups

	Symptoms of Injury/Illness				
	Injury/Illness (child work)[a]	Tiredness/Exhaustion (child work)[a]	Body injuries (child work)[a]	Backache (child work)[a]	Other health problems (child work)[a]
Bivariate Probit					
10-13 years	0.4258	1.0068***	0.0503	1.7503***	1.0855***
N = 8,952	(0.2830)	(0.2397)	(1.3356)	(0.5294)	(0.1745)
Correlation of errors (ρ)	-0.5131***	-0.2678***	-0.0289	-0.7125	-1.0602***
	(0.1015)	(0.0691)	(0.0431)	(0.5553)	(0.2602)
Wald test of ρ = 0	25.56	15.04	0.45	1.65	16.60
	(p = 0.0000)	(p = 0.0001)	(p = 0.5027)	(p = 0.1995)	(p = 0.0000)
14-17 years	0.0834	2.0315***	0.4874**	0.7048	0.5250*
N = 6,400	(0.3891)	(0.0841)	(0.2054)	(0.4648)	(0.3068)
Correlation of errors (ρ)	0.5181***	-4.2316**	-0.0008	-0.2749	0.4826***
	(0.1999)	(1.7401)	(0.0785)	(0.2787)	(0.1257)
Wald test of ρ = 0	6.72	5.91	0.0001	0.9724	14.74
	(p = 0.0096)	(p = 0.0150)	(p = 0.9915)	(p = 0.3241)	(p = 0.0001)
10-17 years	1.2354***	-0.2293	0.8595***	-0.2782	1.1121***
N = 15,352	(0.1487)	(0.4930)	(0.1667)	(0.8862)	(0.1605)
Correlation of errors (ρ)	-0.2751***	0.4236*	-0.1718***	0.2453	-0.3492***
	(0.0663)	(0.2390)	(0.0560)	(0.4422)	(0.0613)
Wald test of ρ = 0	17.24	3.14	9.42	0.3079	32.43
	(p = 0.0000)	(p = 0.0763)	(p = 0.0021)	(p = 0.5790)	(p = 0.0000)

Notes: Data are from NCLS 2002. [a]'Child work' is a binary variable. Standard errors in parentheses and are computed robustly to account for heteroskedasticity. 'Body injury' includes 'loss of limbs'.'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, dummies for sector of employment, urban areas, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.14: Effect of Working Hours on Injury/Illness across Sector - Partial Linear Model Estimates

	Symptoms of Injury/Illness				
	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Semi-parametric model					
	Sector of employment				
Agriculture					
Residual	0.2158	-0.5635**	0.5786***	-0.2507	0.4514**
	(0.2876)	(0.2343)	(0.1668)	(0.1572)	(0.2131)
Significance test on hour	625.30	554.95	328.82	380.62	551.14
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Manufacturing					
Residual	-0.1146	0.2398*	0.2987**	-0.5003***	-0.1527
	(0.1521)	(0.1305)	(0.1420)	(0.1015)	(0.1267)
Significance test on hour	193.25	154.86	134.12	199.34	135.31
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Construction					
Residual	-1.0391***	-0.1014	-0.7715***	-0.5823***	0.4162***
	(0.1356)	(0.1415)	(0.1585)	(0.1054)	(0.1560)
Significance test on hour	168.95	72.22	174.95	53.88	50.98
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Wholesale and Retail					
Residual	0.5846***	0.1769*	0.4205***	-0.0141	0.0014
	(0.1294)	(0.0950)	(0.0911)	(0.0709)	(0.0845)
Significance test on hour	209.34	119.59	237.73	144.27	181.19
	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)

Notes: Data are from NCLS 2002. 'Hour' is (log) of the number of hours worked by the child. Standard errors in parentheses. Body injury' includes 'loss of limbs'. 'Other health problems' include infection, burns and lung diseases. Variables included but not reported for different specifications are child's age (in years) and its square term, sex of child, the interaction between child's age and sex, child's vaccination, age of parents, the number of children for each child in the household, the number of adults over 17 years, dummies for urban areas, parental education, protection at the workplace, dummies for dwelling characteristics and facilities enjoyed by the household and the number of rooms in the household. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table 5.15: Ordered Probit and IV Ordered Probit Estimates of Seriousness of Injury/Illness

			Tobit (First stage)	
Variables	Ordered Probit	IV Ordered Probit	Child labour hours	Child labour hours (squared)
Child's age	-0.8163***	-1.0251***	-4.5690***	-304.8768***
	(0.0499)	(0.1295)	(0.4304)	(32.2427)
Child's age (squared)	0.0330***	0.0485***	0.2822***	18.1993***
	(0.0019)	(0.0077)	(0.0163)	(1.2207)
Female	-1.7906***	-1.7961***	5.9118***	461.5240***
	(0.2581)	(0.3451)	(1.7864)	(134.5742)
Age*female	0.0779***	0.0793**	-0.7616***	-59.8465***
	(0.0190)	(0.0347)	(0.1366)	(10.2923)
Child's vaccination	-0.2306***	-0.1625***	-0.2175	-34.2152*
	(0.0285)	(0.0520)	(0.2366)	(17.6740)
Child labour hours	0.0638***	-0.2457*		
	(0.0033)	(0.1510)		
Child labour hours (squared)	-0.0004***	0.0035		
	(0.0000)	(0.0022)		
Residual_child labour hours		0.3096**		
		(0.1511)		
Residual_child labour hours (squared)		-0.0040*		
		(0.0022)		
Agriculture	-0.4198***	2.3847**	21.3875***	966.4163***
	(0.0714)	(1.1809)	(0.4393)	(33.8733)
Manufacturing	-0.0307	2.8635**	30.1310***	1,628.5697***
	(0.0773)	(1.1725)	(0.5144)	(39.2490)
Construction	-0.1382	2.9979**	35.6895***	2,002.8300***
	(0.0900)	(1.2707)	(0.7972)	(59.9730)
Wholesale and Retail	-0.6049***	2.0765*	23.6285***	1,173.2471***
	(0.0772)	(1.1028)	(0.4770)	(36.5361)
Number of children for each child in the household	-0.0216**	-0.0233	0.7649***	60.3459***
	(0.0110)	(0.0308)	(0.0792)	(5.9114)
Number of adults over 17 years	-0.1085***	-0.1731***	-0.9743***	-60.0134***
	(0.0166)	(0.0329)	(0.1063)	(7.9340)
Father's age	-0.0101***	-0.0145***	0.0140	2.2046
	(0.0029)	(0.0041)	(0.0210)	(1.5714)
Father has primary education	-0.1506***	-0.2217***	-1.2253***	-77.7644***
	(0.0450)	(0.0539)	(0.3206)	(23.9238)
Father has secondary education	-0.1419***	-0.0699	1.7781***	121.1000***
	(0.0391)	(0.0608)	(0.3013)	(22.4959)
Mother's age	0.0171***	0.0236***	-0.0125	-2.6093
	(0.0033)	(0.0052)	(0.0266)	(1.9847)
Mother has primary education	-0.2475***	-0.2936***	-2.1228***	-154.6218***
	(0.0472)	(0.0789)	(0.3337)	(24.9130)
Mother has secondary education	-0.6499***	-0.9135***	-8.0248***	-561.9642***
	(0.0741)	(0.2297)	(0.3556)	(26.5474)
Protection	-0.3086***	-0.3095***		
	(0.0636)	(0.0636)		
Urban	-0.0491	-0.0024	-2.6094***	-216.7568***
	(0.0364)	(0.1208)	(0.2580)	(19.2728)
Safe drinking water	-0.6848***	-0.6860***		
	(0.0631)	(0.0631)		
Electricity	-0.2850***	-0.2857***		
	(0.0342)	(0.0342)		
Sanitation OK	-1.1152***	-1.1247***		
	(0.2102)	(0.2094)		
Number of rooms in the household	0.0049	0.0049		
	(0.0136)	(0.0136)		

Continued on next page

Table 5.15 (continued): Ordered Probit and IV Ordered Probit Estimates of Seriousness of Injury/Illness

			Tobit (First stage)	
Variables	Ordered Probit	IV Ordered Probit	Child labour hours	Child labour hours (squared)
Migration status			20.4802*	1,267.5353
			(10.5630)	(786.5964)
Migration status x urban			-11.5066*	-656.1988
			(5.9717)	(445.0991)
Informal school			2.712	203.4208
			(2.3200)	(172.6890)
Constant			16.0356***	924.2913***
			(2.8966)	(217.0938)
Cut 1	-4.4146***	-5.7244***		
	(0.3339)	(0.6295)		
Cut 2	-3.4330***	-4.7426***		
	(0.3346)	(0.6292)		
Sigma			13.2804***	988.5295***
			(0.0782)	(5.7975)
F -test			1.72	2.50
			(p = 0.0152)	(p = 0.0517)
N	16,010	16,010	16,010	16,010

Notes: Data are from NCLS 2002. Robust standard errors are in parentheses. The omitted categories are male child, no vaccination, service sector, no schooling, no working dress, rural, source of drinking water is ponds/rivers, no electricity, no sanitary latrine, if the household does not leave their place of residence during the last 12 months and the formal public schools and/or the NGO schools. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Chapter 6

Conclusion

This thesis has investigated the extent of gender inequality in the demand for tertiary education and wages as well as the effect of child labour on education and health in Bangladesh. Although this thesis investigated diverse topics using different tools of analysis, at the heart of each analysis is an attempt to address issues essential to improving the well-being of individuals in developing countries.

Chapter 2 examined whether the gender wage gap varies along the wage distribution among full-time employees in Bangladesh. We examined the gender wage gap along the entire distribution of wages by using the Oaxaca-Blinder decomposition based on the unconditional quantile regression estimates (Firpo et al. 2009). Unlike the standard Oaxaca-Blinder approach, this approach enabled us to decompose the gender wage gap into its various components (explained and unexplained) at different points of the wage distribution.

Since full-time workers might not be a random subset of all workers, we also take into account possible sample selection into full-time employment. We use the Heckman (1979) two-step approach to take into account the selection issue in full-time employment and extend the Oaxaca-Blinder decomposition to the unconditional quantile regression framework. The use of the unconditional quantile regression approach to investigate the effect of sample selectivity bias on the gender wage gap at different points of the distribution is new.

On the whole, we find that women are paid less than men everywhere on the wage distribution in Bangladesh, and the gap is higher at the lower end of the distribution. Discrimination explains the major proportion of the wage gap at all quantiles. The gender wage gap, however, widened much more at the lower end of the distribution than at the upper end between 1999 and 2005. Our results also show that not controlling for sample selection is likely to over-estimate the observed wage gap across the wage distribution. The selection-corrected wage gap is found to be predominantly due to discrimination against women in the labour market.

Chapter 3 investigated gender inequality in tertiary enrolment rates in Bangladesh. There are many potential explanations for the gender gap in educational outcomes. This chapter aimed to investigate one explanation – labour market returns – for the gender gap in tertiary enrolment rates. In particular, we test whether wage premiums in returns from secondary to tertiary education explain low female enrolment in tertiary education.

After controlling for household- and individual-level characteristics and using the FE estimator at the district-area level, we find that wage premiums do not have any significant effect on the gender gap in tertiary enrolment rates. However, we note that wage premiums are a significant determinant for male enrolment. We find that labour market returns for females are generally higher, but their tertiary enrolment rates are lower and generally less responsive to labour market returns. The evidence presented in Chapter 3 also suggests that parental characteristics, most notably parents' education, may play a large role in attaining gender parity in tertiary enrolment rates in Bangladesh.

Chapter 4 examined the effect of child labour hours on child education in Bangladesh. The empirical estimates obtained in Chapter 4 suggest that child labour hours are endogenous, and that this endogeneity needs to be taken into account. We use an IV estimation strategy by exploiting the child's sector of employment (for example, agriculture, manufacturing, wholesale and retail and service) as potential instruments for child labour hours. Given the lack of theoretical justification of our instruments, we have verified their validity through empirical tests. These tests support the suitability of the instruments and, hence, the analytical soundness of our results.

Our overall results suggest that child labour hours adversely affect child schooling in Bangladesh, but the marginal impact of child labour hours weakens when working hours increase. This is reflected in reduced school attendance and age-adjusted school attendance rates. We do not have sufficient evidence to explain the reasons for this pattern of relationship between child labour hours and child schooling. We therefore use different degrees of polynomials in working hours and find that child attainment of schooling declines when working hours increase. This technique has largely been able to mitigate the controversy regarding the relationship between the number of working hours and child education. We also investigate the impact of child labour hours on child schooling using a non-parametric approach and find that the results are consistent with the parametric approach when we control for different degrees of polynomials in working hours.

In Chapter 4, we also sought to identify whether parental education affects the child's work-schooling trade-off. Since the mother's and father's education may affect investment in male and female children differently, we use both the mother's and father's level of education. The core finding is that both the mother's and father's education levels shifts the work-schooling trade-off in favour of education; however, we find that the mother's educational attainment has a stronger marginal impact on the work-schooling trade-off than the father's education. We also note that both the mother's and father's education often plays a significant role in the schooling decision for a female child. The same incentive effect is not found for a male child.

Furthermore, we addressed the selectivity bias in Chapter 4 because we restricted our estimating sample to working children. Overall, our analysis shows that a sample of working children may introduce significant bias of IV estimates. We also compared the results of the sample selection model with double-hurdle model estimates in Chapter 4 and found that there is a significant difference between these two estimates. This was largely due to the use of a different methodology.

In Chapter 5, we explored the contemporaneous effect of child labour on child health outcomes, in particular on self-reported injury or illness. In doing so, we considered two different measures of child labour: child labour force participation and the number of hours worked by the child. However, since both measures are endogenous in the child health outcome equation due to simultaneity bias we use an IV approach by exploiting the migration status of the household, as well as a proxy variable for school quality. In addition, we use Robinson's (1988) semi-parametric regression estimator to estimate the impact of child labour hours on child health outcomes, treating child working hours as endogenous. We also attempt to identify the relationship between work heterogeneity of child labour and health, concerning the lack of homogeneity of child labour across a sector of employment.

On the whole, we find that child labour increases the probability of injury or illness, in particular tiredness/exhaustion, body injuries and other health problems (for example, infection, burns and lung diseases) when we correct for potential sources of endogeneity bias in the bivariate probit framework. This result is robust when we restricted our analysis only to rural children and split the sample by sectors of employment (for example, agriculture, manufacturing, wholesale and retail and construction). A similar finding was obtained with the semi-parametric approach. However, as demonstrated in other studies, the semi-parametric estimates suggest that the effect of the number of working hours is not linear for child health outcomes. It is also worth noting that younger children are more likely to suffer from backache and other health problems while older children are more likely to suffer from tiredness/exhaustion. We provide explanations for different health outcomes between younger and older children in Chapter 5.

The results obtained in this study are relevant for policymakers in matters pertaining to reducing gender wage discrimination in the labour market, as well as improving women's incentives to acquire a tertiary education. Furthermore, the findings of this thesis strengthen the need for stronger enforcement of laws that regulate child labour, especially given its adverse consequences on both human capital accumulation and child health.

These results are pertinent for future research in the area of economic well-being and household and/or family economics. Further research can also be accomplished in the area of income inequality, employment, and women's development.

References

Ahmed, S. (2007), "Gender Wage Discrimination in Rural and Urban Labour Markets of Bangladesh". Unpublished Mphil Thesis.

Ahmed, S., and Maitra, P. (2010), "Gender Wage Discrimination in Rural and Urban Labour Markets of Bangladesh", *Oxford Development Studies,* 38(1), 83-112.

Akabayashi, H., and Psacharopoulos, G. (1999), "The Trade-off between Child Labour and Human Capital Formation: A Tanzanian Case Study", *Journal of Development Studies*, 35(5), 120-140.

Al-Samarrai, S. (2007), "Changes in Employment in Bangladesh, 2000-2005: The Impacts on Poverty and Gender Equity", Working paper prepared for the 2007 Bangladesh Poverty Assessment, The World Bank.

Albrecht, J., Bjorkluud, A., and Vroman, S. (2003), "Is there a Glass Ceiling in Sweden?", *Journal of Labor Economics,* 21(1), 145-177.

Albrecht, J., Vuuren, A., and Vroman, S. (2009), "Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands", *Labour Economics,* 16(4), 383-396.

Altonji, J. G. (1991), "The Demand for and Return to Education When Education Outcomes are Uncertain", NBER Working Paper Series, No. 714, National Bureau of Economic Research, Cambridge, MA.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005), "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling", *Journal of Human Resources,* 40(4), 791-821.

Amin, S., Quayes, M. S., and Rives, J. M. (2004), "Poverty and other Determinants of Child Labor in Bangladesh", *Southern Economic Journal*, 70(4), 876-892.

Amin, S., Quayes, M. S., and Rives, J. M. (2006), "Market Work and Household Work as Deterrents to Schooling in Bangladesh", *World Development,* 34(7), 1271-1286.

Anderson, K. H., King, E. M., and Wang, Y. (2003), "Market Returns, Transfers and Demand for Schooling in Malaysia, 1976-89", *Journal of Development Studies,* 39(3), 1-28.

Arends-Kuenning, M., and Amin, S. (2004), "School Incentive Programs and Children's Activities: The Case of Bangladesh", *Comparative Education Review,* 48(3), 295-317.

Aristei, D., Perali, F., and Pieroni, L. (2008), "Cohort, Age and Time Effects in Alcohol Consumption by Italian Households: A Double-Hurdle Approach", *Empirical Economics,* 35(1), 29-61.

Arulampalam, W., Booth, A. L., and Bryan, M. L. (2007), "Is there a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution", *Industrial and Labor Relations Review,* 60(2), 163-186.

Asadullah, M. N. (2006), "Returns to Education in Bangladesh", *Education Economics,* 14(4), 453-468.

Asadullah, M. N., and Chaudhury, N. (2009), "Reverse Gender Gap in Schooling in Bangladesh: Insights from Urban and Rural Households", *Journal of Development Studies,* 45(8), 1360-1380.

Ashagrie, K. (1997),"Statistics on Working Children and Hazardous Child Labour in Brief. Geneva: International Labor Office.". Retrieved from http://www.ilo.org/public/english/child/documentation/trends/stats.htm.

Ashraf, B., and Ashraf, J. (1993), "Estimating the Gender Wage Gap in Rawalpindi City", *Journal of Development Studies,* 24(2), 365 - 376.
Aslam, M. (2009), "Education Gender Gaps in Pakistan: Is the Labor Market to Blame?", *Economic Development and Cultural Change,* 57(4), 747-784.
Baland, J. M., and Robinson, J. A. (1998), "A Model of Child Labor", CRED Serie Recherche 206, University of Namer, Belgium.
Barkat, A., and Majid, M. (2003), "Adolescent Reproductive Health in Bangladesh: Status, Policies, Programs and Issues", Policy Project Report, U.S. Agency for International Development Asia/Near East Bureau, Washington, DC.
Basu, K. (1999), "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards", *Journal of Economic Literature,* 37(3), 1083-1119.
Basu, K. (2006), "Gender and Say: A Model of Household Behaviour with Endogenously Determined Balance of Power", *The Economic Journal,* 116(511), 558-580.
Basu, K., Das, S., and Dutta, B. (2010), "Child Labor and Household Wealth: Theory and Empirical Evidence of an Inverted-U", *Journal of Development Economics,* 91(1), 8-14.
Basu, K., and Ray, R. (2001), "The Collective Model of the Household and An Unexpected Implication for Child Labor: Hypothesis and an Empirical Test", Working Paper 2813, Washing D. C: The World Bank.
Basu, K., and Tzannatos, Z. (2003), "The Global Child Labor Problem: What Do We Know and What Can We Do?", *The World Bank Economic Review,* 17(2), 147-173.
Basu, K., and Van, P. H. (1998), "The Economics of Child Labor", *American Economic Review,* 88(3), 412-427.
Becker, G. (1971), "The Economics of Discrimination", Chicago: University of Chicago Press, Second Edition.
Becker, G. (1981), "A Treatise on the Family", Cambridge, Harvard University Press.
Beegle, K., Dehejia, R., and Gatti, R. (2009), "Why Should We Care about Child Labor?", *Journal of Human Resources,* 44(4), 871-889.
Behrman, J. R., and Deolalikar, A. B. (1995), "Are there Differential Returns to Schooling by Gender? The Case of Indonesian Labour Markets", *Oxford Bulletin of Economics and Statistics,* 57(1), 97-117.
Bhalotra, S. (2007), "Is Child Work Necessary?", *Oxford Bulletin of Economics and Statistics,* 69(1), 29-55.
Bhalotra, S., and Heady, C. (2003), "Child Farm Labor: The Wealth Paradox", *The World Bank Economic Review,* 17(2), 197-227.
Blinder, A. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", *The Journal of Human Resources,* 8(4), 436 - 455.
Boozer, M., and Suri, T. (2001), "Child Labor and Schooling Decisions in Ghana". CT: Yale University. Unpublished.
Burgess, S., Propper, C., and Rigg, J. (2004), "The Impact of Low Income on Child Health: Evidence from a Birth Cohort Study", CASE Paper 85, Centre for Analysis for Social Exclusion, London School of Economics.
Canals-Cerdá, J., and Ridao-Cano, C. (2004), "The Dynamics of School and Work in Rural Bangladesh", World Bank Policy Research Working Paper 333,0.
Chamarbagwala, R. (2008), "Regional Returns to Education, Child Labour and Schooling in India", *Journal of Development Studies,* 44(2), 233-257.

Chzhen, Y., and Mumford, K. (2011), "Gender Gaps across the Earnings Distribution for Full-time Employees in Britain: Allowing for Sample Selection", *Labour Economics*, 18(6), 837-844.

Cockburn, J., and Dostie, B. (2007), "Child Work and Schooling: The Role of Household Asset Profiles and Poverty in Rural Ethiopia", *Journal of African Economies*, 16(4), 519-563.

