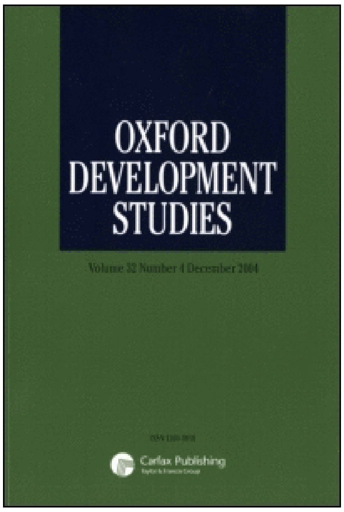

Oxford Development Studies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713439972

Gender Wage Discrimination in Rural and Urban Labour Markets of Bangladesh

Salma Ahmed[a]; Pushkar Maitra[a]
[a] Department of Economics, Monash University, Clayton Campus, VIC, Australia

Online publication date: 19 February 2010

To cite this Article Ahmed, Salma and Maitra, Pushkar(2010) 'Gender Wage Discrimination in Rural and Urban Labour Markets of Bangladesh', Oxford Development Studies, 38: 1, 83 — 112

To link to this Article: DOI: 10.1080/13600810903551611

URL: http://dx.doi.org/10.1080/13600810903551611

Oxford Development Studies,
Vol. 38, No. 1, March 2010

Routledge
Taylor & Francis Group

Gender Wage Discrimination in Rural and Urban Labour Markets of Bangladesh

SALMA AHMED & PUSHKAR MAITRA

ABSTRACT *Female wages in Bangladesh are significantly lower than male wages. This paper quantifies the extent to which discrimination can explain this gender wage gap across the rural and urban labour markets of Bangladesh, using unit record data from the 1999–2000 Labour Force Survey. The gender wage differential is decomposed into a component that can be explained by differences in productive characteristics and a component not explained by observable productive differences, which is attributed to discrimination. An attempt is also made to improve on the standard methodology by implementing a wage-gap decomposition method that accounts for selectivity bias, on top of the usual "explained" and "unexplained" components. Analytical results from this paper show that gender wage differentials are considerably larger in urban areas than in rural areas and a significant portion of this wage differential can be attributed to discrimination against women. The results also show that selectivity bias is an important component of total discrimination.*

1. Introduction

It is now fairly well established that women lag behind men in many domains in developing countries. Gender differences are noticeable in several spheres. For example, women have less access to and control over resources, few opportunities are made available to them at the workplace and they are significantly under-represented in the political sphere. Although women incur the direct costs of these inequalities, the costs are eventually borne by the society as a whole. The nature and extent of this inequality varies considerably across countries, and is related to the basis and structure of the economic system of a nation. One of the most visible examples in this respect is inequality in wages based on gender. Women on average earn less than men for similar work and the gap varies across nations. In developing regions, such as Southern Asia and Sub-Saharan Africa, on average women earn around 50% of men's income (controlling for occupational differences), and this gender wage gap declines substantially, though does not vanish, when we look at developed countries: for example in North America, women on average earn 63% of men's income.

Salma Ahmed, Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. Email: salma.ahmed@buseco.monash.edu.au. Pushkar Maitra (corresponding author), Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. Email: pushkar.maitra@buseco.monash.edu.au

We would like to thank Anu Rammohan, Rebecca Valenzuela, Mark Bryan, Paul Miller and an anonymous referee for their comments and suggestions. The usual caveat applies.

ISSN 1360-0818 print/ISSN 1469-9966 online/10/010083-30

DOI: 10.1080/13600810903551611

Wage inequalities between males and females exist in almost every country and one can think of two main reasons for this. First, males and females might choose to accumulate different levels of productive skills. This may be motivated by, among other things, culture, geographic proximity, historical reasons, etc. Second, even in the presence of equal endowments of productive skills, wage inequality may persist if employers reward productive skills differently depending on the gender of the worker. Such a potential cause of wage inequalities is usually attributed to discrimination at the workplace (Becker, 1957; Phelps, 1972; Arrow, 1972). Becker's work provided the background for subsequent work by Oaxaca (1973), Blinder (1973), Reimers (1983), Neumark (1988) and Cotton (1988). The methodology developed in these papers has been used to examine gender wage discrimination in a number of different (though primarily developed) countries.

A small number of studies now exists on countries in Latin America (Psacharpoulos & Tzannatos, 1992), Africa (Knight & Sabot, 1982; Glick & Sahn, 1997; Appleton *et al.*, 1999) and the transitional economies of Eastern and Central Europe (Brainerd, 1998; Reilly & Newell, 2000; Jurajda, 2001; Adamchik & Bedi, 2003). All these studies identify active discrimination against women in the labour market. Horton (1996), in a seven-country study of women in East Asian labour markets, found that differences in returns to male and female characteristics account for at least half the gap between male and female earnings. Recent studies on South Asia have yielded similar results. Jacob (2006) explored the changes in the wage gap between caste and gender groups in India between 1983 and 1999–2000 using a nationally representative data set. He found that about 55% of the wage gap between men and women in 1999 cannot be explained by differences in productive characteristics and endowments. Akter (2005), in a study of the rural labour market in Bangladesh, found that 70% of the total wage gap is due to within-job discrimination. Kapsos (2008) found that women in non-agricultural sectors in Bangladesh earn 21% less per hour than men. The few other studies that exist in Bangladesh have focused almost exclusively on the urban manufacturing sector and report that differences in the wage rates of men and women cannot be accounted for by differences in productivity-related characteristics, implying that discrimination against women may play a role (Majumder & Zohir, 1993; Majumder & Mahmud, 1994; Zohir, 1998).