Cotton, J. (1988), "On the Decomposition of Wage Differentials", *The Review of Economics and Statistics,* 70(2), 236 - 243.

Coulombe, H. (1998), "Child Labour and Education in Cote d' Ivore", World Bank Working Paper Mimeo.

Cragg, J. G. (1971), "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica,* 39(5), 829-844.

D'Souza, S., and Chen, L. C. (1980), "Sex Differentials in Mortality in Rural Bangladesh", *Population and Development Review,* 6(2), 257-270.

Dar, A., Blunch, N.-H., Kim, B., and Sasaki, M. (2002), "Participation of Children in Schooling and Labor Activities: A Review of Empirical Studies", Social Protection Working Paper Series No. 0221,, The World Bank.

Deb, P., and Rosati, F. (2002), "Determinants of Child Labor and School Attendance: The Role of Household Unobservables", Understanding Children's Work Research Paper.

Dougherty, C. (2005), "Why Are the Returns to Schooling Higher for Women than for Men?", *Journal of Human Resources*, 40(4), 969-988.

Duncan, G., and Leigh, D. (1980), "Wage Determination in the Union and Non-Union Sectors: A Sample Selectivity Approach", *Industrial and Labour Relations Review,* 34(1), 24 - 34.

Duraisamy, P. (2002), "Changes in Returns to Education in India, 1983–94: by Gender, Age-Cohort and Location", *Economics of Education Review,* 21(6), 609-622.

Dyson, T., and Moore, M. (1983), "On Kinship Structure, Female Autonomy, and Demographic Behavior in India", *Population and Development Review,* 9(1), 35-60.

Edmonds, E. V. (2007). Child Labour. In Scultz, T. P. and Strauss, J. (Ed.), *The Handbook of Development Economics* (Vol. 4). Amsterdam: North Holland.

Emerson, P. M., and Souza, A. P. (2003), "Is there a Child Labor Trap? Intergenerational Persistence of Child Labor in Brazil", *Economic Development and Cultural Change,* 51(2), 375-398.

Emerson, P. M., and Souza, A. P. (2007), "Child Labor, School Attendance, and Intrahousehold Gender Bias in Brazil", *The World Bank Economic Review,* 21(2), 301-316.

Ersado, L. (2005), "Child Labor and Schooling Decisions in Urban and Rural Areas: Comparative Evidence from Nepal, Peru, and Zimbabwe", *World Development,* 33(3), 455-480.

Fassa, A. G. (2003), "Health Benefits of Eliminating Child Labour", IPEC Working Paper. Geneva:ILO.

Field, E., and Ambrus, A. (2008), "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh", *Journal of Political Economy,* 116(5), 881-930.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009), "Unconditional Quantile Regressions", *Econometrica*, 77(3), 953-973.

Freeman, R. B. (1975), "Legal Cobwebs": A Recursive Model of the Market for New Lawyers", *Review of Economics and Statistics,* 57(2), 171.

Ganglmair, B. (2006), "Intrinsic Competition and the Labor-Schooling Trade-off in Uganda", Working Paper, Econ WPA.

Ganguli, I., and Terrell, K. (2005), "The Gender Gap in Ukraine's Transition", IZA Discussion Paper No.1776.

García, J., Hernández, P. J., and López-Nicolás, A. (2001), "How Wide is the Gap? An Investigation of Gender Wage Differences using Quantile Regression", *Empirical Economics,* 26(1), 149-167.

Gibson, J., and Fatai, O. K. (2006), "Subsidies, Selectivity and the Returns to Education in Urban Papua New Guinea", *Economics of Education Review,* 25(2), 133-146.

Ginther, D. K., and Hayes, K. J. (2003), "Gender Differences in Salary and Promotion for Faculty in the Humanities 1977-95", *Journal of Human Resources,* 38(1), 34-73.

Glick, P., and Sahn, D. E. (2000), "Schooling of Girls and Boys in a West African Country: the Effects of Parental Education, Income, and Household Structure", *Economics of Education Review,* 19(1), 63-87.

Gormly, S., and Swinnerton, K. A. (2004), "The Effect of Adult Returns to Schooling on Children's School Enrolment, Theory and Evidence from South Africa". Unpublished Paper.

Goulart, P., and Bedi, A. S. (2008), "Child Labour and Educational Success in Portugal", *Economics of Education Review,* 27(5), 575-587.

Green, W. H. (1997), "Econometric Analysis", 3rd ed. Prentice-Hall, Inc.

Grootaert, C. (1999). Child Labor in Cote d' Ivore: Incidence and Determinants. In Grootaert, C. and Patrinos, H. (Ed.), *The Policy Analysis of Child Labor: A Comparative Study*: New York: St. Martin's Press.

Grossman, M. (1972), "The Demand for Health: A Theoretical and Empirical Investigation", Columbia University Press, for the National Bureau of Economic Research, New York.

Guarcello, L., Lyon, S., and Rosati, F. (2004), "Impact of Working Time on Children's Health", Research report, Understanding Children's Work.

Guarcello, L., Lyon, S., and Rosati, F. (2006), "Child Labour and Education for All: An Issue Paper", Understanding Children's Work Project Working Paper Series, November 2006.

Gunnarsson, V., Orazem, P. F., and Sánchez, M. A. (2006), "Child Labor and School Achievement in Latin America", *The World Bank Economic Review,* 20(1), 31-54.

Gupta, N. D., Oaxaca, R. L., and Smith, N. (2006), "Swimming Upstream, Floating Downstream: Comparing Women's Relative Wage Progress in the United States and Denmark", *Industrial and Labor Relations Review,* 59(2), 243-266.

Han, P. (2008), "Human Capital and Hours Worked of Children in Cambodia: Empirical Evidence for Policy Implications", *Asian Economic Journal,* 22(1), 25-46.

Hausmann, R., Tyson, L., and Zaahidi, S. (2008), "The Global Gender Gap Report".

Heady, C. (2003), "The Effect of Child Labor on Learning Achievement", *World Development,* 31(2), 385-398.

Heckman, J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica,* 47(1), 153 - 161.

Holly, A., and Sargan, J. (1982), "Testing for Exogeneity in a Limited Information Framework", *American Economic Review,* 70(2), 268-272.

Huebler, F. (2006), "Child Labour and School Attendance in Developing Countries: Empirical Evidence from National Household Surveys", Unpublished Doctoral Thesis, New School University.

Idler, E. L., and Benyamini, Y. (1997), "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies", *Journal of Health and Social Behavior,* 38(1), 21-37.

Islam, A., and Choe, C. (2011), "Child Labour and Schooling Responses to Access to Microcredit in Rural Bangladesh", *Economic Inquiry, forthcoming.*

Jacob, B. A. (2002), "Where the Boys aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education", *Economics of Education Review,* 21(6), 589-598.

Jensen, P., and Nielsen, H. S. (1997), "Child Labour or School Attendance? Evidence from Zambia", *Journal of Population Economics*, 10(4), 407-424.

Jolliffe, D., and Campos, N. F. (2005), "Does Market Liberalisation Reduce Gender Discrimination? Econometric Evidence from Hungary, 1986–1998", *Labour Economics,* 12(1), 1-22.

Jones, A. M. (1989), "A Double-Hurdle Model of Cigarette Consumption", *Journal of Applied Econometrics,* 4(1), 23-39.

Kambhampati, U. S., and Rajan, R. (2008), "The 'Nowhere' Children: Patriarchy and the Role of Girls in India's Rural Economy", *Journal of Development Studies,* 44(9), 1309-1341.

Kana, M., Phoumin, H., and Seiichi, F. (2010), "Does Child Labour Have a Negative Impact on Child Education and Health? A Case Study in Rural Cambodia", *Oxford Development Studies,* 38(3), 357-382.

Kaplan, G. A., and Camacho, T. (1983), "Perceived Health and Mortality: A Nine-Year Follow-up of the Human Population Laboratory Cohort", *American Journal of Epidemiology,* 117(3), 292-304.

Kapsos, S. (2008), "The Gender Wage Gap in Bangladesh", ILO Asia-Pacific Working Paper Series.

Kassenbohmer, S., and Sinning, M. (2010), "Distributional Changes in the Gender Wage Gap: 1994-2007", IZA Discussion Paper No. 5303.

Khanam, R. (2006), "Child Labour in Bangladesh: Trends, Patterns and Policy Options", *Asian Profile*, 34(6), 1-33.

Khanam, R. (2008), "Child Labour and School Attendance: Evidence from Bangladesh", *International Journal of Social Economics*, 35(1/2), 77-98.

Kidd, M., and Viney, R. (1991), "Sex Discrimination and Non-Random Sampling in the Australian Labour Market", *Australian Economic Papers,* 30(56), 28 - 49.

Kingdon, G. G. (1998), "Does the Labour Market Explain Lower Female Schooling in India? ", *Journal of Development Studies,* 35(1), 39-65.

Kingdon, G. G. (2002), "The Gender Gap in Educational Attainment in India: How Much can be Explained?", *Journal of Development Studies,* 39(2), 25-53.

Kingdon, G. G., and Theopold, N. (2006), "Do Returns to Education Matter to Schooling Participation? Evidence from India", Global Poverty Research Group Working Paper No. 52, Global Poverty Research Group.

Koenker, R., and Bassett, G., Jr. (1978), "Regression Quantiles", *Econometrica,* 46(1), 33-50.

Lokshin, M. (2006), "Difference-based Semiparametric Estimation of Partial Linear Regression Models", *Stata Journal*, 6(3), 377-383.

Machado, J. A., and Mata, J. (2005), "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression", *Journal of Applied Econometrics*, 20(4), 445-465.

Maitra, P. (2003), "Schooling and Educational Attainment: Evidence from Bangladesh", *Education Economics*, 11(2), 129-153.

Maitra, P., and Ray, R. (2002), "The Joint Estimation of Child Participation in Schooling and Employment: Comparative Evidence from Three Continents", *Oxford Development Studies*, 30(1), 41-62.

McElroy, M. B., and Horney, M. J. (1981), "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand", *International Economic Review*, 22(2), 333-349.

Micklewright, J., Pearson, M., and Smith, S. (1990), "Unemployment and Early School Leaving", *The Economic Journal*, 100(400), 163-169.

Miller, P. W. (2005), "The Role of Gender among Low-Paid and High-Paid Workers", *Australian Economic Review*, 38(4), 405-417.

Millimet, D. L., and Wang, L. (2006), "A Distributional Analysis of the Gender Earnings Gap in Urban China", *B.E. Journal of Economic Analysis & Policy: Contributions to Economic Analysis & Policy*, 5(1), 1-50.

Mueller, R. E. (1998), "Public-Private Sector Wage Differentials in Canada: Evidence from Quantile Regressions", *Economics Letters*, 60(2), 229-235.

NCLS. (2002), "Report on the National Child Labour Survey", Bangladesh Bureau of Statistics, Dhaka, Bangladesh.

Neumark, D. (1988), "Employers' Discriminatory Behaviour and the Estimation of Wage Discrimination", *Journal of Human Resources*, 23(3), 279 - 295.

Nopo, H. (2006), "The Gender Wage Gap in Chile 1992-2003 from a Matching Comparisons Perspective", IZA Discussion Paper No. 2698.

O'Donnell, O., Rosati, F., and Doorslaer, E. V. (2005), "Health Effects of Child Work: Evidence from Rural Vietnam", *Journal of Population Economics*, 18(3), 437-467.

Oaxaca, R. L. (1973), "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14(3), 693-709.

Pal, S. (2004), "Child Schooling in Peru: Evidence from a Sequential Analysis of School Progression", *Journal of Population Economics*, 17(4), 657-680.

Parikh, A., and Sadoulet, E. (2005), "The Effects of Parents' Occupation on Child Labour and School Attendance in Brazil". Unpublished Manuscipt. Univesity of California at Berkeley.

Patrinos, H. A., and Psacharopoulos, G. (1995), "Educational Performance and Child Labor in Paraguay", *International Journal of Educational Development*, 15(1), 47-60.

Patrinos, H. A., and Psacharopoulos, G. (1997), "Family Size, Schooling and Child Labor in Peru: An Empirical Analysis", *Journal of Population Economics*, 10(4), 387-405.

Pham, T., and Reilly, B. (2007), "The Gender Pay Gap in Vietnam, 1993-2002: A Quantile Regression Approach", *Journal of Asian Economics*, 18(5), 775-808.

Picchio, M., and Mussida, C. (2010), "Gender Wage Gap: A Semi-Parametric Approach with Sample Selection Correction", IZA Discussion Paper No. 4783.

Pitt, M. M., Rosenzweig, M. R., and Hassan, M. N. (1990), "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries", *The American Economic Review*, 80(5), 1139-1156.

Popli, G. K. (2012),"Gender Wage Differentials in Mexico: A Distributional Approach", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. Retrieved from http://dx.doi.org/10.1111/j.1467-985X.2011.01031.x

Psacharopoulos, G., and Patrinos, H. A. (2004), "Returns to Investment in Education: A Further Update", *Education Economics*, 12(2), 111-134.

Psacharopoulos, G., and Yang, H. (1991), "Educational Attainment among Venezuelan Youth: An Analysis of its Determinants", *International Journal of Educational Development*, 11(4), 289-294.

Rahman, M. M., Khanam, R., and Absar, N. U. (1999), "Child Labor in Bangladesh: A Critical Appraisal of Harkin's Bill and the MOU-Type Schooling Program", *Journal of Economic Issues*, 33(4), 985-1003.

Rammohan, A. (2012), "The Trade-off between Child Labour and Schooling in India", *Education Economics*, *forthcoming*.

Ravallion, M., and Wodon, Q. (2000), "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy", *The Economic Journal*, 110(462), C158-C175.

Ray, R. (2000a), "Analysis of Child Labour in Peru and Pakistan: A Comparative Study", *Journal of Population Economics*, 13(1), 3-19.

Ray, R. (2000b), "Poverty, Household Size and Child Welfare in India", *Economic and Political Weekly*, 35(39), 3511-3520.

Ray, R. (2002), "Simultaneous Analysis of Child Labour and Child Schooling: Comparative Evidence from Nepal and Pakistan", *Economic and Political Weekly*, 37(52), 5215-5224.

Ray, R. (2004), "Child Labour: A Survey of Selected Asian Countries", *Asian - Pacific Economic Literature*, 18(2), 1-18.

Ray, R., and Lancaster, G. (2005), "The Impact of Children's Work on Schooling: Multi-Country Evidence", *International Labour Review*, 144(2), 189-210.

Reimers, C. (1983), "Labour Market Discrimination against Hispanic and Blank Men", *The Review of Economics and Statistics*, 65(4), 570 - 579.

Reinhold, S., and Jürges, H. (2011), "Parental Income and Child Health in Germany", *Health Economics, forthcoming.*

Ridao-Cano, C. (2001), "Child Labor and Schooling in a Low Income Rural Economy", Population Aging Center Working Paper PAC 2001-0001, Institute of Behavioral Science, University of Colorado at Boulder.

Robinson, P. M. (1988), "Root-N-Consistent Semiparametric Regression", *Econometrica*, 56(4), 931-954.

Rosati, F., and Rossi, M. (2003), "Children's Working Hours and School Enrollment: Evidence from Pakistan and Nicaragua", *The World Bank Economic Review*, 17(2), 283-295.

Sakellariou, C. (2004), "The Use of Quantile Regressions in Estimating Gender Wage Differentials: A Case Study of the Philippines", *Applied Economics*, 36(9), 1001-1007.

Sasin, M. (2007), "The Role of Employment and Earnings in Shared Growth", Discussion Paper, A World Bank Labor Market Study.

Sedlacek, G., Duryea, S., Ilahi, N., and Sasaki, M. (2009), "Child Labour, Schooling and Poverty in Latin America", in (Orazem, F. P., Sedlacek, G. and Tzannatos, Z.) (Ed.), *Child Labour and Education in Latin America: An Economic Perspective*: Palgrave Macmillan.

Sen, A. (1985), "Well-Being, Agency and Freedom: The Dewey Lectures 1984", *The Journal of Philosophy*, 82(4), 169-221.

Shafiq, M. N. (2007), "Household Schooling and Child Labor Decisions in Rural Bangladesh", *Journal of Asian Economics*, 18(6), 946-966.

Smith, J. P. (1999), "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status", *The Journal of Economic Perspectives*, 13(2), 145-166.

Steckel, R. H. (1995), "Stature and the Standard of Living", *Journal of Economic Literature*, 33(4), 1903-1940.
Swaminathan, M. (1998), "Economic Growth and the Persistence of Child labor: Evidence from an Indian City", *World Development*, 26(8), 1513-1528.
Trejo, S. J. (1993), "Overtime Pay, Overtime Hours, and Labor Unions", *Journal of Labor Economics*, 11(2), 253.
Udry, C. (2006). Child Labour. In Banerjee, A. V., Benabou, R. and Mookherjee, D. (Eds.), *Understanding Poverty*. New York: Oxford University Press.
Van de Ven, W. P. M. M., and Van Praag, B. M. S. (1981), "The Demand for Deductibles in Private Health Insurance : A Probit Model with Sample Selection", *Journal of Econometrics*, 17(2), 229-252.
Visaria, P. (1971), "The Sex Ratio of the Population in India. Census of India 1961", Monograph No. 10. Manager of Publications, Delhi.
Wahba, J. (2006), "The Influence of Market Wages and Parental History on Child Labour and Schooling in Egypt", *Journal of Population Economics*, 19(4), 823-852.
Waldron, I. (1985), "What Do We Know About Causes of Sex Differences in Mortality? A Review of the Literature", *Population Bulletin of the United Nations*, 18, 59-76.
WDB. (2012), "Gender Statistics Highlights", World Data Bank, World Bank, Washington D.C.
Wei, C., and Bo, L. (2007), "Glass Ceiling or Sticky Floor? Examining the Gender Pay Gap across the Wage Distribution in Urban China, 1987-2004", MPRA Paper No. 3544.
Wellington, A. (1993), "Changes in the Male/Female Wage Gap, 1976-85", *The Journal of Human Resources*, 28(2), 383-411.
Willis, R. J., and Rosen, S. (1979), "Education and Self-Selection", *Journal of Political Economy*, 87(5), 7-36.
Wolff, F.-C., and Maliki. (2008), "Evidence on the Impact of Child Labor on Child Health in Indonesia, 1993-2000", *Economics and Human Biology*, 6(1), 143-169.
Yamauchi-Kawana, F. (1997), "Information, Neighbourhood Effects and Investment in Human Capital: Learning Schooling Returns in a Dynamic Context", Mimeo, University of Pennsylvania.
Zabaleta, M. B. (2011), "The Impact of Child Labor on Schooling Outcomes in Nicaragua", *Economics of Education Review*, 30(6), 1527-1539.

Appendix A

Table A1: Binary Probit Specification of the Probability of Full-time Employment, LFS 1999-2005

	LFS 1999						LFS 2005					
	Males (N = 10,674)			Females (N = 1,978)			Males (N = 43,722)			Females (N = 13,352)		
Variables	ME		Std. Err.	ME		Std. Err.	ME		Std. Err.	ME		Std. Err.
Age 15-19	0.3218	***	0.0305	0.3698	***	0.0951	0.2512	***	0.0169	0.4557	***	0.0509
Age 20-24	0.2730	***	0.0299	0.3506	***	0.0958	0.2059	***	0.0157	0.1512	***	0.0350
Age 25-29	0.2221	***	0.0280	0.3763	***	0.0926	0.1799	***	0.0146	0.1664	***	0.0346
Age 30-34	0.2069	***	0.0276	0.4067	***	0.0861	0.1754	***	0.0145	0.1849	***	0.0361
Age 35-39	0.1956	***	0.0274	0.4200	***	0.0816	0.1681	***	0.0143	0.1785	***	0.0354
Age 40-44	0.1940	***	0.0278	0.3856	***	0.0842	0.1546	***	0.0146	0.1564	***	0.0353
Age 45-49	0.1943	***	0.0285	0.4061	***	0.0795	0.1332	***	0.0147	0.1301	***	0.0349
Age 50-54	0.2101	***	0.0292	0.2889	**	0.1027	0.1268	***	0.0155	0.1008	***	0.0345
Age 55-59	0.0959	***	0.0336	0.1694		0.1349	0.1021	***	0.0167	0.1055	***	0.0376
Primary school	-0.0827	***	0.0125	-0.0534	*	0.0320	-0.0339	***	0.0063	-0.0265	***	0.0077
Secondary school	-0.0902	***	0.0136	-0.2616	***	0.0301	-0.0550	***	0.0066	-0.0326	***	0.0085
Post-secondary school	0.0046	*	0.0172	-0.0750	*	0.0415	0.0710	***	0.0086	0.2413	***	0.0199
Graduate	0.1146	***	0.0247	0.1251	*	0.0725	0.3236	***	0.0113	0.6271	***	0.0280
Married	0.0263		0.0200	-0.0338		0.0409	0.0166	*	0.0091	-0.1119	***	0.0178
Divorced	0.1410		0.0977	0.1982	***	0.0409	0.0111		0.0578	0.1129	***	0.0294
Widowed	-0.0210		0.1472	0.2773	***	0.0673	-0.0219		0.0306	0.0484	***	0.0198
Number of children 0-5	-0.0142	**	0.0062	0.0079		0.0160	-0.0150	***	0.0033	-0.0229	***	0.0044
Number of children 6-12	-0.0111	**	0.0052	-0.0367	***	0.0132	-0.0153	***	0.0027	-0.0125	***	0.0035
Household pays no rent	0.0348		0.0293	0.0301		0.0537	0.2152	***	0.0174	0.3426	***	0.0293
Household pays rent	0.1180	***	0.0141	0.2419	***	0.0289	0.1663	***	0.0093	0.3199	***	0.0197
Number of males 65 or higher	-0.0394		0.0252	0.0672		0.0518	-0.0813	***	0.0098	-0.0085		0.0124
Number of females 65 or higher	-0.0215		0.0280	-0.0142		0.0701	-0.0365	***	0.0099	-0.0095		0.0134
Urban	0.02138	*	0.0124	0.1097	***	0.0298	0.04874	***	0.0057	0.0764	***	0.0079
Quintile2	-0.1836	***	0.0142	-0.0989	**	0.0394	-0.0999	***	0.0069	-0.0309	***	0.0082
Quintile3	-0.2064	***	0.0137	-0.0817	**	0.0383	-0.1329	***	0.0067	-0.0533	***	0.0078
Quintile4	-0.2021	***	0.0142	0.0421		0.0377	-0.1501	***	0.0069	-0.0566	***	0.0080
Quintile5	-0.1975	***	0.0158	0.0258		0.0390	-0.1627	***	0.0081	-0.0180		0.0107

Notes: Data are from LFS 1999 and LFS 2005. ME is marginal effects. Std. Err. is standard errors. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table A2: OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N = 890)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Age 15-19	-0.3180 ***	-0.5697 ***	-0.4707 ***	-0.3326 ***	-0.2790 ***	0.1263	-0.0040	0.4821	0.0277	-0.0112	-0.8865	-0.3590 **
	(0.0612)	(0.1041)	(0.0847)	(0.0999)	(0.1019)	(0.1171)	(0.2567)	(0.6391)	(0.4113)	(0.3835)	(0.6099)	(0.1403)
Age 20-24	-0.0927	-0.0887	-0.0894	-0.0954	-0.2561 ***	0.0405	0.2029	0.6157	0.1661	0.1158	-0.6188	-0.1566
	(0.0574)	(0.0904)	(0.0771)	(0.0948)	(0.0983)	(0.1126)	(0.2525)	(0.6398)	(0.4087)	(0.3823)	(0.6032)	(0.1369)
Age 25-29	-0.0493	-0.0602	-0.0298	0.0228	-0.1336	-0.0392	0.1480	0.5556	0.1167	0.1007	-0.5988	-0.0797
	(0.0529)	(0.0822)	(0.0702)	(0.0871)	(0.0920)	(0.1080)	(0.2520)	(0.6404)	(0.4098)	(0.3828)	(0.6013)	(0.1283)
Age 30-34	0.0432	0.0270	0.0370	0.1666 *	-0.0626	-0.0102	0.3956	0.6887	0.3384	0.3642	-0.2569	-0.0101
	(0.0518)	(0.0781)	(0.0675)	(0.0846)	(0.0908)	(0.1094)	(0.2520)	(0.6408)	(0.4090)	(0.3819)	(0.6023)	(0.1376)
Age 35-39	0.0703	0.0426	0.0185	0.1409 *	0.0531	0.0969	0.2873	0.5975	0.2199	0.1658	-0.4000	0.2410
	(0.0516)	(0.0777)	(0.0675)	(0.0837)	(0.0908)	(0.1111)	(0.2526)	(0.6443)	(0.4106)	(0.3811)	(0.5998)	(0.1498)
Age 40-44	0.1316 **	0.0326	0.0415	0.2101 **	0.1259	0.2277 **	0.3238	0.6190	0.1935	0.2465	-0.1945	0.1659
	(0.0521)	(0.0783)	(0.0679)	(0.0848)	(0.0920)	(0.1143)	(0.2556)	(0.6464)	(0.4146)	(0.3876)	(0.6043)	(0.1336)
Age 45-49	0.1746 ***	0.0648	0.0505	0.1837 **	0.2139 **	0.4097 ***	0.3422	0.6359	0.3188	0.1711	-0.2270	0.5136 ***
	(0.0536)	(0.0786)	(0.0687)	(0.0871)	(0.0958)	(0.1256)	(0.2558)	(0.6474)	(0.4123)	(0.3860)	(0.6053)	(0.1983)
Age 50-54	0.1575 ***	-0.0143	-0.0041	0.1650 *	0.1653 *	0.5578 ***	0.3093	0.6575	0.1145	0.1728	-0.1328	0.1264
	(0.0551)	(0.0835)	(0.0716)	(0.0887)	(0.0966)	(0.1306)	(0.2646)	(0.6628)	(0.4353)	(0.3945)	(0.6180)	(0.1653)
Age 55-59	0.2202 ***	0.1066	0.0841	0.2011 *	0.2485 **	0.4790 ***	0.4512	0.9709	0.2570	0.2352	0.2719	0.3439
	(0.0623)	(0.0860)	(0.0778)	(0.1019)	(0.1123)	(0.1604)	(0.3062)	(0.6522)	(0.4842)	(0.4435)	(0.7246)	(0.3600)
Primary school	0.0871 ***	0.1496 ***	0.1242 ***	0.2399 ***	0.0482	-0.1992 ***	0.0750	-0.0238	0.1624 **	0.2503 ***	0.1546	-0.1716 **
	(0.0217)	(0.0336)	(0.0302)	(0.0378)	(0.0348)	(0.0353)	(0.0589)	(0.0751)	(0.0751)	(0.0890)	(0.1200)	(0.0857)
Secondary school	0.1447 ***	0.1533 ***	0.1895 ***	0.4097 ***	0.2694 ***	-0.3631 ***	0.1143	-0.0293	0.1809 **	0.3673 ***	0.3173 *	-0.2892 *
	(0.0243)	(0.0335)	(0.0300)	(0.0425)	(0.0497)	(0.0519)	(0.0767)	(0.0799)	(0.0812)	(0.1107)	(0.1670)	(0.1693)
Post-secondary school	0.5861 ***	0.1736 ***	0.1978 ***	0.5610 ***	0.7863 ***	0.1462	0.3441 ***	-0.1153	0.0603	0.3056 ***	0.1892	0.0440
	(0.0339)	(0.0304)	(0.0276)	(0.0432)	(0.0642)	(0.1046)	(0.1163)	(0.1138)	(0.0890)	(0.1161)	(0.2144)	(0.3182)
Graduate	0.8696 ***	0.1733 ***	0.1978 ***	0.5594 ***	0.9203 ***	1.0103 ***	0.7453 ***	-0.0599	0.1639 **	0.4415 ***	0.5462 **	0.2635
	(0.0455)	(0.0365)	(0.0304)	(0.0468)	(0.0799)	(0.1691)	(0.1203)	(0.0736)	(0.0698)	(0.1052)	(0.2566)	(0.4856)
General	0.0364	-0.0950	0.0196	0.1304 **	0.1235	0.2472	0.3640 ***	0.1269 *	0.1938 **	0.2239 *	0.4091 *	0.6844
	(0.0497)	(0.0631)	(0.0412)	(0.0652)	(0.0984)	(0.1745)	(0.1354)	(0.0703)	(0.0773)	(0.1326)	(0.2295)	(0.4874)
Vocational	0.2265 ***	0.0622 **	0.0589 **	0.2008 ***	0.3604 ***	0.5646 ***	0.2676 **	-0.0389	0.1778 *	0.2099	0.7798 **	-0.0211
	(0.0382)	(0.0302)	(0.0320)	(0.0562)	(0.0957)	(0.1408)	(0.1227)	(0.1594)	(0.0984)	(0.1914)	(0.3844)	(0.5059)

Continued on next page

Table A2 (continued): OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N = 890)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Married	0.1021 ***	-0.0746	0.0183	0.1325 **	0.1032 *	0.0895	-0.0976	-0.0098	-0.0927	-0.1504 *	-0.1968	-0.2537
	(0.0302)	(0.0474)	(0.0405)	(0.0514)	(0.0553)	(0.0638)	(0.0690)	(0.0795)	(0.0737)	(0.0858)	(0.1583)	(0.1701)
Widowed	-0.1399	-0.8731	-0.6582 **	-0.2623	0.2255	0.2957	-0.2690 **	-0.3856 **	-0.2427 *	-0.3209 **	-0.4805 ***	-0.2678
	(0.2621)	(0.5395)	(0.2830)	(0.2117)	(0.3605)	(0.6736)	(0.1043)	(0.1917)	(0.1449)	(0.1353)	(0.1766)	(0.2128)
Divorced	0.0154	0.0601	0.0003	-0.0780	-0.1493	-0.0806	-0.0296	0.0981	0.0561	0.0592	-0.3839 *	-0.3725 *
	(0.1073)	(0.1895)	(0.2208)	(0.1812)	(0.1301)	(0.3100)	(0.0989)	(0.1300)	(0.1171)	(0.1228)	(0.1973)	(0.2009)
Urban	0.1631 ***	0.1009 ***	0.1408 ***	0.2754 ***	0.2280 ***	0.3292 ***	0.1161 *	0.1179	0.0652	0.1875 **	0.1981 *	0.2710 **
	(0.0191)	(0.0245)	(0.0256)	(0.0356)	(0.0349)	(0.0457)	(0.0603)	(0.0878)	(0.0808)	(0.0874)	(0.1117)	(0.1166)
Professional	0.1269 ***	-0.1090 ***	-0.0954 ***	0.2180 ***	0.6349 ***	0.6102 ***	0.7279 ***	0.0157	-0.0280	0.3657 **	1.6156 ***	2.8613 ***
	(0.0381)	(0.0355)	(0.0342)	(0.0576)	(0.0780)	(0.1087)	(0.1462)	(0.0732)	(0.1330)	(0.1826)	(0.3337)	(0.4593)
Administrative	0.3136 ***	-0.0237	0.0025	0.4055 ***	0.8388 ***	1.6036 ***	0.6004 ***	0.0264	0.1170	0.4859 **	1.2876 ***	2.5561 ***
	(0.0519)	(0.0362)	(0.0362)	(0.0700)	(0.1025)	(0.1744)	(0.1664)	(0.1053)	(0.1605)	(0.1986)	(0.3602)	(0.4894)
Clerical	0.0224	-0.1184 *	-0.1434 **	0.1134	0.0931	0.2967 **	0.0023	-0.4485	-0.3243	-0.1454	0.6718	0.3958
	(0.0567)	(0.0640)	(0.0589)	(0.0909)	(0.1073)	(0.1397)	(0.2541)	(0.3635)	(0.2618)	(0.3309)	(0.5467)	(0.7697)
Sales	-0.1158 ***	-0.2737 ***	-0.2517 ***	-0.0100	0.1302	-0.0923	-0.3814 ***	-0.6205 ***	-0.5889 ***	-0.4896 ***	-0.4237 *	0.0498
	(0.0435)	(0.0556)	(0.0471)	(0.0719)	(0.0836)	(0.0888)	(0.0891)	(0.0934)	(0.1339)	(0.1666)	(0.2421)	(0.1066)
Agriculture	-0.1173 **	-0.2686 ***	-0.2757 ***	-0.1351	0.0842	0.2218 *	-0.1962 *	-0.0471	-0.1416	-0.4543 **	-0.0627	0.3392 *
	(0.0486)	(0.0711)	(0.0573)	(0.0823)	(0.0974)	(0.1121)	(0.1185)	(0.1496)	(0.1801)	(0.2152)	(0.3094)	(0.1922)
Production	0.0029	-0.1176 ***	-0.0843 **	0.0830	0.1452 **	0.1235 **	0.0105	-0.0846	-0.1017	0.1629	-0.2450	0.1458 *
	(0.0281)	(0.0337)	(0.0327)	(0.0570)	(0.0573)	(0.0549)	(0.0853)	(0.0554)	(0.1081)	(0.1633)	(0.2482)	(0.0763)
Agriculture	-0.1837 ***	0.0192	-0.2327 ***	-0.4114 ***	-0.4249 ***	-0.1122	-0.0590	0.1535	-0.1353	-0.0810	-0.4480 **	-0.2102
	(0.0536)	(0.0764)	(0.0603)	(0.0861)	(0.1045)	(0.1219)	(0.1056)	(0.1872)	(0.1917)	(0.1776)	(0.2024)	(0.0763)
Manufacturing	-0.0195	0.0407	0.0032	-0.0555	-0.20435 ***	0.0302	0.0243	0.2535 ***	0.2438 **	-0.0088	-0.0650	-0.0254
	(0.0341)	(0.0398)	(0.0347)	(0.0580)	(0.0679)	(0.0847)	(0.0668)	(0.0867)	(0.1060)	(0.1031)	(0.1410)	(0.0763)

Continued on next page

Table A2 (continued): OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N= 890)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Wholesale and Retail	-0.1256 **	-0.0362	-0.0394	-0.0888	-0.24124 **	-0.0533	-0.2333	0.2343	0.2624	-0.2666	-0.7045	-0.4004
	(0.0571)	(0.0675)	(0.0566)	(0.0881)	(0.1054)	(0.1366)	(0.2592)	(0.2847)	(0.2163)	(0.3576)	(0.4869)	(0.6767)
Transport	0.0796 *	0.0384	-0.0169	0.1589 **	-0.00038	0.1278	0.0438	0.2475 *	0.1371	-0.1477	0.2735	-0.7451
	(0.0407)	(0.0429)	(0.0415)	(0.0715)	(0.0892)	(0.1049)	(0.3015)	(0.1305)	(0.2627)	(0.3810)	(0.8156)	(0.6836)
Financial institution	0.3704 ***	0.1048 **	0.1089 ***	0.3418 ***	0.6291 ***	1.8997 ***	0.4442 ***	0.2729 *	0.3625 ***	0.4786 ***	0.5846	1.7420 ***
	(0.0649)	(0.0435)	(0.0387)	(0.0657)	(0.1114)	(0.2625)	(0.1591)	(0.1408)	(0.1149)	(0.1437)	(0.3821)	(0.5342)
Real estate	0.1625	0.1180 **	0.0780	0.3336 **	0.5136 **	0.1032	0.3975 *	0.4347 ***	0.5827 ***	0.8628 ***	0.5386	0.4410
	(0.1036)	(0.0473)	(0.0930)	(0.1367)	(0.2127)	(0.4245)	(0.2100)	(0.1339)	(0.1465)	(0.2366)	(0.3895)	(1.0270)
Public adminisitration	0.3235 ***	0.0578	0.0937 ***	0.3695 ***	0.4991 ***	1.2140 ***	0.4269 ***	0.2785 ***	0.3911 ***	0.5072 ***	1.2750 ***	0.1598
	(0.0418)	(0.0384)	(0.0316)	(0.0525)	(0.0856)	(0.1446)	(0.1374)	(0.0876)	(0.0975)	(0.1227)	(0.2531)	(0.5581)
Education	-0.0585	0.0602	0.0598	0.1382 *	-0.0092	-0.3796 *	-0.1748	0.1761	0.0809	-0.0545	0.1241	-0.7994
	(0.0536)	(0.0423)	(0.0413)	(0.0789)	(0.1234)	(0.1952)	(0.1562)	(0.1131)	(0.1463)	(0.1553)	(0.3293)	(0.5085)
Health	0.1837 *	0.0033	0.0034	0.1703	0.4968 ***	0.6993 **	0.3205 *	0.1506	0.2175	0.2880	0.9915 ***	-0.3527
	(0.0979)	(0.0772)	(0.0628)	(0.1036)	(0.1576)	(0.2980)	(0.1718)	(0.1984)	(0.1582)	(0.1813)	(0.3520)	(0.4629)
Constant	1.8024 ***	1.4069 ***	1.6990 ***	1.5664 ***	2.2419 ***	2.3951 ***	1.3189 ***	0.2263	0.9556 **	1.3162 ***	2.4636 ***	2.4938 ***
	(0.0652)	(0.0946)	(0.0856)	(0.1154)	(0.1203)	(0.1421)	(0.3345)	(0.6401)	(0.4320)	(0.4227)	(0.6586)	(0.2192)

Notes: Data are from LFS 1999. Robust standard errors are in parentheses. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table A3: OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 2005

	LFS 2005											
	Males (N = 16,106)						Females (N = 2,286)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Age 15-19	-0.1132 ***	-0.3635 ***	-0.0317	0.0065	-0.0796 **	-0.1452 ***	-0.3340 **	0.0701	-0.4555 *	-0.3184	-0.3860 **	-0.1146
	(0.0427)	(0.0945)	(0.0812)	(0.0453)	(0.0395)	(0.0472)	(0.1612)	(0.3920)	(0.2429)	(0.1936)	(0.1874)	(0.2388)
Age 20-24	-0.0517	-0.1566 *	-0.1171	0.0202	-0.0432	-0.0638	-0.3092 *	0.2569	-0.3380	-0.4078 **	-0.4002 **	-0.1754
	(0.0376)	(0.0797)	(0.0735)	(0.0423)	(0.0371)	(0.0443)	(0.1518)	(0.3718)	(0.2266)	(0.1848)	(0.1804)	(0.2265)
Age 25-29	-0.0077	-0.0805	-0.1161	0.0111	-0.0120	-0.0145	-0.1728	0.2585	-0.1687	-0.1519	-0.3102 *	-0.2326
	(0.0344)	(0.0706)	(0.0672)	(0.0394)	(0.0349)	(0.0417)	(0.1492)	(0.3675)	(0.2189)	(0.1794)	(0.1769)	(0.2212)
Age 30-34	0.0414	0.0759	-0.0353 *	0.0462	0.0052	0.0199	-0.0351	0.3901	-0.0645	-0.0463	-0.2231	-0.0997
	(0.0336)	(0.0666)	(0.0656)	(0.0388)	(0.0345)	(0.0413)	(0.1473)	(0.3663)	(0.2163)	(0.1773)	(0.1759)	(0.2205)
Age 35-39	0.0603	0.0930	0.0114	0.0537	0.0380	0.0273	0.0228	0.3528	-0.0050	-0.0203	-0.0001	0.0579
	(0.0329)	(0.0651)	(0.0646)	(0.0385)	(0.0342)	(0.0408)	(0.1469)	(0.3663)	(0.2151)	(0.1758)	(0.1751)	(0.2207)
Age 40-44	0.0769 **	0.0814	0.0916	0.0732 *	0.0378	0.0641	0.1639	0.3249	0.0285	0.2659	0.1528	0.2377
	(0.0338)	(0.0665)	(0.0657)	(0.0391)	(0.0348)	(0.0419)	(0.1484)	(0.3695)	(0.2170)	(0.1767)	(0.1773)	(0.2263)
Age 45-49	0.0885 **	0.0591	0.1309	0.0760 *	0.0433	0.0375	0.1875	0.4016	0.1389	0.2946	0.2244	0.2430
	(0.0342)	(0.0685)	(0.0668)	(0.0400)	(0.0353)	(0.0423)	(0.1491)	(0.3694)	(0.2188)	(0.1813)	(0.1817)	(0.2321)
Age 50-54	0.0735 *	0.0176	0.1060 *	0.0789 *	0.0651 *	0.0336	0.0952	0.3947	-0.0078	0.1280	0.0430	0.0103
	(0.0364)	(0.0727)	(0.0706)	(0.0419)	(0.0370)	(0.0441)	(0.1584)	(0.3801)	(0.2320)	(0.1914)	(0.1893)	(0.2319)
Age 55-59	0.1060 **	0.0985	0.2218 ***	0.0806 *	0.0395	0.0407	0.1622	0.2526	0.1495	0.3516 *	-0.0064	0.2230
	(0.0388)	(0.0767)	(0.0747)	(0.0453)	(0.0394)	(0.0477)	(0.1781)	(0.4177)	(0.2455)	(0.2067)	(0.2009)	(0.2638)
Primary school	-0.0030	-0.0241	-0.1903 ***	-0.0270	0.0844 ***	0.1125 ***	-0.0707	0.1653	-0.0959	-0.2540 ***	-0.1545 ***	0.0573
	(0.0169)	(0.0353)	(0.0301)	(0.0175)	(0.0165)	(0.0212)	(0.0552)	(0.1075)	(0.0918)	(0.0685)	(0.0569)	(0.0789)
Secondary school	-0.1601 ***	-0.1723 ***	-0.4707 ***	-0.2041 ***	-0.0242	-0.0030	-0.0062	0.1572	0.0201	-0.1451 *	-0.0552	0.1054
	(0.0195)	(0.0432)	(0.0358)	(0.0190)	(0.0174)	(0.0215)	(0.0640)	(0.1240)	(0.1083)	(0.0821)	(0.0686)	(0.0911)
Post-secondary school	-0.1894 ***	0.1880 ***	-0.3356 ***	-0.3733 ***	-0.1790 ***	-0.1539 ***	0.3342 ***	0.3997 **	0.5161 ***	0.3833 ***	0.1438 *	0.1874
	(0.0247)	(0.0594)	(0.0516)	(0.0227)	(0.0189)	(0.0246)	(0.0878)	(0.1577)	(0.1381)	(0.1158)	(0.0867)	(0.1227)
Graduate	0.2156 ***	0.5274 ***	0.4192 ***	0.0739 **	0.0800 ***	0.0811 **	0.7189 ***	0.5781 ***	0.8118 ***	0.9282 ***	0.5762 ***	0.6663 ***
	(0.0304)	(0.0650)	(0.0635)	(0.0309)	(0.0254)	(0.0330)	(0.0944)	(0.1667)	(0.1415)	(0.1275)	(0.1027)	(0.1434)
General	-0.1398 ***	-0.1394 *	-0.2712 ***	-0.1481 ***	-0.0535 **	-0.0392	-0.1225 **	0.0781	-0.0036	-0.1675 **	-0.0882	-0.2537 ***
	(0.0328)	(0.0720)	(0.0631)	(0.0290)	(0.0252)	(0.0341)	(0.0551)	(0.1009)	(0.1087)	(0.0760)	(0.0660)	(0.0786)

Continued on next page

Table A3 (continued): OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 2005