The labour market in Bangladesh is gender segregated, with the bulk of women's work taking place in non-market activities in the home or the informal sector. Those in the formal sector (public and private) are generally employed in "female-intensive industries" (e.g. the ready-made garments (RMG) sector, shrimp processing and pharmaceuticals). Additionally, upward mobility of female labourers is limited. Women in paid employment in both rural and urban areas often receive lower wages than men, even after controlling for types of employment, status of employment, occupation and hours of work. For example, in the urban regions female wage rates were 50% that of males in 1995–96, declining further to 46% in 1999–2000. In rural regions the ratio remained constant at 44% over the same period.

We contribute to the understanding of gender inequality in the Bangladesh labour market in the following ways. First, we consider both rural and urban labour markets in order to understand the overall inequality in wages. Second, we use individual-level unit record data collected for the whole country for this purpose, which has not been done previously.[1] Our empirical analysis is conducted using the Labour Force Survey (LFS) data for 1999–2000, which permits us to go beyond previous analyses of the gender wage gap in Bangladesh. The LFS is a nationally representative random sample and contains detailed information on demographic, social and economic characteristics by gender and location. Finally, standard modelling techniques used in most existing studies on the gender wage gap in Bangladesh fail to capture the potential effects of selectivity bias in the male–female wage gap. We attempt to improve on the standard methodology by implementing a wage-gap decomposition method that accounts for selectivity bias, on top of the usual methods that decompose the gender wage differential into a component that can be explained by differences in productive characteristics and a component unexplained by observable productive differences (attributed to discrimination).

2. Empirical Specification

The quantifiable measures of discrimination against women generally focus extensively on the magnitude of the wage gap between males and females. The most commonly used technique is the Oaxaca–Blinder decomposition method (Oaxaca, 1973; Blinder, 1973). They decompose the wage differential into a component explained by differences in personal characteristics of workers that affected their productivity and a component unexplained by observable productive differences which they attribute to discrimination. It is, however, important to note that the entire unexplained portion cannot be attributed to discrimination alone as it might also capture the impact of model misspecification, omitted variables and measurement error. This latter issue might mean that the different outcomes for men and women may be the result of differences of some unobserved variables (e.g. motivation, congeniality, ability to work in a group, sensitivity, etc.) that are not captured by variables included in the analysis. Unfortunately, the available data do not allow us to quantify this effect.

We start by estimating separate (log) hourly wage equations for males (m) and females (f).[2] The wage equations have the following form:

$$\ln Y_{ij} = \alpha_j + \beta_j X_{ij} + \varepsilon_{ij}; \quad i = 1, \ldots, n; \;\; j = m, f. \tag{1}$$

$\ln Y_{ij}$ is the natural log of hourly wages, α_j is an intercept term for gender group $j; j = m, f$; X_{ij} is a vector of characteristics for individual i who belongs to gender category j, β_j is a vector of coefficients to be estimated, and ε_{ij} is the error term with zero mean and constant variance. Equation (1) is estimated using ordinary least squares (OLS), separately for males and females.

Define D as the difference in the expected values of male and female wages obtained by estimating equation (1) separately for males and females. We can then write:

$$\begin{aligned} D = \overline{\ln Y_m} - \overline{\ln Y_f} &= (\hat{\beta}_m \bar{X}_m - \hat{\beta}_f \bar{X}_f) + (\hat{\alpha}_m - \hat{\alpha}_f) \\ &= \hat{\beta}_m(\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f) \end{aligned} \tag{2}$$

where $\hat{\beta}_j$ and $\hat{\alpha}_j$ denote the corresponding estimated values of β_j and α_j, respectively. The first term on the right-hand side $[\hat{\beta}_m(\bar{X}_m - \bar{X}_f)]$ is the explained portion, which is the portion that explains the gap due to observable characteristics at the mean, evaluated by the male wage equation. The second term $[(\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f]$ is the unexplained portion. It is the difference in the return to each wage determinant received by males and females, evaluated at the mean set of women's characteristics. The third term $[\hat{\alpha}_m - \hat{\alpha}_f]$ is the difference between the constants. The latter two terms arise from differences in the coefficients in the wage equations (both intercepts and slopes) for men and women and are usually interpreted as measures of labour market discrimination after adjusting for differences in observable characteristics (in other words, the adjusted wage gap).

An alternative way of writing equation (2) is to take the female wage structure as the non-discriminatory norm. Then the decomposition becomes:

$$D = \overline{\ln Y_m} - \overline{\ln Y_f} = \hat{\beta}_f(\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_m + (\hat{\alpha}_m - \hat{\alpha}_f). \quad (3)$$

Comparison of equations (2) and (3) leads us to an important question: What is the non-discriminatory wage structure to be used for comparison purposes? The original method used both male and female wage structures as the non-discriminatory norm. This creates an "index number" problem because the choice of male and female wage structures as the competitive standard does not yield the same estimate for the discrimination component. Further, the resulting levels of discrimination provide a range within which the actual level of discrimination falls. Reimers (1983) hypothesized that the correct procedure is instead to take an average of both male and female wage structures. Cotton (1988) and Neumark (1988) suggested improving on 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.

Cotton (1988) argued that males are paid a premium for their productivity characteristics as a result of nepotism; in addition, females have their characteristics undervalued as a result of discrimination. He has, therefore, argued for estimating a non-discriminatory wage structure through a linear combination of the male and female wage structures. The resulting decomposition can be written as follows:

$$D = \overline{\ln Y_m} - \overline{\ln Y_f} = \beta^*(\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \beta^*)\bar{X}_m + (\beta^* - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f), \quad (4)$$

where β^* is a non-discriminatory wage structure that is common to both men and women in the economy. The first term $[\beta^*(\bar{X}_m - \bar{X}_f)]$ of the decomposition is the explained component. The second term $[(\hat{\beta}_m - \beta^*)\bar{X}_m]$ and the third term $[(\beta^* - \hat{\beta}_f)\bar{X}_f]$ represent the male treatment advantage and the female treatment disadvantage, respectively. The sum of the last three terms of equation (4) is considered to be an indicator of the extent of discrimination.

Now β^* is unobservable and is defined as

$$\beta^* = \rho_m\hat{\beta}_m + \rho_f\hat{\beta}_f, \quad (5)$$

where ρ_m is the proportion of the male workforce and ρ_f is the proportion of the female workforce in the samples, and $\hat{\beta}_m$ and $\hat{\beta}_f$ are the estimated parameters from the male and female wage regressions.

2.1 Accounting for Sample Selection

One potential problem with the decomposition approaches discussed above is that they do not account for sample selection bias. The wage equations are applied to a sample of employed men and women, whose selection criteria are not random because we do not have wage data for those who are unemployed. The non-inclusion of those not employed might involve the omission of unobservable factors that influence wages (e.g. motivation). If these unobserved factors are correlated with observable factors in the wage equations, then a failure to take account of this will lead to biased wage equation coefficients (Vella, 1998).

In this paper we revisit the "sample selection" problem in terms of employment status for males and females and the likely consequences of this selection issue on the gender wage gap.[3] We use the Heckman (1979) two-step sample selection model to account for selection bias. The first step involves the estimation of a selection equation[4] and the second step involves the estimation of a wage equation, conditional on employment. In the selection equation an individual i is assumed to choose his or her employment status according to a probit model:

$$I_{ij} = \gamma Z_{ij} + \mu_{ij}; \quad i = 1, \ldots, n; \;\; j = m,f, \tag{6}$$

where I_i is a dummy variable denoting employment status. Z is a vector of determinants of employment and γ is a vector of coefficients to be estimated and $\mu_i \sim IIDN(0, \mathbf{1})$. We estimate separate employment status regressions for males and females. Having estimated equation (6), we compute the inverse Mills ratio,

$$\lambda_i = \frac{\phi(\gamma Z_i)}{1 - \Phi(\gamma Z_i)},$$

which is included as an additional explanatory variable in wage equations given by (1). The expanded wage equations can be written as:

$$\ln Y_{ij} = \alpha_j + \beta_j X_{ij} + \theta_j \lambda_{ij} + \varepsilon_{ij}; \quad i = 1, \ldots, n; \;\; j = m,f, \tag{7}$$

θ_j denotes the covariance between the errors in the selection equation and the wage equation (one for each gender group j).[5] Equation (7) takes into account the correlation between ε_{ij} and μ_{ij}.[6]

We can now compute the extended gender wage gap as:

$$D = \overline{\ln Y_m} - \overline{\ln Y_f} = \hat{\beta}_m(\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f) + (\hat{\theta}_m \bar{\lambda}_m - \hat{\theta}_f \bar{\lambda}_f), \tag{8}$$

The observed wage differential $(\overline{\ln Y_m} - \overline{\ln Y_f})$ is now the sum of the following components: the contribution of endowment differences or the explained portion $[\hat{\beta}_m(\bar{X}_m - \bar{X}_f)]$, the unexplained portion that is attributed to labour market discrimination $[(\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f)]$ and the contribution of differences in the average selectivity bias $[\hat{\theta}_m \bar{\lambda}_m - \hat{\theta}_f \bar{\lambda}_f]$.[7]

The incidence of sample selection bias indicates that the observed mean log-wage differential will differ from the unobserved differential in terms of wage offers (Kidd & Viney, 1991). Therefore, it is convenient to rewrite equation (8) in the following form:

$$(\overline{\ln Y_m} - \overline{\ln Y_f}) + (\hat{\theta}_f \bar{\lambda}_f - \hat{\theta}_m \bar{\lambda}_m) = \hat{\beta}_m(\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f). \tag{9}$$

Here the left-hand side provides the measure of differences in the offered wage (the sum of the difference in the observed mean wages and the difference in the average selectivity bias).[8] The only difference between equation (8) and equation (9) is that equation (9) presents a decomposition of the selectivity adjusted wage difference as opposed to a decomposition of the observed wage difference in equation (8).