	LFS 2005											
	Males (N = 16,106)						Females (N = 2,286)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Vocational	-0.0600	-0.1028	0.0283 **	-0.1505 ***	-0.0361	0.0250	-0.5331 ***	-0.6484	-0.7625 ***	-0.5244 ***	-0.4078 ***	-0.4563 ***
	(0.0603)	(0.1256)	(0.1075)	(0.0538)	(0.0480)	(0.0669)	(0.1250)	(0.4204)	(0.2455)	(0.1878)	(0.1103)	(0.0734)
Married	0.1833 ***	0.3875 ***	0.3213 ***	0.1347 ***	0.0679 ***	0.0319	0.0342	0.0257	0.0772	0.0707	-0.0257	0.0198
	(0.0235)	(0.0549)	(0.0464)	(0.0227)	(0.0194)	(0.0245)	(0.0604)	(0.1236)	(0.0987)	(0.0796)	(0.0632)	(0.0843)
Widowed	-0.0137	-0.0836	0.2095 **	-0.0385	0.0322	0.1279	0.0113	-0.3511	0.0880	0.0248	-0.0782	-0.0467
	(0.1235)	(0.2061)	(0.1509)	(0.0863)	(0.0697)	(0.0961)	(0.1089)	(0.2346)	(0.1304)	(0.1119)	(0.0936)	(0.1200)
Divorced	-0.0181	0.0952	0.1073	-0.1144	-0.2382 **	-0.1288	0.0291	0.2120	0.0115	0.1121	-0.0144	-0.1137
	(0.0993)	(0.3413)	(0.2573)	(0.1590)	(0.1030)	(0.1393)	(0.0853)	(0.1679)	(0.1588)	(0.1272)	(0.1036)	(0.1313)
Urban	-0.0006	-0.0486	0.0032	0.0137	0.0414 ***	0.0413 ***	-0.0520	-0.0409	-0.0928	-0.1315 **	0.0299	0.1359 **
	(0.0138)	(0.0310)	(0.0268)	(0.0133)	(0.0117)	(0.0150)	(0.0416)	(0.0816)	(0.0637)	(0.0544)	(0.0436)	(0.0546)
Professional	0.1009 ***	0.2250 **	0.2177 ***	0.0083	-0.0135	-0.0256	0.3957 ***	0.4558 *	0.2110	0.5526 ***	0.4572 ***	0.3027 *
	(0.0375)	(0.0901)	(0.0829)	(0.0343)	(0.0257)	(0.0346)	(0.1207)	(0.2471)	(0.1983)	(0.1604)	(0.1130)	(0.1602)
Administrative	-0.0091	0.2021	0.2294 **	-0.0715	-0.0850	-0.1662 **	0.7465 ***	0.6028 **	0.5704 **	0.6175 *	0.8349 **	1.1052
	(0.0720)	(0.1809)	(0.1701)	(0.0788)	(0.0625)	(0.0720)	(0.2554)	(0.2961)	(0.2809)	(0.3485)	(0.3857)	(0.7561)
Clerical	0.2037 ***	0.1625 *	0.3168 ***	0.2222 ***	0.0668 **	0.0320	0.2673 **	0.4806 **	0.1890	0.3457 **	0.4214 ***	-0.0658
	(0.0366)	(0.0837)	(0.0779)	(0.0357)	(0.0262)	(0.0340)	(0.1132)	(0.2319)	(0.1841)	(0.1662)	(0.1226)	(0.1571)
Sales	-0.1935 ***	-0.2967 ***	-0.2252 ***	-0.1960 ***	-0.1321 ***	-0.1523 ***	0.0929	0.4068	0.2060	0.0102	0.1379	0.0006
	(0.0391)	(0.1029)	(0.0814)	(0.0313)	(0.0235)	(0.0309)	(0.1373)	(0.3055)	(0.2453)	(0.1946)	(0.1419)	(0.1774)
Agriculture	0.3262 ***	0.7716 ***	0.9965 ***	0.1945 ***	0.0499	-0.0862 *	0.4845 ***	0.9032 ***	0.9621 ***	1.2337 ***	-0.1408	-0.1733
	(0.0496)	(0.1139)	(0.0898)	(0.0469)	(0.0390)	(0.0480)	(0.1227)	(0.2528)	(0.1767)	(0.1564)	(0.1282)	(0.1573)
Production	0.2822 ***	0.3417 ***	0.4158 ***	0.2665 ***	0.1855 ***	0.1491 ***	0.2486 **	0.3833 *	0.1981	0.2934 **	0.2870 ***	0.1573
	(0.0311)	(0.0748)	(0.0627)	(0.0276)	(0.0223)	(0.0295)	(0.0985)	(0.2125)	(0.1613)	(0.1274)	(0.0986)	(0.1386)
Agriculture	0.2525 ***	0.4611 ***	0.4425 ***	0.2651 ***	-0.0724 *	-0.1069 **	0.7878 ***	1.0717 ***	0.8477 ***	0.3658 **	0.6536 ***	0.3965 **
	(0.0550)	(0.1275)	(0.0980)	(0.0502)	(0.0422)	(0.0518)	(0.1378)	(0.2823)	(0.1909)	(0.1656)	(0.1405)	(0.1806)
Manufacturing	0.0832 **	0.3887 ***	0.1227 **	0.0218	-0.0293	-0.0182	0.3872 ***	1.2922 ***	0.7692 ***	0.0024	-0.0466	-0.1027
	(0.0396)	(0.0967)	(0.0754)	(0.0327)	(0.0274)	(0.0352)	(0.1095)	(0.2389)	(0.1740)	(0.1338)	(0.1038)	(0.1442)

Continued on next page

Table A3 (continued): OLS and Unconditional Quantile Regression Estimates without Selectivity Bias Correction, by Gender, LFS 2005

LFS 2005												
	Males (N = 16,106)						Females (N= 2,286)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Wholesale and Retail	-0.1413 ***	-0.4485 ***	-0.3132 ***	-0.0236	-0.0779 **	-0.0226	0.8786 ***	0.9671 **	1.2664 ***	1.0706 ***	0.4715 *	-0.0903
	(0.0505)	(0.1278)	(0.0935)	(0.0391)	(0.0320)	(0.0418)	(0.2086)	(0.4430)	(0.2954)	(0.2628)	(0.2763)	(0.2796)
Transport	0.1638 ***	0.5846 ***	0.3426 ***	0.0006	-0.0484	0.0182	0.8616 ***	1.6406 ***	1.5009 ***	0.7208 ***	0.3773 *	0.0387
	(0.0423)	(0.0976)	(0.0793)	(0.0363)	(0.0312)	(0.0421)	(0.1521)	(0.1999)	(0.1749)	(0.2476)	(0.2231)	(0.3172)
Financial institution	0.1738 ***	0.5157 ***	0.2693 **	0.0637	0.0297	0.0648	0.2839 **	0.7541 **	0.6720 ***	-0.0552	-0.1770	0.0209
	(0.0529)	(0.1138)	(0.1050)	(0.0486)	(0.0413)	(0.0553)	(0.1336)	(0.2943)	(0.2057)	(0.1661)	(0.1300)	(0.1988)
Real estate	0.0478	0.2170	0.0875 **	-0.0406	-0.0733	-0.0082	0.4661 **	1.2255 ***	0.5997	-0.1593	0.0407	0.0595
	(0.0853)	(0.1980)	(0.1692)	(0.0760)	(0.0652)	(0.0907)	(0.2714)	(0.2498)	(0.3885)	(0.3110)	(0.2934)	(0.4697)
Public adminisitration	0.0758 *	0.5767 ***	0.5810 ***	-0.1034 ***	-0.2299 ***	-0.2003 ***	0.4495 ***	1.1785 ***	0.8077 ***	0.4312 **	-0.2698 **	-0.2675
	(0.0415)	(0.1018)	(0.0845)	(0.0353)	(0.0270)	(0.0349)	(0.1173)	(0.2430)	(0.1917)	(0.1720)	(0.1254)	(0.1630)
Education	-0.2289 ***	0.1846	-0.2350 **	-0.3330 ***	-0.3059 ***	-0.2564 ***	0.0602	1.0067 ***	0.7681 ***	-0.2727 *	-0.6733 ***	-0.8131 ***
	(0.0474)	(0.1177)	(0.0978)	(0.0404)	(0.0320)	(0.0427)	(0.1235)	(0.2608)	(0.1980)	(0.1566)	(0.1208)	(0.1786)
Health	-0.0234	0.2867 **	-0.0774 **	-0.1320 ***	-0.1250 ***	-0.0913	0.2449 **	0.9077 ***	0.7727 ***	0.1041	-0.3048 ***	-0.3299 **
	(0.0600)	(0.1423)	(0.1238)	(0.0504)	(0.0417)	(0.0557)	(0.1277)	(0.2812)	(0.2060)	(0.1623)	(0.1156)	(0.1623)
Constant	3.0416 ***	1.3204 ***	2.2050 ***	3.3666 ***	3.9231 ***	4.3106 ***	2.1728 ***	-0.7763 *	0.6505 ***	1.9783 ***	2.9769 ***	3.2916 ***
	(0.0522)	(0.1168)	(0.0985)	(0.0508)	(0.0446)	(0.0550)	(0.1675)	(0.4098)	(0.2482)	(0.1974)	(0.1872)	(0.2333)

Notes: Data are from LFS 2005. Robust standard errors are in parentheses. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table A4: OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N = 890)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Age 15-19	-0.2246 ***	-0.6329 ***	-0.4941 ***	-0.1235	-0.0455	0.2183 *	-0.0508	0.3542	-0.0412	-0.0528	-0.7697	-0.2887
	(0.0668)	(0.1113)	(0.0924)	(0.1079)	(0.1109)	(0.1245)	(0.2631)	(0.6369)	(0.4192)	(0.3938)	(0.6156)	(0.1938)
Age 20-24	-0.0155	-0.1409	-0.1087	0.0774	-0.0632	0.1166	0.1580	0.5083	0.1082	0.0809	-0.5206	-0.0974
	(0.0614)	(0.0963)	(0.0835)	(0.1010)	(0.1048)	(0.1184)	(0.2585)	(0.6346)	(0.4128)	(0.3895)	(0.6011)	(0.1715)
Age 25-29	0.0138	-0.1028	-0.0456	0.1640	0.0241	0.0230	0.0992	0.4642	0.0674	0.0710	-0.5152	-0.0294
	(0.0559)	(0.0866)	(0.0749)	(0.0911)	(0.0964)	(0.1121)	(0.2591)	(0.6333)	(0.4107)	(0.3882)	(0.5992)	(0.1605)
Age 30-34	0.1013 *	-0.0123	0.0225	0.2968 ***	0.0828	0.0471	0.3410	0.6014	0.2913	0.3358	-0.1772	0.0379
	(0.0544)	(0.0823)	(0.0715)	(0.0878)	(0.0945)	(0.1115)	(0.2608)	(0.6346)	(0.4107)	(0.3865)	(0.5987)	(0.1554)
Age 35-39	0.1232 **	0.0068	0.0053	0.2593 ***	0.1853 *	0.1490	0.2294	0.5144	0.1751	0.1387	-0.3241	0.2867 *
	(0.0537)	(0.0812)	(0.0708)	(0.0861)	(0.0936)	(0.1128)	(0.2625)	(0.6358)	(0.4097)	(0.3850)	(0.5953)	(0.1641)
Age 40-44	0.1828 ***	-0.0020	0.0286	0.3248 ***	0.2540 ***	0.2782 **	0.2732	0.5455	0.1539	0.2226	-0.1274	0.2063
	(0.0542)	(0.0816)	(0.0713)	(0.0871)	(0.0946)	(0.1159)	(0.2630)	(0.6394)	(0.4146)	(0.3917)	(0.5985)	(0.1382)
Age 45-49	0.2262 ***	0.0299	0.0376	0.2991 ***	0.3428 ***	0.4605 ***	0.2872	0.5554	0.2754	0.1449	-0.1535	0.5579 ***
	(0.0555)	(0.0816)	(0.0719)	(0.0896)	(0.0987)	(0.1272)	(0.2646)	(0.6402)	(0.4121)	(0.3896)	(0.6017)	(0.2158)
Age 50-54	0.2128 ***	-0.0518	-0.0180	0.2889 ***	0.3037 ***	0.6123 ***	0.2719	0.5741	0.0696	0.1457	-0.0566	0.1722
	(0.0572)	(0.0868)	(0.0753)	(0.0917)	(0.0994)	(0.1321)	(0.2686)	(0.6538)	(0.4347)	(0.3998)	(0.6156)	(0.1909)
Age 55-59	0.2434 ***	0.0909	0.0783	0.2529 **	0.3064 ***	0.5018 ***	0.4277	0.9300	0.2350	0.2219	0.3092	0.3664
	(0.0625)	(0.0867)	(0.0785)	(0.1016)	(0.1127)	(0.1605)	(0.3076)	(0.6428)	(0.4805)	(0.4468)	(0.7190)	(0.3616)
Primary school	0.0613 ***	0.1671 ***	0.1307 ***	0.1822 ***	-0.0163	-0.2246 ***	0.0842	0.0116	0.1814 **	0.2618 ***	0.1223	-0.1910 *
	(0.0229)	(0.0356)	(0.0319)	(0.0398)	(0.0367)	(0.0378)	(0.0600)	(0.0822)	(0.0822)	(0.0936)	(0.1327)	(0.0976)
Secondary school	0.1118 ***	0.1755 ***	0.1978 ***	0.3360 ***	0.1870 ***	-0.3955 ***	0.1565 *	0.0135	0.2040 **	0.3812 ***	0.2782	-0.3128 *
	(0.0261)	(0.0369)	(0.0327)	(0.0454)	(0.0513)	(0.0558)	(0.0925)	(0.0899)	(0.0857)	(0.1141)	(0.1730)	(0.1738)
Post-secondary school	0.5826 ***	0.1789 ***	0.1997 ***	0.5436 ***	0.7669 ***	0.1385	0.3507 ***	-0.1072	0.0646	0.3082 ***	0.1819	0.0396
	(0.0310)	(0.0306)	(0.0278)	(0.0431)	(0.0642)	(0.1048)	(0.0953)	(0.1119)	(0.0886)	(0.1160)	(0.2153)	(0.3177)
Graduate	0.9028 ***	0.1564 ***	0.1915 ***	0.6156 ***	0.9831 ***	1.0350 ***	0.7277 ***	-0.0933	0.1459 *	0.4307 ***	0.5767 **	0.2818
	(0.0393)	(0.0388)	(0.0323)	(0.0479)	(0.0805)	(0.1696)	(0.1168)	(0.0819)	(0.0764)	(0.1101)	(0.2592)	(0.4875)
General	0.0359	-0.0948	0.0197	0.1297 **	0.1227	0.2469	0.3664 ***	0.1294 *	0.1952 **	0.2247 *	0.4068 *	0.6830
	(0.0482)	(0.0630)	(0.0411)	(0.0650)	(0.0992)	(0.1745)	(0.1286)	(0.0719)	(0.0778)	(0.1330)	(0.2296)	(0.4882)

Continued on next page

Table A4 (continued): OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N = 890)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Vocational	0.2252 ***	0.0630 **	0.0592 *	0.1982 ***	0.3575 ***	0.5635 ***	0.2648 *	-0.0441	0.1750 *	0.2082	0.7845 **	-0.0182
	(0.0420)	(0.0301)	(0.0321)	(0.0554)	(0.0950)	(0.1407)	(0.1440)	(0.1615)	(0.0996)	(0.1923)	(0.3838)	(0.5050)
Married	0.1184 ***	-0.0842 *	0.0147	0.1645 ***	0.1389 **	0.1036	-0.0957	-0.0234	-0.1000	-0.1548 *	-0.1845	-0.2462
	(0.0304)	(0.0478)	(0.0410)	(0.0514)	(0.0559)	(0.0645)	(0.0672)	(0.0792)	(0.0747)	(0.0871)	(0.1605)	(0.1718)
Widowed	0.0716	-0.8852	-0.6627 **	-0.2222	0.2703	0.3134	-0.0494	-0.3864 **	-0.2432 *	-0.3211 **	-0.4797 ***	-0.2674
	(0.1387)	(0.5439)	(0.2861)	(0.1901)	(0.3393)	(0.6677)	(0.0951)	(0.1917)	(0.1448)	(0.1354)	(0.1769)	(0.2132)
Divorced	-0.1228	0.0218	-0.0139	0.0488	-0.0077	-0.0248	-0.2965 ***	0.0357	0.0225	0.0389	-0.3269	-0.3382
	(0.2331)	(0.1907)	(0.2217)	(0.1797)	(0.1357)	(0.3119)	(0.1122)	(0.1434)	(0.1282)	(0.1325)	(0.2114)	(0.2156)
Urban	0.1786 ***	0.0886 ***	0.1362 ***	0.3162 ***	0.2735 ***	0.3471 ***	0.0892	0.0972	0.0541	0.1808 **	0.2169 *	0.2823 **
	(0.0203)	(0.0256)	(0.0265)	(0.0365)	(0.0360)	(0.0470)	(0.0702)	(0.0897)	(0.0835)	(0.0895)	(0.1117)	(0.1190)
Professional	0.1193 ***	-0.1038 ***	-0.0934 ***	0.2006 ***	0.6155 ***	0.6025 ***	0.7203 ***	0.0292	-0.0208	0.3701 **	1.6033 ***	2.8539 ***
	(0.0345)	(0.0360)	(0.0343)	(0.0575)	(0.0780)	(0.1089)	(0.1415)	(0.0765)	(0.1347)	(0.1833)	(0.3336)	(0.4590)
Administrative	0.3055 ***	-0.0175	0.0048	0.3848 ***	0.8157 ***	1.5945 ***	0.5945 ***	0.0439	0.1265	0.4916 **	1.2716 ***	2.5465 ***
	(0.0472)	(0.0366)	(0.0363)	(0.0700)	(0.1026)	(0.1743)	(0.1591)	(0.1085)	(0.1611)	(0.1989)	(0.3598)	(0.4895)
Clerical	0.0184	-0.1157 *	-0.1424 **	0.1044	0.0830	0.2927 **	-0.0046	-0.4405	-0.3200	-0.1428	0.6646	0.3914
	(0.0493)	(0.0641)	(0.0589)	(0.0902)	(0.1062)	(0.1395)	(0.2171)	(0.3695)	(0.2642)	(0.3308)	(0.5465)	(0.7707)
Sales	-0.1196 ***	-0.2716 ***	-0.2509 ***	-0.0171	0.1222	-0.0955	-0.3851 ***	-0.6076 ***	-0.5820 ***	-0.4854 ***	-0.4356 *	0.0427
	(0.0385)	(0.0558)	(0.0472)	(0.0715)	(0.0830)	(0.0889)	(0.1044)	(0.0949)	(0.1345)	(0.1668)	(0.2426)	(0.1075)
Agriculture	-0.1219 ***	-0.2656 ***	-0.2746 ***	-0.1448	0.0732	0.2175 *	-0.2017	-0.0297	-0.1322	-0.4487 **	-0.0785	0.3297 *
	(0.0455)	(0.0707)	(0.0571)	(0.0824)	(0.0972)	(0.1121)	(0.1398)	(0.1504)	(0.1809)	(0.2161)	(0.3128)	(0.1948)
Production	-0.0017	-0.1144 ***	-0.0832 **	0.0726	0.1336 **	0.1189 **	0.0038	-0.0848	-0.1018	0.1628	-0.2448	0.1459 *
	(0.0292)	(0.0339)	(0.0327)	(0.0568)	(0.0571)	(0.0550)	(0.0991)	(0.0558)	(0.1079)	(0.1631)	(0.2490)	(0.0765)
Agriculture	-0.1803 ***	0.0172	-0.2335 ***	-0.4047 ***	-0.4173 ***	-0.1092	-0.0505	0.1487	-0.1379	-0.0826	-0.4437 **	-0.2076
	(0.0474)	(0.0761)	(0.0602)	(0.0862)	(0.1040)	(0.1218)	(0.1229)	(0.1861)	(0.1913)	(0.1776)	(0.2047)	(0.1979)
Manufacturing	-0.0181	0.0398	0.0029	-0.0527	-0.2011 ***	0.0314	0.0219	0.2591 ***	0.2468 **	-0.0069	-0.0701	-0.0285
	(0.0312)	(0.0400)	(0.0348)	(0.0577)	(0.0673)	(0.0845)	(0.0722)	(0.0872)	(0.1058)	(0.1031)	(0.1410)	(0.1455)

Continued on next page

Table A4 (continued): OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 1999

	LFS 1999											
	Males (N = 4,632)						Females (N = 890)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Wholesale and Retail	-0.1296 ***	-0.0337	-0.0384	-0.0971	-0.2505 **	-0.0570	-0.2200	0.2426	0.2669	-0.2639	-0.7120	-0.4050
	(0.0470)	(0.0679)	(0.0566)	(0.0871)	(0.1045)	(0.1366)	(0.2060)	(0.2864)	(0.2166)	(0.3574)	(0.4864)	(0.6774)
Transport	0.0743 *	0.0421	-0.0155	0.1466 **	-0.0142	0.1223	0.0445	0.2567 *	0.1421	-0.1447	0.2651	-0.7502
	(0.0385)	(0.0429)	(0.0415)	(0.0714)	(0.0885)	(0.1051)	(0.2140)	(0.1316)	(0.2649)	(0.3829)	(0.8146)	(0.6854)
Financial institution	0.3630 ***	0.1106 **	0.1111 ***	0.3227 ***	0.6078 ***	1.8913 ***	0.4489 **	0.2898 **	0.3716 ***	0.4841 ***	0.5692	1.7327 ***
	(0.0580)	(0.0439)	(0.0388)	(0.0650)	(0.1105)	(0.2624)	(0.2284)	(0.1415)	(0.1153)	(0.1448)	(0.3800)	(0.5342)
Real estate	0.1577	0.1218 **	0.0795	0.3208 **	0.4993 **	0.0975	0.3940	0.4187 ***	0.5741 ***	0.8576 ***	0.5532	0.4497
	(0.1122)	(0.0478)	(0.0927)	(0.1402)	(0.2163)	(0.4243)	(0.2673)	(0.1385)	(0.1501)	(0.2381)	(0.3929)	(1.0296)
Public adminisitration	0.3234 ***	0.0577	0.0937 ***	0.3700 ***	0.4996 ***	1.2142 ***	0.4255 ***	0.2877 ***	0.3961 ***	0.5102 ***	1.2666 ***	0.1547
	(0.0376)	(0.0385)	(0.0316)	(0.0523)	(0.0853)	(0.1446)	(0.1289)	(0.0894)	(0.0982)	(0.1233)	(0.2524)	(0.5588)
Education	-0.0626	0.0640	0.0612	0.1255	-0.0234	-0.3852 *	-0.1706	0.1892	0.0880	-0.0502	0.1121	-0.8066
	(0.0504)	(0.0423)	(0.0414)	(0.0784)	(0.1235)	(0.1953)	(0.1292)	(0.1153)	(0.1481)	(0.1568)	(0.3286)	(0.5088)
Health	0.1784 **	0.0073	0.0049	0.1573	0.4823 ***	0.6936 **	0.3218 **	0.1549	0.2198	0.2894	0.9876 ***	-0.3551
	(0.0716)	(0.0768)	(0.0625)	(0.1040)	(0.1582)	(0.2981)	(0.1391)	(0.2009)	(0.1596)	(0.1821)	(0.3512)	(0.4626)
inverse Mill's ratio	0.5649 ***	-0.3820	-0.1416	1.2646 ***	1.4123 ***	0.5565 *	-0.3042	-0.7297	-0.3933	-0.2373	0.6666	0.4013
	(0.1617)	(0.2593)	(0.2243)	(0.2727)	(0.2593)	(0.2966)	(0.3732)	(0.6316)	(0.5973)	(0.6030)	(0.8095)	(0.6726)
Constant	1.5908 ***	1.6521 ***	1.7899 ***	0.7548 ***	1.3355 ***	2.0379 ***	1.4369 ***	0.6914	1.2062 **	1.4675 ***	2.0388 **	2.2380 ***
	(0.0897)	(0.1924)	(0.1713)	(0.2087)	(0.2022)	(0.2241)	(0.3287)	(0.7292)	(0.5716)	(0.5627)	(0.8240)	(0.5018)

Notes: Data are from LFS 1999. Robust standard errors are in parentheses. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table A5: OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 2005

	LFS 2005											
	Males (N = 16,106)						Females (N= 2,286)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Age 15-19	-0.0278	-0.3385 ***	-0.3979	0.1038 **	0.0127	-0.0373	-0.2413	-0.0659	-0.3979	-0.3358	-0.3555 *	-0.2054
	(0.0477)	(0.1029)	(0.2592)	(0.0491)	(0.0426)	(0.0514)	(0.1748)	(0.4101)	(0.2592)	(0.2063)	(0.1975)	(0.2510)
Age 20-24	0.0150	-0.1371	-0.2921	0.0962 **	0.0289	0.0205	-0.2697 *	0.1486	-0.2921	-0.4216 **	-0.3759 **	-0.2478
	(0.0435)	(0.0858)	(0.2374)	(0.0447)	(0.0391)	(0.0470)	(0.1613)	(0.3823)	(0.2374)	(0.1937)	(0.1860)	(0.2323)
Age 25-29	0.0493	-0.0638	-0.1307	0.0761	0.0497	0.0576	-0.1318	0.1686	-0.1307	-0.1634	-0.2900	-0.2926
	(0.0404)	(0.0759)	(0.2273)	(0.0414)	(0.0363)	(0.0437)	(0.1574)	(0.3747)	(0.2273)	(0.1852)	(0.1803)	(0.2258)
Age 30-34	0.0956 **	0.0918	-0.0284	0.1080 ***	0.0639 *	0.0884 **	0.0071	0.3049	-0.0284	-0.0572	-0.2040	-0.1566
	(0.0398)	(0.0722)	(0.2242)	(0.0406)	(0.0359)	(0.0434)	(0.1571)	(0.3732)	(0.2242)	(0.1829)	(0.1785)	(0.2230)
Age 35-39	0.1107 ***	0.1078	0.0305	0.1111 ***	0.0925 ***	0.0909 **	0.0637	0.2689	0.0305	-0.0310	0.0187	0.0018
	(0.0392)	(0.0704)	(0.2231)	(0.0401)	(0.0354)	(0.0426)	(0.1561)	(0.3719)	(0.2231)	(0.1814)	(0.1781)	(0.2220)
Age 40-44	0.1241 ***	0.0952	0.0615	0.1270 ***	0.0888 **	0.1238 ***	0.2021	0.2468	0.0615	0.2560	0.1703	0.1855
	(0.0396)	(0.0708)	(0.2233)	(0.0405)	(0.0358)	(0.0434)	(0.1570)	(0.3745)	(0.2233)	(0.1817)	(0.1797)	(0.2278)
Age 45-49	0.1290 ***	0.0710	0.1670	0.1222 ***	0.0871 **	0.0887 **	0.2187	0.3350	0.1670	0.2861	0.2393	0.1985
	(0.0399)	(0.0713)	(0.2233)	(0.0409)	(0.0360)	(0.0432)	(0.1596)	(0.3737)	(0.2233)	(0.1842)	(0.1832)	(0.2331)
Age 50-54	0.1130 ***	0.0292	0.0180	0.1239 ***	0.1077 ***	0.0834 *	0.1200	0.3336	0.0180	0.1202	0.0567	-0.0305
	(0.0416)	(0.0752)	(0.2357)	(0.0428)	(0.0375)	(0.0449)	(0.1654)	(0.3834)	(0.2357)	(0.1939)	(0.1899)	(0.2319)
Age 55-59	0.1380 ***	0.1079	0.1696	0.1171 **	0.0741 *	0.0811 *	0.1874	0.2052	0.1696	0.3456	0.0043	0.1913
	(0.0444)	(0.0787)	(0.2474)	(0.0457)	(0.0397)	(0.0481)	(0.1759)	(0.4190)	(0.2474)	(0.2083)	(0.2012)	(0.2651)
Primary school	-0.0206	-0.0293	-0.1053	-0.0471 ***	0.0653 ***	0.0902 ***	-0.0737	0.1874 *	-0.1053	-0.2512 ***	-0.1594 ***	0.0721
	(0.0171)	(0.0363)	(0.0925)	(0.0179)	(0.0169)	(0.0215)	(0.0532)	(0.1093)	(0.0925)	(0.0687)	(0.0573)	(0.0791)
Secondary school	-0.1895 ***	-0.1809 ***	0.0035	-0.2377 ***	-0.0560 ***	-0.0402 *	-0.0131	0.1965	0.0035	-0.1400	-0.0640	0.1317
	(0.0192)	(0.0454)	(0.1107)	(0.0200)	(0.0182)	(0.0226)	(0.0622)	(0.1272)	(0.1107)	(0.0838)	(0.0703)	(0.0932)
Post-secondary school	-0.1753 ***	0.1921 ***	0.5225 ***	-0.3574 ***	-0.1639 ***	-0.1363 ***	0.3740 ***	0.3845 **	0.5225 ***	0.3814 ***	0.1472 *	0.1773
	(0.0235)	(0.0602)	(0.1388)	(0.0229)	(0.0191)	(0.0248)	(0.0846)	(0.1583)	(0.1388)	(0.1167)	(0.0872)	(0.1238)
Graduate	0.3078 ***	0.5554 ***	0.8670 ***	0.1829 ***	0.1834 ***	0.2020 ***	0.7984 ***	0.4478 **	0.8670 ***	0.9116 ***	0.6055 ***	0.5792 ***
	(0.0358)	(0.0816)	(0.1676)	(0.0376)	(0.0314)	(0.0399)	(0.1066)	(0.1968)	(0.1676)	(0.1502)	(0.1231)	(0.1652)
General	-0.1365 ***	-0.1385 *	-0.0053	-0.1444 ***	-0.0500 **	-0.0351	-0.1194 *	0.0820	-0.0053	-0.1670 **	-0.0891	-0.2511 ***
	(0.0271)	(0.0720)	(0.1087)	(0.0289)	(0.0251)	(0.0341)	(0.0631)	(0.1015)	(0.1087)	(0.0761)	(0.0661)	(0.0788)

Continued on next page

Table A5 (continued): OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 2005

	LFS 2005											
	Males (N = 16,106)						Females (N= 2,286)					
Variables	OLS	10th	25th	50th	75th	90th	OLS	10th	25th	50th	75th	90th
Vocational	-0.0580	-0.1021	-0.7545 ***	-0.1480 ***	-0.0338	0.0277	-0.5206 ***	-0.6675	-0.7545 ***	-0.5269 ***	-0.4035 ***	-0.4691 ***
	(0.0517)	(0.1256)	(0.2458)	(0.0535)	(0.0478)	(0.0667)	(0.1550)	(0.4202)	(0.2458)	(0.1880)	(0.1105)	(0.0756)
Married	0.1927 ***	0.3901 ***	0.0841	0.1451 ***	0.0777 ***	0.0435 *	0.0167	0.0093	0.0841	0.0686	-0.0220	0.0089
	(0.0224)	(0.0550)	(0.0994)	(0.0227)	(0.0194)	(0.0246)	(0.0619)	(0.1248)	(0.0994)	(0.0804)	(0.0643)	(0.0860)
Widowed	-0.0140	0.0978	0.0190	-0.1041	-0.2284 **	-0.1174	0.0467	-0.3686	0.0190	0.1099	-0.0104	-0.1254
	(0.0815)	(0.3410)	(0.1593)	(0.1605)	(0.1021)	(0.1391)	(0.0819)	(0.2356)	(0.1593)	(0.1276)	(0.1043)	(0.1323)
Divorced	-0.0041	-0.0823	0.0899	-0.0337	0.0367	0.1332	0.0361	0.2075	0.0899	0.0242	-0.0772	-0.0497
	(0.1476)	(0.2058)	(0.1305)	(0.0852)	(0.0686)	(0.0952)	(0.0937)	(0.1679)	(0.1305)	(0.1119)	(0.0937)	(0.1203)
Urban	0.0195	-0.0429	-0.0798	0.0361 **	0.0627 ***	0.0661 ***	-0.0172	-0.0717	-0.0798	-0.1354 **	0.0368	0.1154 **
	(0.0138)	(0.0324)	(0.0672)	(0.0140)	(0.0124)	(0.0156)	(0.0457)	(0.0874)	(0.0672)	(0.0580)	(0.0460)	(0.0567)
Professional	0.0968 ***	0.2238 **	0.2086	0.0037	-0.0178	-0.0307	0.3964 ***	0.4615 *	0.2086	0.5534 ***	0.4559 ***	0.3065 *
	(0.0358)	(0.0901)	(0.1988)	(0.0343)	(0.0258)	(0.0346)	(0.1091)	(0.2466)	(0.1988)	(0.1604)	(0.1130)	(0.1599)
Administrative	-0.0110	0.2016	0.5534 *	-0.0732	-0.0867	-0.1681 **	0.7334 **	0.6428 **	0.5534 *	0.6226	0.8259 **	1.1319
	(0.0804)	(0.1807)	(0.2820)	(0.0787)	(0.0626)	(0.0723)	(0.3705)	(0.2994)	(0.2820)	(0.3484)	(0.3839)	(0.7657)
Clerical	0.2020 ***	0.1619 *	0.1836	0.2201 ***	0.0648 **	0.0296	0.2638 **	0.4933 **	0.1836	0.3473 **	0.4186 ***	-0.0573
	(0.0349)	(0.0837)	(0.1847)	(0.0358)	(0.0262)	(0.0340)	(0.1157)	(0.2322)	(0.1847)	(0.1662)	(0.1226)	(0.1569)
Sales	-0.1960 ***	-0.2973 ***	0.2056	-0.1984 ***	-0.1345 ***	-0.1550 ***	0.0884	0.4077	0.2056	0.0103	0.1377	0.0012
	(0.0347)	(0.1029)	(0.2451)	(0.0313)	(0.0235)	(0.0308)	(0.1387)	(0.3054)	(0.2451)	(0.1947)	(0.1416)	(0.1782)
Agriculture	0.3363 ***	0.7746 ***	0.9765 ***	0.2060 ***	0.0608	-0.0735	0.5018 ***	0.8692 ***	0.9765 ***	1.2294 ***	-0.1331	-0.1960
	(0.0440)	(0.1142)	(0.1782)	(0.0464)	(0.0389)	(0.0478)	(0.1179)	(0.2559)	(0.1782)	(0.1594)	(0.1306)	(0.1597)
Production	0.2780 ***	0.3406 ***	0.1944	0.2618 ***	0.1812 ***	0.1440 ***	0.2432 ***	0.3921 *	0.1944	0.2945 **	0.2850 ***	0.1632
	(0.0265)	(0.0748)	(0.1617)	(0.0276)	(0.0223)	(0.0295)	(0.0912)	(0.2124)	(0.1617)	(0.1273)	(0.0985)	(0.1383)
Agriculture	0.2525 ***	0.4610 ***	0.8371 ***	0.2646 ***	-0.0730 *	-0.1076 **	0.7679 ***	1.0969 ***	0.8371 ***	0.3690 **	0.6480 ***	0.4133 **
	(0.0469)	(0.1275)	(0.1913)	(0.0498)	(0.0420)	(0.0516)	(0.1243)	(0.2844)	(0.1913)	(0.1669)	(0.1417)	(0.1822)
Manufacturing	0.0877 ***	0.3900 ***	0.7730 ***	0.0267	-0.0247	-0.0128	0.3951 ***	1.2832 ***	0.7730 ***	0.0012	-0.0446	-0.1087
	(0.0311)	(0.0968)	(0.1743)	(0.0327)	(0.0274)	(0.0353)	(0.0961)	(0.2385)	(0.1743)	(0.1340)	(0.1040)	(0.1441)

Continued on next page

Table A5 (continued): OLS and Unconditional Quantile Regression Estimates with Selectivity Bias Correction, by Gender, LFS 2005

	LFS 2005											
	Males (N = 16,106)						Females (N= 2,286)					
Variables	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	OLS	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Wholesale and Retail	-0.1408***	-0.4484***	1.2620***	-0.0231	-0.0774**	-0.0220	0.8779***	0.9773**	1.2620***	1.0719***	0.4692*	-0.0835
	(0.0393)	(0.1278)	(0.2951)	(0.0391)	(0.0320)	(0.0418)	(0.2023)	(0.4453)	(0.2951)	(0.2633)	(0.2752)	(0.2813)
Transport	0.1672***	0.5855***	1.4959***	0.0043	-0.0448	0.0224	0.8572***	1.6523***	1.4959***	0.7223***	0.3747*	0.0465
	(0.0335)	(0.0977)	(0.1748)	(0.0364)	(0.0312)	(0.0421)	(0.1615)	(0.2002)	(0.1748)	(0.2476)	(0.2231)	(0.3182)
Financial institution	0.1701***	0.5147***	0.6716***	0.0596	0.0257	0.0602	0.2856**	0.7552**	0.6716***	-0.0550	-0.1772	0.0216
	(0.0474)	(0.1138)	(0.2057)	(0.0487)	(0.0415)	(0.0556)	(0.1205)	(0.2938)	(0.2057)	(0.1661)	(0.1302)	(0.1984)
Real estate	0.0493	0.2175	0.5974	-0.0387	-0.0715	-0.0061	0.4640*	1.2310***	0.5974	-0.1586	0.0395	0.0631
	(0.0746)	(0.1981)	(0.3861)	(0.0760)	(0.0653)	(0.0904)	(0.2500)	(0.2500)	(0.3861)	(0.3119)	(0.2924)	(0.4744)
Public adminisitration	0.0757**	0.5767***	0.8046***	-0.1033***	-0.2298***	-0.2002***	0.4482***	1.1860***	0.8046***	0.4322**	-0.2715**	-0.2625
	(0.0359)	(0.1018)	(0.1918)	(0.0354)	(0.0270)	(0.0349)	(0.1183)	(0.2426)	(0.1918)	(0.1723)	(0.1255)	(0.1631)
Education	-0.2320***	0.1836	0.7607***	-0.3371***	-0.3098***	-0.2610***	0.0561	1.0241***	0.7607***	-0.2705	-0.6772***	-0.8014***
	(0.0421)	(0.1177)	(0.1988)	(0.0406)	(0.0321)	(0.0428)	(0.1116)	(0.2608)	(0.1988)	(0.1572)	(0.1215)	(0.1793)
Health	-0.0213	0.2873**	0.7692***	-0.1295**	-0.1226***	-0.0885	0.2421**	0.9161***	0.7692***	0.1052	-0.3067***	-0.3243**
	(0.0529)	(0.1423)	(0.2065)	(0.0506)	(0.0419)	(0.0559)	(0.1145)	(0.2808)	(0.2065)	(0.1625)	(0.1159)	(0.1626)
inverse Mill's ratio	0.6445***	0.1888	0.3922	0.7347***	0.6969***	0.8148***	0.3849*	-0.9266	0.3922	-0.1182	0.2077	-0.6188
	(0.1379)	(0.3140)	(0.6196)	(0.1429)	(0.1279)	(0.1600)	(0.2177)	(0.7463)	(0.6196)	(0.5302)	(0.4587)	(0.5804)
Constant	2.7745***	0.6989***	0.3875	2.3769***	2.9590***	3.2669***	2.0218***	-0.1550	0.3875	2.0575***	2.8376***	3.7066***
	(0.0746)	(0.2441)	(0.4870)	(0.1092)	(0.0966)	(0.1210)	(0.1986)	(0.6490)	(0.4870)	(0.4107)	(0.3605)	(0.4436)

Notes: Data are from LFS 2005. Robust standard errors are in parentheses. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table A6: Detailed Decomposition of the Gender Wage Gap at the Mean and at Different Quantiles

	LFS 1999						LFS 2005					
	Mean	10th	25th	50th	75th	90th	Mean	10th	25th	50th	75th	90th
Panel A: No Selection												
Differences in observed wages	0.4542	0.5026	0.5423	0.4688	0.5551	0.2261	0.6488	0.7303	0.8475	0.6530	0.4747	0.4049
Contribution of Characteristics												
Age	0.0425	0.0350	0.0295	0.0427	0.0593	0.0493	0.0070	0.0098	0.0172	0.0035	0.0050	0.0049
Education	0.0295	0.0232	0.0263	0.0613	0.0557	-0.0164	-0.0144	-0.0414	-0.0494	-0.0069	0.0048	0.0069
Training	0.0042	0.0011	0.0009	0.0032	0.0058	0.0090	0.0046	0.0045	0.0090	0.0047	0.0017	0.0013
Marital Status	0.0366	0.0318	0.0431	0.0496	0.0248	0.0087	0.0275	0.0574	0.0068	0.0296	0.0179	-0.0065
Urban	-0.0421	-0.0203	-0.0282	-0.0553	-0.0457	-0.0660	0.0001	0.0080	-0.0005	-0.0023	-0.0068	-0.0068
Occupation	-0.0065	0.0113	-0.0009	-0.0404	-0.0625	0.0031	0.0286	0.0891	0.1318	0.0160	-0.0015	-0.0250
Industry	-0.0323	0.0010	-0.0513	-0.0684	-0.0340	0.0647	0.0699	0.0416	0.1151	0.0878	0.0148	0.0079
Total	0.0318	0.0831	0.0194	-0.0073	0.0033	0.0523	0.1231	0.1691	0.2301	0.1324	0.0359	-0.0173
	(7%)	(17%)	(4%)	(-2%)	(0.01%)	(23%)	(19%)	(22%)	(27%)	(20%)	(8%)	(-4%)
Discrimination	0.4224	0.4195	0.5229	0.4761	0.5518	0.1738	0.5257	0.5612	0.6174	0.5206	0.4388	0.4223
	(93%)	(83%)	(96%)	(102%)	(99%)	(77%)	(81%)	(77%)	(73%)	(80%)	(92%)	(104%)
Panel B: With Selection												
Differences in observed wages	0.4542	0.5026	0.5423	0.4688	0.5551	0.2261	0.6488	0.7303	0.8475	0.6530	0.4747	0.4049
Differences in offered wages	0.0068	0.3416	0.4216	-0.3155	0.1418	0.1321	0.5047	0.1310	1.0571	0.1751	0.1952	-0.3801
Contribution of Characteristics												
Age	0.0336	0.0388	0.0309	0.0303	0.0454	0.0438	0.0045	0.0090	0.0172	0.0006	0.0022	0.0017
Education	0.0540	0.0254	0.0271	0.0542	0.0477	-0.0196	-0.0218	-0.0435	-0.0494	-0.0153	-0.0032	-0.0025
Training	0.0036	0.0011	0.0009	0.0031	0.0057	0.0090	0.0045	0.0045	0.0090	0.0046	0.0016	0.0012
Marital Status	0.0181	0.0361	0.0447	0.0352	0.0087	0.0023	0.0276	0.0575	0.0068	0.0298	0.0180	-0.0063
Urban	-0.0463	-0.0178	-0.0273	-0.0634	-0.0549	-0.0696	-0.0031	0.0071	-0.0005	-0.0060	-0.0103	-0.0109
Occupation	-0.0095	0.0111	-0.0010	-0.0398	-0.0333	0.0034	0.0309	0.0898	0.1318	0.0187	0.0011	-0.0220
Industry	-0.0400	0.0008	-0.0513	-0.0677	-0.0618	0.0650	0.0700	0.0416	0.1151	0.0879	0.0150	0.0081
Total	0.0135	0.0955	0.0239	-0.0482	-0.0424	0.0343	0.1126	0.1660	0.2301	0.1203	0.0244	-0.0308
	(199%)	(28%)	(6%)	(15%)	(-30%)	(26%)	(22%)	(127%)	(22%)	(70%)	(13%)	(8%)
Discrimination	-0.0067	0.2461	0.3977	-0.2673	0.1842	0.0978	0.3922	-0.0338	0.8270	0.0548	0.1708	-0.3493
	(-99%)	(72%)	(94%)	(85%)	(130%)	(74%)	(78%)	(-27%)	(78%)	(30%)	(88%)	(92%)

Notes: Data are from LFS 1999 and LFS 2005. Numbers in parentheses in panel A indicate the percentage of each component's contribution to the overall wage gap while numbers in parentheses in panel B indicate the percentage of each component's contribution to the offered wage gap. The following explanatory variables are included in each group: Age: Age 15-19, Age 20-24, Age 25-29, Age 30-34, Age 35-39, Age 40-44, Age 45-49, Age 50-54, Age 55-59. Education: Primary school, secondary school, post-secondary school, and graduate. Training: Vocational and general. Marital status: Married, divorced and widowed. Occupation: Professional, administrative, clerical, sales, agriculture, and production. Industry: Agriculture, manufacturing, health, public administration, transport, financial institutional, real estate, wholesale and retail, and education.