We can rewrite equation (9) by taking the female wage structure as the non-discriminatory norm:

$$(\overline{\ln Y_m} - \overline{\ln Y_f}) + (\hat{\theta}_f \bar{\lambda}_f - \hat{\theta}_m \bar{\lambda}_m) = \hat{\beta}_f (\bar{X}_m - \bar{X}_f) + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_m + (\hat{\alpha}_m - \hat{\alpha}_f). \quad (10)$$

The information derived from equations (9) and (10) will allow us to decompose differences in observed and offered wages into the "explained" and "discriminatory" components.

3. The Data

The data set that we use for empirical estimation is obtained from the LFS 1999–2000 in the form of unit record files. This data set has information on 34 998 individuals aged 10 years and above, and provides information about whether an individual is employed, unemployed (but actively looking for work) and out of the labour force (e.g. retired, disabled or a full-time student). The estimating sample is restricted to 18 979 employed and unemployed individuals.[9] In this sample, 12 394 (66%) are males and 6585 (34%) are females. Out of this sample, 5951 males and 2368 females reside in urban areas, while 6443 males and 4217 females reside in rural areas. Of this sample, 18 237 are employed, consisting of 11 919 males and 6318 females. Of the sample of employed individuals, 7875 individuals reside in urban areas while 10 362 individuals reside in rural areas; 72% of the urban sample and 60% of the rural sample are male.

In the wage regressions the dependent variable is the log of hourly wages. We focus on four classes of workers, the self-employed, wage employees, casual workers[10] and unpaid family workers,[11] to obtain as complete a population coverage as possible.[12] While the survey provides monthly wages for the self-employed and the wage-employed population, it provides daily wages for casual workers. We convert the daily wage income of casual workers to monthly income assuming that they work 6 days in a week. However, the survey does not provide information on the wage of unpaid family workers. We replace the unobserved (missing) information on wages of unpaid family workers by the mean monthly wage of wage labourers. Further, the survey collected information on the usual hours of work per week but not the number of weeks worked during a month. The monthly hours of work is computed by multiplying the usual hours of work per week by 4.3. The hourly wages is then obtained by dividing monthly wages by the imputed monthly hours of work.

The set of explanatory variables[13] includes variables that measure individual productivity (remember this is unobserved and we use the individual's educational attainment, age and the square of age[14] as proxies), log of monthly hours worked,[15] the number of months unemployed,[16] dummies for being married, job status, training and skill.[17] We also control for occupation and industry of employment.

3.1 Identification of the Selection Equation

Identification of the selection (employment) equation requires the inclusion of (at least) one variable in the selection equation, which is not included in the wage equation. Although the identification comes from the non-linearity of the inverse Mills ratio, the $X = Z$ case (see equations (6) and (7)) often results in substantial collinearity between the predicted inverse Mills ratio term and the remaining covariates in the outcome (wage) equation. This collinearity will, as always, lead to large standard errors. The best solution to this problem is to incorporate a set of variables that belong to the selection equation but not in the wage equation. Admittedly, it is very difficult to identify appropriate exclusion restrictions, and there is no particular method to test whether the exclusion restrictions are valid. Typically it depends on intuition and on the data at hand. There are some formal tests to do but they are not exclusive and should be used only when we justify the variables to be excluded.[18]

We include asset ownership (an indirect proxy for non-labour income), home ownership status to proxy wealth, and some family status variables as identifying variables.[19] These variables are assumed to affect the probability of employment 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 beyond the control of any individual). For example, as a source of non-labour income the ownership of various types of asset could reduce the probability of employment by raising the shadow value of a person's time in non-market activities. An increase in the number of working age adults and elderly in the household (i.e. the number of other men (*NMEN*) and women (*NFEM*) aged between 15 and 64 years) is likely to affect significantly the likelihood of working, though the direction of the effect is not clear.

Next we report descriptive statistics for the urban sample and the rural sample of males and females (see Table 1). We also present *t*-tests for gender differences. The following results are worth noting: in both rural and urban areas more than 90% of males and females are likely to be employed (*PART*); while there is significant difference between males' and females' employment status in the urban sample, the difference in employment status is not statistically significant by gender for the rural sample. Females are on average younger (*AGE*) and are generally less educated than males. Gaps in educational attainment between males and females are significant at all levels of education except for primary education (*PRIM*) in urban areas. A higher proportion of females are married (*MARRIED*) compared with males in rural areas. This is quite a common scenario in rural Bangladesh where adult male members of the household migrate to urban areas but are unable to bring their family with them or owing to adverse economic conditions married female members are forced to accept wage employment outside the home. Women are more likely to own assets such as sewing machines, shallow machines and tractors (*ASST03*) compared with men, though men are more likely to own larger assets such as shops or businesses (*ASST01*).

In the urban sample, the log wage gap at the mean is 0.6703. That is, on average, women earn 95% less per hour than men.[20] The difference is statistically significant at the 1% significance level. In the rural sample, the mean log wage difference for men and women is 0.3551 at the 1% significance level. In percentage terms, on average, women earn 43% less than men in rural areas.[21] Finally, women are predominantly employed in agriculture (*OCCUP_AGRI*) and services (*OCCUP_SER*), whereas men dominate sales (*OCCUP_SALES*) and production (*OCCUP_PROD*) related occupations.