Table A7: Detailed Decomposition of Change in Gender Wage Gap at the Mean and at Different Quantiles

	Mean	10th	25th	50th	75th	90th
Panel A: No Selection						
Changes in observed wage gap	0.1946	0.2277	0.3052	0.1852	-0.0804	0.1788
Contribution of characteristics						
Age	-0.0305	-0.0133	-0.0268	-0.0368	-0.0318	-0.0262
Education	-0.0906	-0.0610	-0.1094	-0.1252	-0.0684	-0.0784
Training	0.0010	-0.0105	-0.0129	0.0055	0.0016	0.0098
Marital Status	-0.0053	-0.0124	-0.0091	-0.0044	-0.0010	-0.0011
Urban	-0.0070	-0.0007	-0.0129	-0.0192	-0.0001	0.0143
Occupation	0.0706	0.0871	0.1046	0.0217	0.0628	0.1279
Industry	-0.0356	-0.1613	-0.1426	0.0056	0.0783	0.0810
Total	-0.0973	-0.1720	-0.2090	-0.1529	0.0413	0.1273
	(-50%)	(-75%)	(-68%)	(-83%)	(-51%)	(71%)
Discrimination	0.2919	0.3997	0.5142	0.3371	-0.1217	0.0515
	(150%)	(175%)	(168%)	(183%)	(151%)	(29%)
Panel B: With Selection						
Changes in observed wage gap	0.1946	0.2277	0.3052	0.1852	-0.0804	0.1788
Changes in offered wage gap	0.4978	-0.2096	0.6355	0.4955	0.0535	-0.5121
Contribution of characteristics						
Age	-0.0286	-0.0162	-0.0254	-0.0366	-0.0305	-0.0276
Education	-0.1010	-0.0460	-0.1152	-0.1188	-0.0671	-0.0639
Training	0.0009	-0.0109	-0.0128	0.0055	0.0017	0.0096
Marital Status	-0.0051	-0.0123	-0.0091	-0.0046	-0.0013	-0.0012
Urban	-0.0030	-0.0054	-0.0111	-0.0219	-0.0013	0.0091
Occupation	0.0674	0.0895	0.1035	0.0210	0.0613	0.1286
Industry	-0.0357	-0.1622	-0.1422	0.0057	0.0787	0.0806
Total	-0.1051	-0.1634	-0.2124	-0.1498	0.0414	0.1352
	(-21%)	(78%)	(-33%)	(-30%)	(77%)	(-26%)
Discrimination	0.6029	-0.0462	0.8479	0.6446	0.0121	-0.6473
	(121%)	(22%)	(133%)	(130%)	(23%)	(126%)

Notes: Data are from LFS 1999 and LFS 2005. Numbers in parentheses in panel A indicate the percentage of each component's contribution to the overall wage gap while numbers in parentheses in panel B indicate the percentage of each component's contribution to the offered wage gap. The following explanatory variables are included in each group: Age: Age 15-19, Age 20-24, Age 25-29, Age 30-34, Age 35-39, Age 40-44, Age 45-49, Age 50-54, Age 55-59. Education: Primary school, secondary school, post-secondary school, and graduate. Training: Vocational and general. Marital status: Married, divorced and widowed. Occupation: Professional, administrative, clerical, sales, agriculture, and production. Industry: Agriculture, manufacturing, health, public administration, transport, financial institutional, real estate, wholesale and retail, and education.

Appendix B

Table B1: Descriptive Statistics, by Year and Gender

	1999				2004			
	Males		Females		Males		Females	
Variables	Mean/Median	Std. Dev.	Mean/Median	Std. Dev.	Mean/Median	Std. Dev.	Mean/Median	Std. Dev.
Fertility preference in district	0.0741	0.0354	0.2424	0.0495	0.0189	0.0116	0.2584	0.0205
Fertility preference in district for age group 18-21	0.1819	0.2740	0.6199	0.1867	0.0345	0.1016	0.6393	0.1038
Fertility preference in district for those over 21 years	0.0719	0.0348	0.1587	0.0521	0.0177	0.0115	0.1620	0.0220
Median age at first marriage in district	22.8276	1.8476	14.5431	1.1221	22.8186	2.0167	14.2832	0.9470
Median age at first marriage in district for age group 18-21	18.7857	2.2761	14.9605	1.2304	19.4554	2.0786	14.9063	1.0895
Median age at first marriage in district for those over 21 years	22.9224	1.8657	14.5690	1.1904	22.9425	1.9852	14.2566	0.9754

Notes: Data are from DHS 1999 and DHS 2004. Std. Dev. is standard deviation.

Table B2: Enrolment in Tertiary Education: Linear Probability Models
(Sample: 18-21 years in both survey years)

Variables	FE-OLS	Variables	FE-OLS
Male	0.0453	Mother has primary education	0.1097***
	(0.0336)		(0.0264)
Premium 25-35	0.0673	Mother has secondary education	0.1743***
	(0.0513)		(0.0286)
Male*premium 25-35	0.0620	Mother has tertiary education	0.1607***
	(0.0492)		(0.0388)
Age 19	-0.0534**	Father has primary education	0.1172***
	(0.0242)		(0.0324)
Age 20	-0.0975***	Father has secondary education	0.0831***
	(0.0248)		(0.0264)
Age 21	-0.0891***	Father has tertiary education	0.1531***
	(0.0315)		(0.0311)
Muslim	0.0272	At least one parent employed	0.0464**
	(0.0320)		(0.0222)
Number of children 0-4	-0.0699	Mean head count ratios in district	0.0035
	(0.0433)		(0.0032)
Number of school children 5-17	-0.0179**	Mean youth unemployment rate in district	8.2019
	(0.0085)		(6.1611)
Number of males 15-60	-0.0076	Mean of fertility preference in district	-1.4665***
	(0.0096)	(aged over 21 years)	(0.2996)
Number of females 15-60	0.0154	Median age at first marriage in district	0.0399*
	(0.0136)	(aged over 21 years)	(0.0219)
Number of males over 60	0.0108	Mean sex ratio at birth in district	-0.0017
	(0.0298)		(0.0010)
Number of females over 60	0.0314	Constant	-4.4683
	(0.0296)		(3.2094)

Notes: Data are from pooled LFS 1999 and DHS 1999 and LFS 2005 and DHS 2004. FE-OLS is fixed-effect OLS. Figures in parentheses are robust standard errors, corrected for clustering at the district-area level. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table B3: Gender Differences in Enrolment in Tertiary Education: Linear Probability Models

Variables	(1) FE-OLS (Interactions)	(2) FE-OLS (Males)	(3) FE-OLS (Females)
Premium 25-35	0.0440	0.1187	0.0747
	(0.0468)		(0.0513)
Age 19	0.0098	-0.0516	-0.0614*
	(0.0505)		(0.0344)
Age 20	0.1102***	-0.0451	-0.1553***
	(0.0406)		(0.0346)
Age 21	0.0499	-0.0625	-0.1124**
	(0.0519)		(0.0519)
Muslim	0.0459	0.0499	0.0040
	(0.0456)		(0.0415)
Number of children 0-4	0.0791	-0.0411	-0.1202*
	(0.1100)		(0.0720)
Number of school children 5-17	-0.0182	-0.0289	-0.0107
	(0.0187)		(0.0121)
Number of males 15-60	0.0109	0.0012	-0.0097
	(0.0208)		(0.0129)
Number of females 15-60	-0.0452*	-0.0076	0.0376**
	(0.0242)		(0.0183)
Number of males over 60	0.0540	0.0399	-0.0141
	(0.0514)		(0.0447)
Number of females over 60	0.0435	0.0336	-0.0099
	(0.0756)		(0.0529)
Mother has primary education	-0.1512***	0.0435	0.1947***
	(0.0554)		(0.0436)
Mother has secondary education	-0.1054*	0.3318	0.2264***
	(0.0557)		(0.0404)
Mother has tertiary education	-0.1454*	0.0887	0.2341***
	(0.0766)		(0.0675)
Father has primary education	-0.1698*	0.0332	0.2030***
	(0.0884)		(0.0687)
Father has secondary education	-0.0535	0.0527	0.1062**
	(0.0595)		(0.0458)
Father has tertiary education	-0.0487	0.1196	0.1683***
	(0.0542)		(0.0450)

Continued on next page

Table B3 (continued): Gender Differences in Enrolment in Tertiary Education: Linear Probability Models

Variables	(1) FE-OLS (Interactions)	(2) FE-OLS (Males)	(3) FE-OLS (Females)
At least one parent employed	-0.0287	0.0239	0.0526
	(0.0411)		(0.0367)
Mean head count ratios in district	-0.0000	0.0031	0.0031
	(0.0023)		(0.0033)
Mean youth unemployment rate in district	-0.0585	7.2556	7.3141
	(0.2175)		(5.9477)
Mean of fertility preference in district	1.0588*	-0.9253	-1.9841***
(aged over 21 years)	(0.6250)		(0.4923)
Median age at first marriage in district	0.0097	0.0413	0.0316
(aged over 21 years)	(0.0132)		(0.0239)
Mean sex ratio at birth in district	-0.0021*	-0.0028	-0.0007
	(0.0011)		(0.0012)
Constant	-3.9871	-3.7571	-3.9871
	(3.0916)		(3.0916)
Male	0.2300		
	(0.3378)		
R-squared	0.17		
F -test	5.20		
p>F	(0.0000)		
N	1,927	1,078	849

Notes: Data are from pooled LFS 1999 and DHS 1999 and LFS 2005 and DHS 2004. FE-OLS is fixed-effect OLS. Figures in parentheses are robust standard errors, corrected for clustering at the district-area level. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table B4: Enrolment in Tertiary Education by Gender: Linear Probability Models using Median Wage Premiums
(Sample: 18-21 years in both survey years)

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
Premium 25-35	0.1755**	0.1567**	0.1473*	0.1584**	0.1336	0.1135*	0.0913	0.0888
	(0.0705)	(0.0785)	(0.0790)	(0.0740)	(0.0846)	(0.0663)	(0.0736)	(0.0693)
Age 19	-0.0429	-0.0482	-0.0463	-0.0498	-0.0403	-0.0374	-0.0410	-0.0491
	(0.0422)	(0.0391)	(0.0392)	(0.0391)	(0.0339)	(0.0317)	(0.0320)	(0.0336)
Age 20	-0.0242	-0.0329	-0.0313	-0.0428	-0.1992***	-0.1369***	-0.1345***	-0.1405***
	(0.0264)	(0.0299)	(0.0302)	(0.0292)	(0.0368)	(0.0353)	(0.0342)	(0.0341)
Age 21	-0.0274	-0.0434	-0.0422	-0.0563*	-0.1028**	-0.0667	-0.0688	-0.0861
	(0.0307)	(0.0299)	(0.0307)	(0.0287)	(0.0464)	(0.0496)	(0.0498)	(0.0525)
Muslim	0.0320	0.0609	0.0631	0.0668	-0.0282	-0.0039	0.0045	-0.0002
	(0.0438)	(0.0424)	(0.0418)	(0.0420)	(0.0544)	(0.0474)	(0.0465)	(0.0441)
Number of children 0-4		-0.0387	-0.0384	-0.0414		-0.0886	-0.0990	-0.1025
		(0.0650)	(0.0648)	(0.0627)		(0.0767)	(0.0752)	(0.0766)
Number of school children 5-17		-0.0278**	-0.0280**	-0.0263**		-0.0097	-0.0074	-0.0031
		(0.0126)	(0.0127)	(0.0128)		(0.0125)	(0.0118)	(0.0130)
Number of males 15-60		-0.0043	-0.0061	-0.0047		-0.0106	-0.0096	-0.0078
		(0.0154)	(0.0153)	(0.0154)		(0.0147)	(0.0143)	(0.0141)
Number of females 15-60		-0.0138	-0.0141	-0.0146		0.0374*	0.0404**	0.0401**
		(0.0176)	(0.0176)	(0.0182)		(0.0202)	(0.0197)	(0.0195)
Number of males over 60		0.0405	0.0377	0.0344		-0.0096	-0.0154	-0.0153
		(0.0371)	(0.0375)	(0.0373)		(0.0474)	(0.0471)	(0.0465)
Number of females over 60		0.0497	0.0474	0.0363		0.0107	0.0073	-0.0061
		(0.0423)	(0.0428)	(0.0420)		(0.0509)	(0.0506)	(0.0534)
Mother has primary education		0.0154	0.0143	0.0146		0.1659***	0.1671***	0.1609***
		(0.0324)	(0.0324)	(0.0325)		(0.0465)	(0.0442)	(0.0449)
Mother has secondary education		0.1076***	0.1020***	0.0976***		0.2108***	0.2094***	0.2090***
		(0.0365)	(0.0369)	(0.0369)		(0.0437)	(0.0424)	(0.0432)
Mother has tertiary education		0.0929**	0.0907**	0.0725*		0.2170***	0.2159***	0.2145***
		(0.0408)	(0.0411)	(0.0398)		(0.0709)	(0.0722)	(0.0733)

Continued on next page

Table B4 (continued): Enrolment in Tertiary Education by Gender: Linear Probability Models using Median Wage Premiums (Sample: 18-21 years in both survey years)

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
Father has primary education		0.0384	0.0396	0.0303		0.2334***	0.2370***	0.2298***
		(0.0418)	(0.0420)	(0.0434)		(0.0684)	(0.0673)	(0.0695)
Father has secondary education		0.0662*	0.0680*	0.0574		0.1158**	0.1159**	0.1162**
		(0.0360)	(0.0360)	(0.0359)		(0.0465)	(0.0460)	(0.0476)
Father has tertiary education		0.1313***	0.1320***	0.1346***		0.1738***	0.1727***	0.1668***
		(0.0378)	(0.0378)	(0.0360)		(0.0448)	(0.0452)	(0.0475)
At least one parent employed		0.0313	0.0347	0.0305		0.0647	0.0551	0.0448
		(0.0249)	(0.0240)	(0.0232)		(0.0409)	(0.0393)	(0.0388)
Mean head count ratios in district			0.0005	0.0034			0.0052	0.0083*
			(0.0027)	(0.0032)			(0.0037)	(0.0045)
Mean youth unemployment rate in district			8.2572*	5.0778			-14.3818***	-7.5994
			(4.4079)	(4.5921)			(4.8407)	(4.9943)
Mean of fertility preference in district				-0.7434				-1.1658***
(aged over 21 years)				(0.5235)				(0.4062)
Median age at first marriage in district				0.0332**				0.0062
(aged over 21 years)				(0.0151)				(0.0238)
Mean sex ratio at birth in district				-0.0036**				-0.0005
				(0.0014)				(0.0014)
Constant	0.7244***	0.6418***	-1.8276	-1.3647	0.8035***	0.4469***	11.4090***	6.0413
	(0.0625)	(0.0737)	(1.3058)	(1.2686)	(0.0536)	(0.0860)	(3.6896)	(3.8563)
N	1,078	1,078	1,078	1,078	849	849	849	849
R-squared	0.0108	0.0646	0.0673	0.0813	0.0399	0.2163	0.2260	0.2363

Notes: Data are from pooled LFS 1999 and LFS 2005. Analyses of Columns 4 and 8 are derived from pooled LFS 1999 and DHS 1999 and LFS 2005 and DHS 2004. FE-OLS is fixed-effect OLS. Figures in parentheses are robust standard errors, corrected for clustering at the district-area level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. p denotes p-value.

Appendix C

Table C1: First Stage OLS Estimates of Child Labour Hours (Sample 5-17)

	(1)		(2)		(3)		(4)		(5)	
Variables	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2
Child's age	-4.6797***	-364.2844***	-4.6535***	-362.5228***	-4.2318***	-335.8697***	-4.1576***	-331.1257***	-2.6349***	-239.3926***
	(0.5685)	(45.3490)	(0.5682)	(45.3341)	(0.5627)	(45.0852)	(0.5611)	(45.0095)	(0.5543)	(44.8632)
Child's age (squared)	0.2537***	18.2197***	0.2527***	18.1515***	0.2361***	17.0991***	0.2337***	16.9498***	0.1838***	13.9708***
	(0.0217)	(1.7345)	(0.0217)	(1.7340)	(0.0215)	(1.7246)	(0.0215)	(1.7216)	(0.0212)	(1.7121)
Female	-7.8234***	-558.0371***	-7.8609***	-560.1888***	-8.3044***	-587.3813***	-8.6962***	-612.1867***	-9.1009***	-639.4684***
	(0.3671)	(29.2830)	(0.3745)	(29.8793)	(0.3869)	(30.9968)	(0.4037)	(32.3822)	(0.3951)	(31.9826)
Number of children 0-4									0.5522***	39.4679**
									(0.2064)	(16.7104)
Number of school children 5-17									0.3561***	21.5989**
									(0.1263)	(10.2205)
Number of adult males over 17 years									-0.2842	-17.6645
									(0.2043)	(16.5361)
Number of adult females over 17 years									-0.3300	-24.3213
									(0.2593)	(20.9867)
Occupation of father									-1.6757***	-127.0299***
									(0.4063)	(32.8828)
Household uses piped water									5.2520***	440.8984***
									(0.9445)	(76.4504)
Household has a TV									-1.9084***	-102.5120***
									(0.4732)	(38.2988)
Household has a radio									0.0424	-8.8819
									(0.3497)	(28.3014)
Household has a bicycle									-1.9197***	-143.8875***
									(0.3819)	(30.9097)
Formal school									2.4040***	149.1312***
									(0.3807)	(30.8143)
NGO school									-0.1983	8.9358
									(0.5056)	(40.9233)
Own marginal land, less than 0.5 acre									-4.7312***	-282.1031***
									(0.3285)	(26.5927)
Own large land, greater than 2 acre									-6.6715***	-498.2775***
									(1.5208)	(123.0984)
Urban									-1.4713***	-149.3000***
									(0.3813)	(30.8615)

Continued on next page

Table C1 (continued): First Stage OLS Estimates of Child Labour Hours (Sample 5-17)

	(1)		(2)		(3)		(4)		(5)	
Variables	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2
Mother more educated than father			-2.8556***	-188.2868***						
			(0.8267)	(65.9641)						
Mother more educated than father x Female			0.5226	26.5405						
			(1.8021)	(143.7829)						
Mother's education (highest grade)					-3.1184***	-197.2523***	-1.7419***	-1.8912***	-1.4778***	-98.2428***
					(0.2211)	(17.7113)	(0.3227)	(0.2736)	(0.2694)	(21.8069)
Mother's education x Female					1.5313***	92.9690**	0.4469	0.5108	0.8163	48.1522
					(0.4776)	(38.2628)	(0.6692)	(0.5754)	(0.5606)	(45.3733)
Father's education (highest grade)							-1.4409***	-1.3730***	-0.6711***	-45.9580***
							(0.1974)	(0.1815)	(0.1809)	(14.6401)
Father's education x Female							1.1887***	1.1531***	1.0188***	65.1151**
							(0.3729)	(0.3476)	(0.3388)	(27.4260)
Agriculture	-10.6330***	-712.9078***	-10.6944***	-717.0280***	-10.4795***	-703.2634***	-10.3606***	-694.7708***	-9.4509***	-630.3137***
	(0.8604)	(68.6288)	(0.8600)	(68.6162)	(0.8505)	(68.1382)	(0.8483)	(68.0233)	(0.8593)	(69.5562)
Manufacturing	1.0169	199.9034***	0.9776	197.2110***	1.2118	212.0977***	1.3700	222.6538***	1.6431*	235.4989***
	(0.9168)	(73.1279)	(0.9164)	(73.1141)	(0.9064)	(72.6169)	(0.9044)	(72.5158)	(0.8818)	(71.3771)
Wholesale and Retail	-1.8629**	-17.7724	-1.8673**	-18.1278	-1.3689	13.5565	-1.2142	23.8175	-0.7314	54.8788
	(0.8889)	(70.9055)	(0.8884)	(70.8833)	(0.8792)	(70.4402)	(0.8770)	(70.3249)	(0.8565)	(69.3256)
Services	-2.3840**	-68.2391	-2.3645**	-67.0198	-1.9527*	-40.6452	-1.8156	-32.0466	-1.3494	-6.4719
	(1.1652)	(92.9469)	(1.1645)	(92.9156)	(1.1524)	(92.3259)	(1.1496)	(92.1957)	(1.1205)	(90.6986)
Constant	51.9231***	3,077.5452***	51.9346***	3,078.1816***	49.8776***	2,947.6196***	49.6883***	2,932.3314***	43.8524***	2,718.9261***
	(3.7588)	(299.8323)	(3.7564)	(299.7104)	(3.7178)	(297.8616)	(3.7076)	(297.3305)	(3.8030)	(307.8235)
F -test on instruments	257.13	244.74	259.24	246.31	271.15	254.7	256.45	273.69	149.68	137.15
p>F	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Adj. R-squared	0.24	0.20	0.24	0.20	0.26	0.21	0.21	0.26	0.30	0.24
N	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900	8,900

Notes: Data are from NCLS 2002. Robust standard errors are in parentheses. Adj. R-squared is adjusted R-squared. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table C2: First Stage OLS Estimates of Child Labour Hours (Sample 7-17)

Variables	(1) Child labour hours	(1) Child labour $hours^2$	(2) Child labour hours	(2) Child labour $hours^2$	(3) Child labour hours	(3) Child labour $hours^2$	(4) Child labour hours	(4) Child labour $hours^2$	(5) Child labour hours	(5) Child labour $hours^2$
Child's age	-5.8860***	-444.9689***	-5.8488***	-442.4572***	-5.1783***	-400.1323***	-5.0317***	-390.8067***	-3.0480***	-269.9108***
	(0.7151)	(57.1144)	(0.7147)	(57.0987)	(0.7086)	(56.8410)	(0.7067)	(56.7590)	(0.6973)	(56.5286)
Child's age (squared)	0.2973***	21.1350***	0.2959***	21.0397***	0.2703***	19.4209***	0.2653***	19.1054***	0.1988***	15.0732***
	(0.0268)	(2.1423)	(0.0268)	(2.1417)	(0.0266)	(2.1322)	(0.0265)	(2.1290)	(0.0261)	(2.1165)
Female	-7.8539***	-561.0331***	-6.6195*	-500.7728*	-8.3566***	-591.6035***	-8.7422***	-616.3955***	-9.1892***	-646.1983***
	(0.3687)	(29.4485)	(3.5622)	(284.5842)	(0.3888)	(31.1860)	(0.4056)	(32.5745)	(0.3969)	(32.1717)
Number of children 0-4									0.5578***	39.2850**
									(0.2073)	(16.8088)
Number of school children 5-17									0.3476***	21.0636**
									(0.1267)	(10.2699)
Number of adult males over 17 years									-0.2984	-18.6911
									(0.2043)	(16.5629)
Number of adult females over 17 years									-0.3478	-25.4499
									(0.2593)	(21.0209)
Occupation of father									-1.7000***	-129.1848***
									(0.4071)	(32.9996)
Household uses piped water									5.2083***	437.4386***
									(0.9434)	(76.4765)
Household has a television									-1.8963***	-101.9894***
									(0.4729)	(38.3332)
Household has a radio									0.0127	-10.5106
									(0.3502)	(28.3876)
Household has a bicycle									-1.8823***	-140.8615***
									(0.3825)	(31.0064)
Formal school									-4.7858***	-285.2718***
									(0.3289)	(26.6612)

Continued on next page

Table C2 (continued): First Stage OLS Estimates of Child Labour Hours (Sample 7-17)

	(1)		(2)		(3)		(4)		(5)	
Variables	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2	Child labour hours	Child labour hours^2
NGO school									-6.7112***	-500.8623***
									(1.5187)	(123.1178)
Own marginal land, less than 0.5 acre									2.4325***	151.3531***
									(0.3811)	(30.8972)
Own large land, greater than 2 acre									-0.2411	6.2245
									(0.5069)	(41.0905)
Urban									-1.5230***	-153.6383***
									(0.3816)	(30.9355)
Mother more educated than father			2.8295***	186.3791***						
			(0.8272)	(66.0845)						
Mother more educated than father x Female			-0.6386	-31.3245						
			(1.8093)	(144.5424)						
Mother's education (highest grade)					-3.1084***	-196.3802***	-1.8727***	-117.5719***	-1.4635***	-97.2046***
					(0.2220)	(17.8122)	(0.2756)	(22.1328)	(0.2712)	(21.9818)
Mother's education x Female					1.5734***	95.1152**	0.5576	29.7853	0.8737	50.2085
					(0.4794)	(38.4607)	(0.5778)	(46.4095)	(0.5627)	(45.6130)
Father's education (highest grade)							-1.3732***	-87.5733***	-0.6663***	-45.7605***
							(0.1824)	(14.6467)	(0.1816)	(14.7200)
Father's education x Female							1.1380***	73.1562***	1.0129***	65.7800**
							(0.3489)	(28.0181)	(0.3398)	(27.5488)
Agriculture	-10.5471***	-703.6908***	-10.6075***	-707.7743***	-10.3962***	-694.2360***	-10.2658***	-685.9097***	-9.3371***	-619.2119***
	(0.8624)	(68.8775)	(0.8620)	(68.8667)	(0.8526)	(68.3943)	(0.8501)	(68.2791)	(0.8610)	(69.8000)
Manufacturing	1.1056	210.0515***	1.0679	207.4371***	1.3057	222.5491***	1.4666	232.8447***	1.7605**	247.5094***
	(0.9187)	(73.3763)	(0.9183)	(73.3639)	(0.9084)	(72.8743)	(0.9061)	(72.7728)	(0.8834)	(71.6150)
Wholesale and Retail	-1.7863**	-8.5738	-1.7913**	-8.9875	-1.2984	22.3474	-1.1330	32.8950	-0.6303	65.6976
	(0.8907)	(71.1440)	(0.8903)	(71.1230)	(0.8812)	(70.6859)	(0.8787)	(70.5721)	(0.8580)	(69.5570)
Services	-2.3868**	-63.3046	-2.3668**	-62.0749	-1.9575*	-35.8353	-1.8249	-27.4401	-1.3162	0.7437
	(1.1669)	(93.1979)	(1.1662)	(93.1681)	(1.1542)	(92.5885)	(1.1511)	(92.4507)	(1.1217)	(90.9281)
Constant	60.0304***	3,616.7075***	54.3067***	3,239.3812***	56.2252***	3,375.4621***	55.4972***	3,329.1966***	46.7709***	2,932.4285***
	(4.7689)	(380.8950)	(5.0453)	(403.0681)	(4.7215)	(378.7573)	(4.7083)	(378.1515)	(4.7505)	(385.1002)
F -test on instruments	255.28	242.82	257.34	244.35	269.24	252.72	271.88	254.56	148.69	136.14
p>F	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)	(p = 0.0000)
Adj. R-squared	0.24	0.20	0.24	0.20	0.26	0.21	0.26	0.21	0.30	0.24
N	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848	8,848

Notes: Data are from NCLS 2002. Robust standard errors are in parentheses. Adj. R-squared is adjusted R-squared. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Appendix D

Table D1: Description of Key Variables used in Regression, by Child Work Status

	Workers			Non-workers			
Variables	N	Mean	Std. Dev.	N	Mean	Std. Dev.	*t*-test
Child covariates							
Age (in years)	14,437	13.1799	1.8093	1,573	12.3814	3.2704	15.03 ***
Age (squared)	14,437	176.9826	47.5771	1,573	163.9886	78.3740	9.52 ***
Female (1= female)	14,437	0.2103	0.4075	1,573	0.3884	0.4875	-16.1 ***
Vaccination (1= yes)	14,437	0.5248	0.4994	1,573	0.6796	0.4668	-11.8 ***
Hours (weekly hours of work)	14,437	21.6545	14.6729	-	-	-	-
Protection (1 = if child receives working dress)	14,437	0.0101	0.1001	-	-	-	-
Household covariates							
Number of children for each child in the household	14,437	2.0449	1.3805	1,573	1.9784	1.3581	1.82 *
Number of adults over 17 years	14,437	2.7842	1.1648	1,573	2.8252	1.2343	-1.31
Father's age (in years)	14,437	47.5804	9.7937	1,573	46.9288	10.6240	2.48 **
Father has no education (1=yes)	14,437	0.5577	0.4967	1,573	0.6421	0.4795	-6.42 ***
Father has primary education (1= if father has completed Grade 5)	14,437	0.1633	0.3697	1,573	0.1093	0.3122	5.58 ***
Father has secondary education (1=if father has completed Grade 10 or more)	14,437	0.2306	0.4212	1,573	0.1589	0.3657	6.49 ***
Mother's age (in years)	14,437	38.5372	7.9360	1,573	37.8265	8.7369	3.34 ***
Mother has no education (1=yes)	14,437	0.6955	0.4602	1,573	0.7667	0.4231	-5.87 ***
Mother has primary education (1= if mother has completed Grade 5)	14,437	0.1516	0.3586	1,573	0.1113	0.3145	4.28 ***
Mother has secondary education (1=if mother has completed Grade 10 or more)	14,437	0.1454	0.3525	1,573	0.1017	0.3024	4.73 ***
Sanitation OK (1 = if the household has a sanitary toilet)	14,437	0.0197	0.1391	1,573	0.0006	0.0252	5.44 ***
Safe drinking water (1 = if the main source of household drinking water is tape water/tube well)	14,437	0.9483	0.2214	1,573	0.9142	0.2802	5.65 ***
Electricity (1 = if the main source of household lighting is electricity)	14,437	0.3440	0.4751	1,573	0.4650	0.4650	2.23 **
Number of rooms in the household	14,437	2.3412	1.2178	1,573	2.3560	1.3073	-0.45
Communal covariates							
Urban (1 = urban, 0 = rural)	14,437	0.3028	0.4595	1,573	0.2543	0.4356	4.00 ***
Instruments							
Informal school (1 = informal education activities)	14,437	0.8818	0.3228	1,573	0.7699	0.4211	1.90 *
Migration status (1 = if the household leaves the usual place of residence to find work)	14,437	0.0012	0.0353	1,573	0.0089	0.0939	-6.46 ***

Notes: Data are from NCLS 2002. Std. Dev. is standard deviation. *t*-test for difference (Working-Non-working children). *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table D2: First Stage OLS Estimates of Child Labour Hours

Variables	(1) All	(2) Rural Sample
Child's age	-0.1906***	-0.1948***
	(0.0238)	(0.0259)
Child's age (squared)	0.0114***	0.0117***
	(0.0009)	(0.0010)
Female	0.3872***	0.6651***
	(0.1030)	(0.1206)
Age*female	-0.0433***	-0.0664***
	(0.0080)	(0.0095)
Child's vaccination	-0.0120	-0.0197*
	(0.0097)	(0.0108)
Number of children for each child in the household	0.0179***	0.0053
	(0.0029)	(0.0034)
Number of adults over 17 years	-0.0278***	-0.0281***
	(0.0040)	(0.0047)
Agriculture	-0.1498***	-0.2835***
	(0.0366)	(0.0472)
Manufacturing	0.1680***	-0.0738
	(0.0385)	(0.0494)
Construction	0.3565***	0.3383***
	(0.0443)	(0.0608)
Wholesale and Retail	-0.0708*	-0.2558***
	(0.0374)	(0.0492)
Father's age	0.0011	0.0028***
	(0.0007)	(0.0009)
Father has primary education	-0.0365***	0.0136
	(0.0112)	(0.0116)
Father has secondary education	0.0378***	0.0451***
	(0.0106)	(0.0106)
Mother's age	-0.0008	-0.0036***
	(0.0010)	(0.0011)
Mother has primary education	-0.0957***	-0.1108***
	(0.0112)	(0.0126)
Mother has secondary education	-0.2746***	-0.2463***
	(0.0108)	(0.0114)
Urban	-0.0443***	
	(0.0104)	
Electricity	-0.0837***	-0.0834***
	(0.0094)	(0.0097)
Migration status	0.1874	0.1523*
	(0.1302)	(0.0851)
Informal school	0.0779***	0.0921***
	(0.0295)	(0.0293)
Constant	3.6972***	3.8162***
	(0.1645)	(0.1845)
F -test on instruments	4.53	6.65
p>F	(0.0108)	(0.0013)
R-squared	0.2718	0.2736
N	14,437	10,065

Notes: Data are from NCLS 2002. Robust standard errors are in parentheses.
*** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table D3: Bivariate Probit Estimates of Injury/Illness and Child Work

Variables	Injury/Illness	Work	Tiredness/Exhaustion	Work	Body injuries	Work	Backache	Work	Other health problems	Work
Child's age	-1.2218***	1.2462***	-1.2840***	1.2306***	-0.2505**	1.2284***	-0.3364*	1.2273***	-0.8033***	1.2323***
	(0.0548)	(0.0506)	(0.0874)	(0.0547)	(0.1140)	(0.0502)	(0.1860)	(0.0503)	(0.0641)	(0.0504)
Child's age (squared)	0.0530***	-0.0477***	0.0510***	-0.0470***	0.0144***	-0.0469***	0.0151**	-0.0469***	0.0369***	-0.0471***
	(0.0020)	(0.0019)	(0.0029)	(0.0021)	(0.0041)	(0.0019)	(0.0069)	(0.0019)	(0.0024)	(0.0019)
Female	-1.4696***	0.0757	-0.9129**	0.1191	-0.7234	0.0662	-2.5770***	0.0577	-1.9526***	0.0534
	(0.2286)	(0.2014)	(0.4031)	(0.2190)	(0.5230)	(0.2018)	(0.3949)	(0.2043)	(0.7244)	(0.2020)
Age*female	0.0436**	-0.0386**	0.0473**	-0.0428**	-0.0088	-0.0372**	0.1575***	-0.0364**	0.0724	-0.0362**
	(0.0174)	(0.0155)	(0.0194)	(0.0175)	(0.0377)	(0.0155)	(0.0303)	(0.0157)	(0.0506)	(0.0155)
Child's vaccination	-0.2100***		-0.1503**		-0.1111**		-0.1414**		-0.0676*	
	(0.0286)		(0.0751)		(0.0432)		(0.0557)		(0.0405)	
Agriculture	-0.1911**		-0.1336		-0.6999***		0.1308		0.0825	
	(0.0844)		(0.1365)		(0.1260)		(0.1116)		(0.1198)	
Manufacturing	0.6678***		0.1039		0.4454***		0.6890***		0.5067***	
	(0.0878)		(0.1394)		(0.1254)		(0.1242)		(0.1262)	
Construction	0.8561***		0.1750		0.9042***		0.0865		0.1416	
	(0.1070)		(0.1438)		(0.1350)		(0.1804)		(0.1487)	
Wholesale and Retail	-0.1865**		-0.4240***		-0.1196		0.2663**		0.0316	
	(0.0902)		(0.0909)		(0.1242)		(0.1172)		(0.1334)	
Child's work	1.3265***		1.9037***		0.7836***		-0.0697		0.6724***	
	(0.1308)		(0.4078)		(0.1534)		(0.5943)		(0.1645)	
Number of children for each child in the household	0.0124	0.0334***	0.0608**	0.0403***	0.0654***	0.0332***	0.0094	0.0338***	-0.1082***	0.0332***
	(0.0105)	(0.0111)	(0.0285)	(0.0135)	(0.0147)	(0.0111)	(0.0209)	(0.0112)	(0.0176)	(0.0110)
Number of adults over 17 years	-0.0693***	-0.0569***	-0.1554**	-0.0578***	0.0302	-0.0586***	-0.0623**	-0.0591***	0.0624***	-0.0596***
	(0.0158)	(0.0137)	(0.0725)	(0.0137)	(0.0220)	(0.0137)	(0.0276)	(0.0137)	(0.0209)	(0.0137)
Father's age	-0.0098***	-0.0013	0.0053	-0.0015	0.0015	-0.0017	-0.0368***	-0.0018	-0.0127***	-0.0017
	(0.0028)	(0.0029)	(0.0034)	(0.0028)	(0.0045)	(0.0029)	(0.0049)	(0.0029)	(0.0042)	(0.0030)
Father has primary education	-0.2746***	0.1908***	-0.1119*	0.1980***	-0.5063***	0.2005***	-0.1414*	0.2022***	-0.2460***	0.1984***
	(0.0413)	(0.0466)	(0.0597)	(0.0458)	(0.0814)	(0.0464)	(0.0820)	(0.0466)	(0.0555)	(0.0465)

Continued on next page

Table D3 (continued): Bivariate Probit Estimates of Injury/Illness and Child Work

Variables	Injury/Illness	Work	Tiredness/Exhaustion	Work	Body injuries	Work	Backache	Work	Other health problems	Work
Father has secondary education	-0.0612	0.1658***	-0.2486***	0.1837***	0.2172***	0.1852***	0.0952	0.1893***	-0.3149***	0.1835***
	(0.0411)	(0.0436)	(0.0470)	(0.0435)	(0.0535)	(0.0436)	(0.0728)	(0.0431)	(0.0633)	(0.0431)
Mother's age	0.0151***	0.0091**	0.0019	0.0096**	0.0039	0.0096***	0.0417***	0.0096***	0.0036	0.0096***
	(0.0034)	(0.0037)	(0.0046)	(0.0038)	(0.0054)	(0.0037)	(0.0054)	(0.0037)	(0.0051)	(0.0037)
Mother has primary education	-0.3416***	0.1057**	-0.1761***	0.0779	-0.4150***	0.0991**	-0.6181***	0.0989**	-0.0061	0.1076**
	(0.0455)	(0.0478)	(0.0505)	(0.0523)	(0.0721)	(0.0484)	(0.1242)	(0.0485)	(0.0609)	(0.0486)
Mother has secondary education	-1.0300***	0.1161**	-0.3064***	0.1167**	-1.1086***	0.1041**	-1.2290***	0.1017**	-0.9539***	0.1088**
	(0.0741)	(0.0506)	(0.1090)	(0.0588)	(0.1198)	(0.0504)	(0.2305)	(0.0503)	(0.1283)	(0.0506)
Protection	0.5688***		-0.7933***		-1.0397***		0.5547***		1.2916***	
	(0.1252)		(0.2593)		(0.2597)		(0.1489)		(0.1128)	
Urban	-0.0476	-0.1591***	-0.0463	-0.1444***	-0.0490	-0.1658***	-0.3915***	-0.1655***	0.4883***	-0.1636***
	(0.0353)	(0.0346)	(0.1129)	(0.0448)	(0.0470)	(0.0349)	(0.0593)	(0.0349)	(0.0541)	(0.0348)
Safe drinking water	-0.5063***		-0.6897***		0.5330***		0.3517**		-0.2121***	
	(0.0611)		(0.1687)		(0.1572)		(0.1374)		(0.0806)	
Electricity	-0.3504***	-0.0935***	-0.0483	-0.0897**	-0.2285***	-0.0984***	-0.5710***	-0.0997***	-0.1160**	-0.0986***
	(0.0364)	(0.0331)	(0.0640)	(0.0366)	(0.0495)	(0.0331)	(0.0712)	(0.0331)	(0.0498)	(0.0330)
Number of rooms in the household	-0.0867***		-0.0972***		-0.0627***		0.0807***		-0.1136***	
	(0.0135)		(0.0364)		(0.0206)		(0.0221)		(0.0201)	
Sanitation OK	-1.1004***		-0.5521***		-5.0658***		-4.4522***		-5.2457***	
	(0.1682)		(0.2009)		(0.1163)		(0.0993)		(0.1190)	
Migration status		9.4177***		-0.5241***		9.1034***		9.5335***		9.7429***
		(0.3523)		(0.1987)		(0.3764)		(0.3304)		(0.3745)
Migration status x urban		-4.9654***		-4.4817		-4.8381***		-5.0717***		-5.1468***
		(0.2540)		-0.2672		(0.2717)		(0.3061)		(0.3045)
Informal school		4.0714***		3.4670*		4.1485***		4.2962***		4.2031***
		(0.0896)		(2.0886)		(0.0487)		(0.0482)		(0.0652)
Constant	5.8036***	-6.4459***	5.9117***	-6.4003***	-1.9209***	-6.3398***	0.2823	-6.3292***	1.9460***	-6.3556***
	(0.3387)	(0.3248)	(0.3152)	(0.3177)	(0.7115)	(0.3254)	(0.7530)	(0.3271)	(0.4236)	(0.3254)
N	16,010	16,010	16,010	16,010	16,010	16,010	16,010	16,010	16,010	16,010

Notes: Data are from NCLS 2002. Robust standard errors are in parentheses. The omitted categories are male child, no vaccination, service sector, no schooling, no working dress, rural, source of drinking water is ponds/rivers, no electricity, no sanitary latrine, if the household does not leave their place of residence during the last 12 months and the formal public schools and/or the NGO schools. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table D4: Partially Linear Model Estimates of Injury/Illness