In Figures 1 and 2, we present the kernel density estimates of the log of hourly wages for males and females for the urban sample (Figure 1) and the rural sample (Figure 2). In general, the peak of the distribution of male (log) hourly wages lies to the right of the distribution of the female (log) hourly wages. The distribution of hourly wages looks different for males and females, and using a Kolmogorov-Smirnov (K-S) test for the

Table 1. Mean values of variables used in estimation

Variable	Urban sample: Males	Urban sample: Females	Urban sample: *t*-test for difference	Rural sample: Males	Rural sample: Females	Rural sample: *t*-test for difference
LNHRSINC	2.6206 (0.8706)	1.9503 (0.8322)	− 31.13***	2.2744 (0.8292)	1.9193 (0.6489)	− 23.16***
LN_MHRS	5.2805 (0.3286)	4.7458 (0.5782)	− 51.51***	5.2220 (0.3363)	4.5172 (0.4903)	− 86.77***
AGE	35.6194 (12.3258)	31.4759 (11.236)	− 14.18***	36.2339 (14.3398)	33.0868 (12.1557)	− 11.75***
AGESQ	14.2064 (9.6394)	11.1693 (8.4092)	− 13.43***	15.1849 (11.836)	12.4246 (9.362)	− 12.76***
UNEMPS	0.0347 (0.9757)	0.0045 (0.0948)	− 1.45	0.0155 (0.6582)	0.0154 (0.5871)	− 0.01
PRIM	0.2106 (0.4077)	0.2226 (0.4160)	1.2	0.2696 (0.4438)	0.2407 (0.4276)	− 3.34***
SECOND	0.2526 (0.4345)	0.1943 (0.3957)	− 5.66***	0.1785 (0.3830)	0.1077 (0.3100)	− 10.05***
PSECON	0.1617 (0.3682)	0.0938 (0.2915)	− 8.03***	0.0695 (0.2544)	0.0313 (0.1742)	− 8.54***
GRAD	0.1207 (0.3257)	0.0448 (0.2068)	− 10.52***	0.0244 (0.1542)	0.0055 (0.0737)	− 7.43***
TRA_VOC	0.0097 (0.0982)	0.0068 (0.0819)	− 1.27	0.0069 (0.0826)	0.0037 (0.0604)	− 2.14**
TRA_TECH	0.0248 (0.1554)	0.0077 (0.0872)	− 4.89***	0.0043 (0.0655)	0.0002 (0.0156)	− 3.9***
TRA_GEN	0.0207 (0.1424)	0.0153 (0.1228)	− 1.57	0.0137 (0.1164)	0.0085 (0.0920)	− 2.4**
SKIL	0.3405 (0.4739)	0.1175 (0.3221)	− 20.4***	0.3673 (0.4821)	0.0886 (0.2842)	− 33.4***
MARRIED	0.7846 (0.4112)	0.7744 (0.4181)	− 0.98	0.7796 (0.4146)	0.8621 (0.3448)	10.6***
SELF_EMPD	0.4609 (0.4985)	0.2017 (0.4014)	− 21.9***	0.5047 (0.5000)	0.0952 (0.2935)	− 47.4***
W_EMPD	0.3417 (0.4743)	0.3489 (0.4767)	0.61	0.1015 (0.3020)	0.0410 (0.1983)	− 11.3***
C_WKR	0.1369 (0.3438)	0.0743 (0.2623)	7.74***	0.2787 (0.4484)	0.0737 (0.2613)	− 26.5***
PART	0.9501 (0.2178)	0.9379 (0.2413)	− 2.23**	0.9724 (0.1639)	0.9715 (0.1663)	− 0.25
ASST01	0.2242 (0.4171)	0.1765 (0.3813)	− 4.82***	0.1083 (0.3108)	0.092 (0.2891)	− 2.73***
ASST02	0.0413 (0.1991)	0.0346 (0.1829)	− 1.42	0.047 (0.2117)	0.0424 (0.2016)	− 1.11
ASST03	0.5579 (1.5744)	0.5954 (1.6198)	0.97	0.8237 (1.8546)	0.9134 (1.9319)	2.4**
HRENT	0.3917 (0.4882)	0.3454 (0.4756)	− 3.93***	0.011 (0.1044)	0.0095 (0.0969)	− 0.76
HFREE	0.0363 (0.1870)	0.0401 (0.1963)	0.83	0.0194 (0.1379)	0.0204 (0.1414)	0.36
NMEN	1.6801 (0.9828)	1.2610 (0.8472)	− 18.2***	1.6485 (0.9303)	1.3353 (0.8430)	− 17.6***
NFEM	0.5135 (0.6647)	1.2829 (0.5822)	49.03***	0.91 (0.7280)	1.3427 (0.6472)	31.3***
OLD01	0.0422 (0.2010)	0.0418 (0.2002)	− 0.08	0.0897 (0.2858)	0.0714 (0.2575)	− 3.37***
OLD02	0.0049 (0.0696)	0.0152 (0.1224)	4.83***	0.0107 (0.1029)	0.0225 (0.1484)	4.85***
MALE_HHEAD	0.9178 (0.2746)	0.8045 (0.3967)	− 14.9***	0.9393 (0.2388)	0.8802 (0.3247)	− 10.8***
FEMALE_HHEAD	0.0155 (0.1234)	0.1136 (0.3174)	20.31***	0.0166 (0.1278)	0.0669 (0.2498)	13.7***
OCCUP_PROF	0.0212 (0.1441)	0.0041 (0.0635)	− 5.41***	0.0011 (0.0334)	0.0002 (0.0156)	− 1.57