Variables	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Child's age	-0.2193***	-0.1157***	-0.1819***	0.1503***	-0.0720**
	(0.0502)	(0.0399)	(0.0353)	(0.0283)	(0.0362)
Child's age (squared)	0.0120***	0.0056**	0.0112***	-0.0081***	0.0033
	(0.0028)	(0.0023)	(0.0020)	(0.0016)	(0.0020)
Female	0.0059	-0.0259	0.2318***	-0.1974***	-0.0026
	(0.1008)	(0.0800)	(0.0709)	(0.0567)	(0.0727)
Age*female	-0.0137	-0.0019	-0.0327***	0.0221***	-0.0012
	(0.0104)	(0.0083)	(0.0073)	(0.0059)	(0.0075)
Child's vaccination	-0.1022***	-0.0522***	-0.0570***	-0.0158***	0.0228***
	(0.0106)	(0.0084)	(0.0075)	(0.0060)	(0.0077)
Agriculture	0.0712*	0.0351	-0.0757***	0.1335***	-0.0218
	(0.0395)	(0.0313)	(0.0278)	(0.0222)	(0.0285)
Manufacturing	0.3336***	0.1197***	0.2496***	-0.0127	-0.0230
	(0.0423)	(0.0336)	(0.0298)	(0.0238)	(0.0305)
Construction	0.5794***	0.1896***	0.4503***	-0.0740	0.0135
	(0.0837)	(0.0664)	(0.0588)	(0.0471)	(0.0604)
Wholesale and Retail	0.0559**	0.0173	-0.0070	0.0927***	-0.0470**
	(0.0257)	(0.0204)	(0.0181)	(0.0145)	(0.0185)
Number of children for each child in the household	0.0196***	0.0182***	0.0190***	-0.0126***	-0.0050
	(0.0051)	(0.0040)	(0.0036)	(0.0029)	(0.0037)
Number of adults over 17 years	-0.0337***	-0.0320***	-0.0190***	0.0147***	0.0026
	(0.0070)	(0.0056)	(0.0049)	(0.0040)	(0.0051)
Father's age	-0.0028***	-0.0015***	0.0012***	-0.0024***	-0.0001
	(0.0006)	(0.0005)	(0.0004)	(0.0003)	(0.0004)
Father has primary education	-0.0407***	-0.0143	-0.0240***	0.0286***	-0.0310***
	(0.0114)	(0.0091)	(0.0080)	(0.0064)	(0.0082)
Father has secondary education	0.0005	-0.0219**	0.0521***	-0.0037	-0.0260***
	(0.0122)	(0.0097)	(0.0085)	(0.0068)	(0.0088)
Mother's age	0.0047***	0.0019***	-0.0003	0.0020***	0.0010*
	(0.0007)	(0.0006)	(0.0005)	(0.0004)	(0.0005)
Mother has primary education	-0.0994***	-0.0431**	-0.0867***	0.0273**	0.0031
	(0.0236)	(0.0187)	(0.0166)	(0.0133)	(0.0170)
Mother has secondary education	-0.2403***	-0.0920*	-0.2517***	0.1238***	-0.0204
	(0.0631)	(0.0501)	(0.0444)	(0.0355)	(0.0455)
Protection	0.1376***	0.0399	-0.0049	0.1066***	-0.0040
	(0.0342)	(0.0272)	(0.0240)	(0.0192)	(0.0247)
Urban	-0.0593***	-0.0353***	-0.0504***	0.0128*	0.0135
	(0.0122)	(0.0097)	(0.0086)	(0.0069)	(0.0088)
Safe drinking water	-0.0229*	-0.0880***	0.0472***	0.0063	0.0116
	(0.0130)	(0.0103)	(0.0091)	(0.0073)	(0.0094)
Electricity	-0.0927***	-0.0147	-0.0852***	0.0310***	-0.0237*
	(0.0198)	(0.0157)	(0.0139)	(0.0112)	(0.0143)
Number of rooms in the household	-0.0020	-0.0026	0.0005	0.0039**	-0.0039**
	(0.0027)	(0.0022)	(0.0019)	(0.0015)	(0.0020)
Sanitation OK	-0.0283	0.0049	-0.0110	-0.0175	-0.0048
	(0.0247)	(0.0196)	(0.0174)	(0.0139)	(0.0178)
Residual	0.7940***	0.3491*	0.8677***	-0.4932***	0.0703
	(0.2256)	(0.1791)	(0.1586)	(0.1269)	(0.1627)
N	14,436	14,436	14,436	14,436	14,436

Notes: Data are from NCLS 2002. Standard errors in parentheses. The omitted categories are male child, no vaccination, service sector, no schooling, no working dress, rural, source of drinking water is ponds/rivers, no electricity and no sanitary latrine. The model is fitted by first order differencing. Thus, the sample size is reduced to 14,436 instead of 14,437 (see Lokshin 2006 for more discussion on this issue).
*** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

Table D5: Partially Linear Model Estimates of Injury/Illness, Rural Sample

Variables	Injury/Illness	Tiredness/Exhaustion	Body injuries	Backache	Other health problems
Child's age	-0.2327***	-0.1593***	-0.1518***	0.1891***	-0.1107**
	(0.0613)	(0.0489)	(0.0413)	(0.0335)	(0.0456)
Child's age (squared)	0.0117***	0.0072**	0.0081***	-0.0106***	0.0070***
	(0.0036)	(0.0028)	(0.0024)	(0.0019)	(0.0027)
Female	0.0853	0.0334	0.0910	-0.4185***	0.3794***
	(0.1941)	(0.1546)	(0.1307)	(0.1061)	(0.1444)
Age*female	-0.0168	-0.0091	-0.0153	0.0454***	-0.0379***
	(0.0189)	(0.0151)	(0.0127)	(0.0103)	(0.0141)
Child's vaccination	-0.0854***	-0.0255**	-0.0425***	-0.0076	-0.0098
	(0.0131)	(0.0104)	(0.0088)	(0.0071)	(0.0097)
Agriculture	0.0237	-0.0608	-0.0727	0.2658***	-0.1087*
	(0.0847)	(0.0675)	(0.0570)	(0.0463)	(0.0630)
Manufacturing	0.1068***	-0.0280	0.0325	0.1226***	-0.0203
	(0.0371)	(0.0295)	(0.0250)	(0.0203)	(0.0276)
Construction	0.5535***	0.0753	0.3192***	-0.1360**	0.2950***
	(0.1028)	(0.0819)	(0.0692)	(0.0562)	(0.0765)
Wholesale and Retail	-0.0231	-0.0819	-0.0425	0.2435***	-0.1422**
	(0.0775)	(0.0617)	(0.0522)	(0.0423)	(0.0576)
Number of children for each child in the household	0.0182***	0.0161***	0.0042*	0.0031	-0.0051*
	(0.0037)	(0.0030)	(0.0025)	(0.0020)	(0.0028)
Number of adults over 17 years	-0.0294***	-0.0263***	-0.0113*	0.0149***	-0.0066
	(0.0087)	(0.0069)	(0.0058)	(0.0047)	(0.0065)
Father's age	-0.0000	-0.0005	0.0022***	-0.0030***	0.0013*
	(0.0010)	(0.0008)	(0.0007)	(0.0006)	(0.0008)
Father has primary education	0.0044	0.0001	0.0193***	-0.0051	-0.0099
	(0.0105)	(0.0084)	(0.0071)	(0.0058)	(0.0078)
Father has secondary education	0.0073	-0.0116	0.0276**	-0.0193**	0.0106
	(0.0162)	(0.0129)	(0.0109)	(0.0089)	(0.0121)
Mother's age	-0.0014	-0.0004	-0.0024***	0.0033***	-0.0019**
	(0.0013)	(0.0010)	(0.0009)	(0.0007)	(0.0010)
Mother has primary education	-0.0931***	-0.0371	-0.0532**	0.0784***	-0.0811***
	(0.0331)	(0.0264)	(0.0223)	(0.0181)	(0.0246)
Mother has secondary education	-0.1727**	-0.0769	-0.1072**	0.1711***	-0.1596***
	(0.0709)	(0.0565)	(0.0477)	(0.0387)	(0.0527)
Protection	0.1947***	0.0823**	0.1145***	0.0481*	-0.0502
	(0.0489)	(0.0390)	(0.0329)	(0.0267)	(0.0364)
Safe drinking water	-0.0463***	-0.1256***	0.0524***	0.0099	0.0169
	(0.0142)	(0.0113)	(0.0095)	(0.0077)	(0.0105)
Electricity	-0.0307	-0.0116	-0.0215	0.0521***	-0.0498***
	(0.0248)	(0.0197)	(0.0167)	(0.0135)	(0.0184)
Sanitation OK	-0.0009	0.0465	-0.0013	-0.0324*	-0.0137
	(0.0359)	(0.0286)	(0.0242)	(0.0196)	(0.0267)
Number of rooms in the household	-0.0036	-0.0009	-0.0014	0.0023	-0.0036
	(0.0033)	(0.0026)	(0.0022)	(0.0018)	(0.0024)
Residual	0.5105*	0.2972	0.3968**	-0.7179***	0.5344**
	(0.2811)	(0.2239)	(0.1893)	(0.1536)	(0.2091)
N	10,064	10,064	10,064	10,064	10,064

Notes: Data are from NCLS 2002. Standard errors in parentheses. The omitted categories are male child, no vaccination, service sector, no schooling, no working dress, source of drinking water is ponds/rivers, no electricity and no sanitary latrine. The model is fitted by first order differencing. Thus, the sample size is reduced to 10,064 instead of 10,065. *** p<0.01,** p<0.05, * p<0.1. p denotes p-value.

ADDENDUM

Chapter 2

p 10 para 2: delete after sentence 1 and read "Firpo et al. (2009) show that the Oaxaca-Blinder decomposition can be extended to any distributional statistic. It involves two steps. The first stage decomposed into explained (*the endowment effect*) and unexplained components (*the discrimination component*). To decompose the gender wage gap into the two components mentioned before, we produce a counterfactual wage distribution, which represents the distribution of wages for male workers in full-time employment if they had the same distribution of characteristics as females. The counterfactual can be obtained by a reweighting method (i.e. reweight the distribution of workers in one group to control for composition). In the second stage, both components are further divided into two stages. The first stage of decomposition requires estimation of the so-called *recentered influence function* (RIF) regression for each distributional statistics which is the core of the Firpo et al. (2009) method. Firpo et al. (2009) show that one can obtain the average effects of explanatory variables on a distributional statistic (for example, wage quantiles) by running regression with original response replaced by the RIF of the statistics. The regression is known as the RIF regression. We run separate RIF regressions for males, females and for the counterfactual. In the second stage, we decompose the gender wage gap into explained (*the endowment effect*) and unexplained components (*the discrimination effect*) at each quantile, as it is usually done with the Oaxaca-Blinder decomposition. Specifically, the predicted wage differential $D_t(\nu)$ measured at quantile ν can be decomposed as follows".

p 10: Add at the end of para 2 : "These quantiles were selected for comparability to quantiles that are most commonly chosen in the previous studies that have used RIF regressions framework, such as Wei and Bo (2007)".

p 11: Add after sentence 1: "We choose to use Wellington (1993) decomposition for its simplicity using the RIF regression procedure".

p 11: delete sentence 2 and read "The rationale behind this extension is to accommodate changes in male-female productivity related characteristics (for example, changes in the characteristics of the workforce) that would cause changes in the gender wage gap over time (Wellington 1993)".

p 11: delete sentence 3 and read "To do this, we subtracted the difference in (log) wages in period $\tau(\tau = 1999)$ from the corresponding difference in period t $(t = 2005)$".

p 12 para 2: Add footnote after sentence 2: "A self-employed is officially defined as a person who operates an enterprise or business on his/her own account or operates jointly with others in the form of a partnership for profit or family gain. Work on the family farm or in family businesses refers to employment without pay or profit in family businesses".

p 17: read para 3 "A large number of explanatory variables that are expected to affect wages are included in the wage regressions. For example, the most commonly employed controls for a worker's productive skill is educational attainment (Albrecht et al. 2009; Chzhen and Mumford 2011). The basic theoretical justification for education is provided by the theory of human capital in which the personal earnings are a positive function of educational attainment. Therefore, this explanatory variable is likely to directly affect productivity and thereby the wage rate. In addition to the education variable, two dummy variables for training (for example, vocational and general) are also included as a supplementary proxy for productivity (Ahmed and Maitra 2010). Following the literature, the wage equation is augmented with variables for different age groups, occupation (Albrecht et al. 2009; Arulampalam et al. 2007; Chzhen and Mumford 2011; Gupta et al. 2006; Kapsos 2008; Sakellariou 2004), industry (Gupta et al. 2006; Kapsos 2008), marital status (Albrecht et al. 2009; Chzhen and Mumford 2011) and region of residence (rural residence is the reference category). The marital status is included as a proxy for maturity and dedication, and therefore productivity of the worker. Since it is believed that marriage diverts a women's time away from the market, the expected effect of marriage on wage rates is negative for women and positive for men, respectively".[108]

p 20 para 1: Add footnote at the end of sentence 1: "As outlined above, to achieve identification of the probit model, the following variables are included: dummy variables for ownership of dwelling and wealth quintiles of the household, number of young children and number of men and women over 65 years of age

[108] Information on variables is detailed in footnotes 11-14 (p.17).

in the household. The exclusion restrictions of these additional variables are also checked. We re-ran regressions but this time we included all of the excluded variables in the wage equation and examined whether these additional variables are statistically significant in the wage equation. The unreported results suggest that these excluded variables are not statistically significant in the wage equations, implying that these variables do not have a direct effect on wage rates".

p 20 para 3: Add after sentence 1: "The counterfactual wage is not reported in the thesis".

p 22: Add footnote at the end of para 2: "The gender wage gap can be analysed in a number of different dimensions, for example, gender wage differentials across religion/ethnicity or race. However, we have included religion affiliation (for example, Muslim dummy variable) as a proxy for ethnicity in both OLS and unconditional quantile regressions and the unreported results suggest that the results did not deviate much from OLS and unconditional quantile regression estimates reported in Tables A2-A5, p.176-187".

p 26: read para 2 "Turning next to decomposition results for the change in the wage gap from 1999 to 2005 (Table 2.7, p.36), we see that inclusion of the selection term does not change the results pertaining to the endowment effects compared to results without the selection term (Table 2.5, p.35). However, we find that women benefitted from a decline in discrimination both at the lower and upper ends of the wage distribution after accounting for sample selectivity bias. Direct comparison of these results with any existing studies is difficult because of the different nature of each study. In particular, markedly different methodology and datasets have been employed and emphasis has been placed on the sample selection bias into full-time employment. However, it is interesting to note that these results are sensitive to the choice of quantiles; hence suggesting these results should be interpreted with care".

Chapter 3

p 43 para 1: Add after sentence 3: "Wage premiums are estimated at the district-area level; where area indicates either a rural or an urban area. Because this is the lowest level of geographical unit reported in the dataset. We have utilised this information to allow for more variation. The alternative procedure could be wage premiums at the aggregate level, such as district or province".

p 44 para 3: delete sentence 3 and read "The LFS and DHS datasets are representative at the district level".

p 45 para 2: Add after sentence 2: "We used nominal wages instead of real wages".

p 48 para 2: Add after sentence 1: "Sample of the descriptive statistics correspond to main regression by gender is reported in Table 3.7".

p 48 para 3: Add after sentence 10: "This is because shortage in the supply of women can lead to increases in their demand in household activities due to gender division of labour in a developing country setting such as Bangladesh. As a consequence, parents may be reluctant to invest in a female child's education. Such gender-differentiated patterns contribute to the gender gap in education. However, it is also possible that shortage of women can increase their wages and may lead parents to invest in a female child's education".

p 48 para 3: delete sentence 11 "This may apply in Bangladesh....."

p 48 para 3: delete sentence 12 "Furthermore, from these female mortality rates....."

p 52 para 2: Add footnote at the end of sentence 4: "We further check the robustness of our results by using probit models. Overall, the results are similar".

p 52: Add at the end of para 2: "We examine the robustness of our results when we restrict ourselves to the sample of non-married individuals aged 18-21. The unreported results suggest that females are more likely to be enrolled in tertiary education than males though the gender effect is not statistically different from zero at conventional levels of significance. The results are unchanged with respect to wage premiums".

p 52: Add footnote at the end of para 2: "We re-estimate our baseline regressions by using sample for 17-21 years. In general, the results reiterate our findings from Table 3.6. In addition, these results do not vary much when we estimated with real wages. To conserve space, we do not report these results in the thesis".

p 54 at the end of para 3: delete "waged" and read "wage".

p 55: Add at the end of para 2: “Since there appears to be some overlap between school children aged 5-17 and adult members of the household aged 15-60. We further estimate regressions by using variables such as number of adult males and females in the household aged 18-60. The results replicate our main findings reported in Table 3.8. To conserve space, we do not report these results in the thesis”.

p 55: Add footnote at the end of para 3: “To control for household income, we included a variable that indicates whether at least one parent is employed. We examine the robustness of our results by including average monthly household income and a dummy variable that indicates if household owns more than one acre of land (Chamarbagwala 2008) as a proxy for household income and wealth. We note that the results are unchanged after inclusion of a proxy of household income and wealth. To conserve space, we do not report these results in the thesis”.

p 55 para 3: delete “sentence 2” and read “It is worth noting that the effect of both mother’s and father’s education has a stronger effect on the tertiary enrolment rate of females compared with males”.

p 56 Section 3.7.1: Add after sentence 2: “We do not know how many age groups we should use for median wage premiums. To make our analysis comparable to mean wage premiums we choose to use similar age groups (25-35, 25-45, 25+)”.

p 57: Add at the end of para 2: “This result suggests that other factors are more crucial in determining female enrolment in tertiary education”.

p 57 para 3: delete sentence 1 and read “Indeed, we find that males and females substantially benefit when their parents are educated, even after controlling for wage premiums”.

Chapter 4

p 76 para 1: Add after sentence 1: “Alternative school measurements, such as enrolment or attainment are not included due to data limitation”.

p 83 para 4: Add after sentence 1: “We assume similar effects of parental education on both school attendance and GAGE”.

p 85: delete Section 4.7.1 and read as Section 4.7.2 in p.90; delete Section 4.7.2 in p.90 and read as Section 4.7.1 in p.85.

p 88: Add footnote at the end of para 2: “We also estimate probit models for school attendance as a baseline model. In general, we find a strong negative effect of working hours on the probability of school attendance. However, this effect weakens when working hours increase. The results regarding parental education remain in most cases. However, it is worth noting that probit estimates are reasonably smaller than IV probit estimates, suggesting that standard probit estimates are biased downward if the endogeneity of working hours is not taken into consideration. To conserve space, we do not report these results in the thesis”.

p 90: Add footnote at the end of para 2: “Tobit models for GAGE are also estimated as a benchmark. The results replicate our main findings that are included in the thesis; however, Tobit estimates are larger than IV Tobit estimates. To conserve space, we do not report these results in the thesis”.

p 96 para 1: read after sentence 6 “A similar effect of parents’ education on child’s school attendance and grade attainment is also found in urban areas. It may be possible that the income effect from a child working in paid employment in urban areas counteracts and weakens the effect of parental education on a child’s schooling”.

p 97: read para 2 “We find that parental characteristics, especially their level of education, affect the work-schooling trade-off across genders. The mother’s level of education shows a greater positive effect on the school attendance of male and female children but has a differentially higher effect on the female child. A similar effect is found for age-adjusted grade attainment. On the other hand, our results confirm that if the mother’s education is higher than that of the father’s, it increases school attendance; but interestingly decreases grade attainment. This is in contrast with our conjecture. One possible explanation is that even if the mother is highly educated than that of the father she cannot participate in household decision-making on equal terms and hence, children are more likely to be exploited as child labour. We also find that the father’s education (as measured by the highest grade attained) is positively correlated with daughter’s school attendance relative to sons, even after controlling for the mother's education. The interpretation of these results is unchanged when the model is estimated with the full set of covariates. The magnitude of the impact of the father's education is larger in age-adjusted grade attainment, but the effect of the father's education on the grade attainment is not significantly larger for a female child with and without the full set of covariates. After correcting for potential sources of selection bias, the qualitative results remain for school attendance but moderately change for age-adjusted grade attainment equations".

p 102: Add notes below Table 4.3: "Work refers to full-time employment, while school implies when children attend school full-time and avoid child labour".

Chapter 5

p 120 para 1: delete "for a policy perspective" and read "from a policy perspective".

p 126: Add at the end of para 1: "This figure is comparable to Bangladesh LFS 1999".

p 126 para 1: Add footnote after sentence 1: "Regarding the definition of child labour, we follow NCLS 2002. Child labour as referred to in the NCLS consists of children aged 5-17 who are economically active except (i) those who are under five years old and (ii) those between 12-14 years old who spend less than 14 hours a week on their jobs, unless their activities or occupations are hazardous by nature or circumstance. Added to this are 15-17 year old children in the worst form of child Labour (i.e. work 43 hours or more per week). Ray (2004) also followed similar definition in his study on child labour".

p 126 para 1: Add after sentence 3: "Furthermore, we cannot include any information related to precise nature of child's work (for instance, whether children involve operating any machine or heavy manual job) due to data limitation".

p 126 para 2: Add after sentence 3: "The choice of these two health indices is mainly based on questions available in NCLS 2002. These are the typical questions used for identifying morbidity status for working children in the developing country (see, for example, the Vietnam Living Standards Survey, the Cambodia Child Labour Survey)".

p 132 para 2: delete sentence 2 and read "Taking conditional expectations given ℓ_n in Equation (5.6) gives us:

$$E(\mathcal{H}_n|\ell_n) = \pi_0 + E(Q_n|\ell_n)\pi_1 + \mathcal{F}(\ell_n) \qquad (5.6)'$$

The difference between Equations (5.6) and (5.6)' yields

$$\mathcal{H}_n - E(\mathcal{H}_n|\ell_n) = \big(Q_n - E(Q_n|\ell_n)\big)\pi_1 + \epsilon_n; \;\; n = 1,\ldots,\mathcal{N} \qquad (5.7)"$$

p 134 para 1: Add at the end of sentence 5: "(O'Donnell et al. 2005)".

p 134 para 1: Add at the end of sentence 9: "(O'Donnell et al. 2005)".

p135: Add at the end of footnote 88: "Informal school refers to informal education activities (for instance, family education and others) as indicated by NCLS 2002 reported in Chapter 4 (p.101)".

p 140 Section 5.6: Add footnote at the end of further analysis: "It is difficult to account for omitted variable bias in our context as we do not have access to panel data. However, to take into account omitted variable bias, we conducted a small investigation by using household fixed effects. We have performed regressions using the fixed effect logit models with the number of hours worked by the child. Insights from fixed effect logit model indicate that controlling for unobserved heterogeneity does not affect our previous conclusion: we obtain a significantly positive coefficient of child labour hours on the probability of reporting injury or illness. For example, the point estimates for reporting any injury or illness are 2.337 (z = 30.79); the corresponding values are 1.302 (z = 13.92) for tiredness/exhaustion; 1.478 (z = 15.81) for body injuries; 1.092 (z = 9.66) for backache; 2.340 (z = 18.51) for other health problems. To conserve space, we do not report these results in the thesis".

p 141 Section 5.6.1: Add footnote at the end of para 2: “These results are robust when we included regional dummy variables (Chittagong, Rajshahi, Khulna, Barisal, Sylhet and Noakhali. The reference category is Dhaka)”.