OCCUP_CLERIC	0.0875 (0.2827)	0.0257 (0.1582)	−9.74***	0.0193 (0.1376)	0.0037 (0.0604)	−6.86***
OCCUP_SALES	0.2890 (0.4533)	0.0459 (0.2094)	−24.3***	0.1346 (0.3413)	0.0188 (0.1358)	−20.7***
OCCUP_SER	0.0624 (0.2420)	0.1607 (0.3674)	13.87***	0.0233 (0.1509)	0.0488 (0.2155)	7.08***
OCCUP_AGRI	0.1197 (0.3247)	0.3661 (0.4818)	26.18***	0.6428 (0.4792)	0.8511 (0.3560)	23.9***
OCCUP_PROD	0.3601 (0.4801)	0.3206 (0.4668)	−3.31***	0.1481 (0.3553)	0.0659 (0.2481)	−12.9***
INDS_MANU	0.2045 (0.4033)	0.2188 (0.4135)	1.41	0.0838 (0.2771)	0.0630 (0.2429)	−3.92***
INDS_WHRET	0.2757 (0.4469)	0.0486 (0.2151)	−22.9***	0.1326 (0.3392)	0.0200 (0.1401)	−20.2***
INDS_HOTEL	0.0269 (0.1618)	0.0081 (0.0897)	−5.17***	0.0120 (0.1088)	0.0012 (0.0349)	−6.12***
INDS_COMMU	0.1498 (0.3569)	0.0086 (0.0921)	−18.4***	0.0528 (0.2237)	0.0017 (0.0413)	−14.5***
INDS_FINST	0.0200 (0.1400)	0.0041 (0.0635)	−5.16***	0.0029 (0.0535)	0.0005 (0.0221)	−2.71***
INDS_REALEST	0.0090 (0.0946)	0.0041 (0.0635)	−2.28**	0.0019 (0.0437)	0.0005 (0.0221)	−1.93*
INDS_PUBADMN	0.0591 (0.2358)	0.0225 (0.1484)	−6.8**	0.0115 (0.1066)	0.0007 (0.0271)	−6.33***
INDS_HLTH	0.1028 (0.3037)	0.2688 (0.4434)	19***	0.0327 (0.1779)	0.0552 (0.2283)	5.6**
INDS_EDU	0.0264 (0.1602)	0.0531 (0.2243)	5.92***	0.0239 (0.1529)	0.0071 (0.0838)	−6.45***
λ	0.0785 (0.1316)	0.0931 (0.1437)	4.17***	0.0504 (0.0851)	0.046 (0.1007)	3.38***

Note: See Table A1 for definition of variables.
*Significant at 10%; **significant at 5%; ***significant at 1%. Standard deviations reported in parentheses.
Data source: Unit record data, LFS 1999–2000.

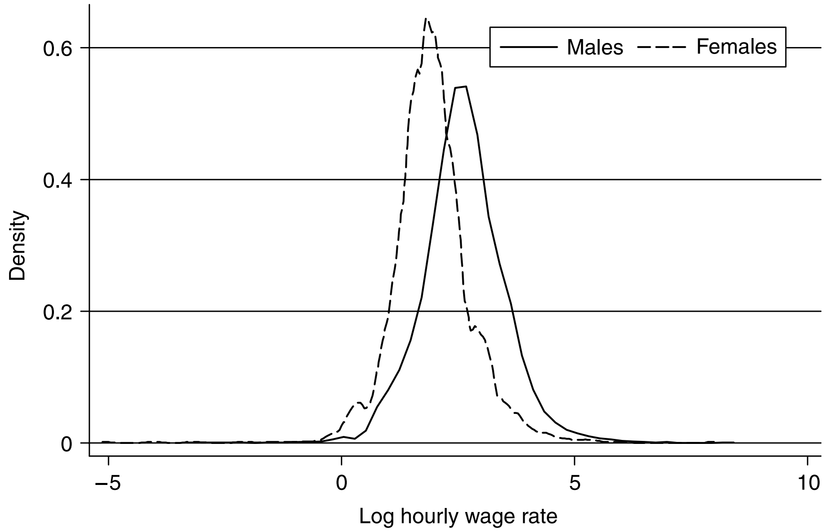

Figure 1. The distribution of hourly wages for males and females. Urban sample (1999–2000).

equality of distributions we see that indeed they are: using the K-S test, the null hypothesis of equality of distribution is always rejected (in every case the *p*-value is 0.000).

4. Empirical Results

Given that we argue (and the regression results presented agree) that ignoring selection bias could result in biased estimates, we focus our discussion on the selectivity corrected regression results. We start with a discussion of the selection (employment) equation and then turn to the selectivity corrected log wage regressions for both

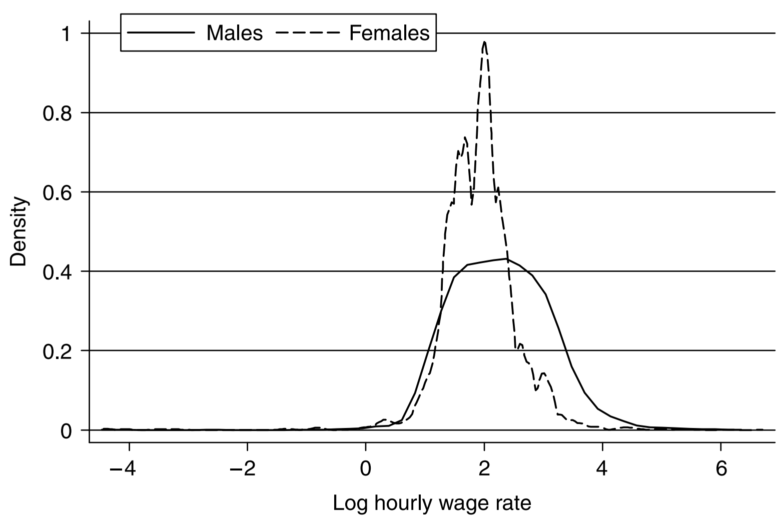

Figure 2. The distribution of hourly wages for males and females. Rural sample (1999–2000).

males and females. The uncorrected log wage regression results are presented in the Appendix.

4.1 Regression Results

4.1.1 Results from the probit regression (selection). The selection (employment) equation is estimated using a probit regression model, where the dependent variable *PART* takes the value of one if the worker is employed and zero if the worker is unemployed. Separate regressions are estimated for males and females.

Age (*AGE*) has a positive and statistically significant effect on the employment probabilities of males and females across all samples. Being married (*MARRIED*) has a stronger effect on the probability of employment of women compared with that of men in the rural sample (the marginal effects presented in Table 2 show that relative to being unmarried, a married male has a 1 percentage point higher probability of being employed; the corresponding figure for married females is 3.5), but not in the urban sample. This might be a reflection of household budget constraints in rural areas that force married women to enter the labour force in order to support their families. Interestingly, being married has a similar effect on the probability of employment for urban males and females (5.57 percentage points for males compared with 5.51 percentage points for females). The four educational attainment dummies (i.e. *PRIM*, *SECOND*, *PSECON*, *GRAD*) are always negative and statistically significant. For the urban sample, males with primary schooling (*PRIM*) have a 2 percentage point lower probability of being employed in the labour market compared with men with no schooling (the omitted category). This probability increases to 4.65 percentage points if the highest education attained is level 6–10 (*SECOND*) and rises further to 9.54 percentage points if the highest education attained is *SSC*/*HSC*[22] or equivalent (*PSECON*) and to 9.81 percentage points if the individual has attained a graduate or higher degree (*GRAD*). The corresponding probabilities for the sample of urban females are 3.39, 6.66, 10.70 and 10.05. For rural males the corresponding probabilities are 0.65, 1.19, 4.12 and 4.51, whereas for rural females the corresponding probabilities are 0.54, 1.47, 4.85 and 9.71.

The effects of wealth on the probability of employment are measured by non-labour income such as ownership of assets. Asset ownership (for example, if the individual has a shop/business (*ASST01*) and if the individual has a rickshaw/van/pushcart/boat (*ASST02*) in the urban sample and if the individual has a shop/business (*ASST01*) in the rural sample) has a statistically significant effect on probability of male employment.[23] However, the marginal effects of asset ownership are not statistically significant in either sample for women. In the urban sample, for both males and females, the probability of employment is higher for those who live in a rented house (*HRENT*) compared with those who own the house they reside in (the omitted category). This result is surprisingly similar to that typically found for developed countries. For example, Kidd & Viney (1991) using data from Australia reported that the probability of employment in the labour market increases for women if they live in rented accommodation (*HRENT*). On the other hand, it is surprising that residing in rented accommodation has a negative and statistically significant impact on a male's employment in the rural sample. In the case of men, it might be capturing the fact that in the rural sample those individuals working are also those more likely to own a home. Págán (2002) has identified a similar association in rural Guatemala,

Table 2. Estimates of selection equations: LFS 1999–2000

	Urban sample				Rural sample			
	Males		Females		Males		Females	
Variable	Coef.	ME[a]	Coef.	ME[a]	Coef.	ME[a]	Coef.	ME[a]
AGE	0.0716 (0.0178)	0.0023***	0.1018 (0.0250)	0.0050***	0.0762 (0.0255)	0.0012***	0.1198 (0.0180)	0.0022***
AGESQ	−0.0007 (0.0002)	0.0000***	−0.0008 (0.0004)	0.0000**	−0.0004 (0.0004)	0.0000	−0.0011 (0.0002)	0.0000***
PRIM	−0.4617 (0.1564)	−0.0203***	0.5053 (0.1500)	−0.0339***	−0.3349 (0.1092)	−0.0065***	−0.2530 (0.1341)	−0.0054*
SECOND	−0.8503 (0.1435)	−0.0465***	−0.7962 (0.1417)	−0.0666***	−0.4957 (0.1139)	−0.0119***	−0.4928 (0.1431)	−0.0147***
PSECON	−1.1716 (0.1511)	−0.0954***	−0.9766 (0.1629)	−0.1070***	−0.9242 (0.1357)	−0.0412***	−0.9147 (0.1792)	−0.0485***
GRAD	−1.1412 (0.1622)	−0.0981***	−0.9038 (0.2238)	−0.1005***	−0.9276 (0.2082)	−0.0451***	−1.2336 (0.3467)	−0.0971***
MARRIED	0.9459 (0.1077)	0.0557***	0.7393 (0.1169)	0.0551***	0.4312 (0.1281)	0.0092***	0.8628 (0.1136)	0.0349***
ASST01	0.2833 (0.0840)	0.0077***	0.0736 (0.1254)	0.0035	0.2172 (0.1309)	0.0027*	−0.0211 (0.1666)	−0.0004
ASST02	0.5587 (0.2355)	0.0104**	−0.3217 (0.2607)	−0.0216	0.1869 (0.2082)	0.0023	0.2920 (0.3106)	0.0038
ASST03	−0.0032 (0.0202)	−0.0001	0.0105 (0.0321)	0.0005	0.0159 (0.0212)	0.0002	0.0288 (0.0271)	0.0005
HRENT	0.2922 (0.0806)	0.0088***	0.4236 (0.1187)	0.0188***	−0.5570 (0.3078)	−0.0176*	−0.5064 (0.3648)	−0.0171
HFREE	−0.0587 (0.1886)	−0.0020	−0.0236 (0.2718)	−0.0012	0.1207 (0.3700)	0.0016	0.3156 (0.5189)	0.0039
NMEN	–[b]		−0.1111 (0.0575)	−0.0055*	–[b]		−0.1120 (0.0552)	−0.0020**
NFEM	−0.1591 (0.0479)	0.0479***	–[b]		−0.0199 (0.0514)	−0.0003	–[b]	
OLD01	−0.0725 (0.1621)	0.1621	0.0987 (0.2312)	0.0045	−0.2410 (0.1287)	−0.0048*	−0.1174 (0.1896)	−0.0024
OLD02	0.6197 (0.5525)	0.5525	0.3279 (0.6094)	0.0118	−0.2448 (0.2932)	−0.0052	−0.4439 (0.3401)	−0.0137
MALE_HHEAD	0.3258 (0.0925)	0.0142***	0.1026 (0.1543)	0.0054	0.2249 (0.1287)	0.0045*	0.1756 (0.1745)	0.0037
FEMALE_HHEAD	0.2821 (0.2257)	0.0067	0.3167 (0.2437)	0.0123	0.5319 (0.2683)	0.0045**	0.0412 (0.2391)	0.0007
Observations	5951		2368		6443		4217	

Note: Dependent variable: *PART*—binary variable of employment in the labour force.
[a] ME, the marginal effect for dummy variables, is computed as dF/dx as x changes from zero to unity, holding the other explanatory variables at their sample means.
[b] Variable not included.
Standard errors in parentheses. *Significant at 10%; **significant at 5%; ***significant at 1%.

where owning the house where the person resides has a positive and statistically significant impact on employment probabilities for males.

4.1.2 Results from the selectivity-corrected wage regression. The estimated wage equations corrected for selection bias are reported for males and females separately for the rural and the urban samples in Table 3. Age (*AGE*) and age squared (*AGESQ*) are statistically significant at conventional levels except for rural female employees. As expected, wage rates increase with age, but the relationship is concave. In the urban sample wage rates peak at age 59 for males and at age 43 for females. The corresponding numbers for the rural sample are 46 and 29, respectively.

The estimated rates of return to educational attainment increase monotonically for both men and women. However, it is interesting to note that the returns to education for females are highest for women residing in urban areas. This result is possibly attributable to sectoral differences in labour market opportunities as well as differences in productivity that provide higher wage rates for women with more educational attainment. Training appears to have a stronger role to play in the wage determination of women compared with that of men. In the urban sample females with general (*TRA_GEN*) and technical (*TRA_TECH*) training earn higher premiums compared with their counterparts with no training (the omitted category), whereas for males, though the wage effect is small, wage rates for those with general (*TRA_GEN*) and technical (*TRA_TECH*) training are statistically significantly different from the wage rates of males with no training. In the rural sample females with general training (*TRA_GEN*) are paid higher wage rates compared with females with no training, whereas for males the opposite is true.

The wage rates of self-employed males and females (*SELF_EMPD*), of wage employees (*W_EMPD*) and of casual workers (*C_WKR*) in the urban sample are positive and statistically significantly different from the wage rate of unpaid family workers (the omitted category). However, we find a similar result in the rural sample, except for the female wage employee. Although in all specifications the coefficients for males are larger than for females, gender differences are well pronounced in self-employment (*SELF_EMPD*) in both the rural and the urban samples. One possible explanation for the observed difference is that males who select into self-employment are generally endowed with higher levels of education (i.e. some secondary and post secondary education) compared with females and, accordingly, are more likely to succeed in the labour market.

Being married (*MARRIED*) significantly reduces female wage rates in the rural sample. One explanation could be that married women involve higher costs of employment via maternity leave entitlements. On the other hand, marriage (*MARRIED*) leads to an increase in men's wage rates above the reference group (i.e. never married/separated/divorced/widowed) in the urban sample, whereas *MARRIED* becomes statistically insignificant for males in the rural sample.

Finally, it is worth noting that λ is always negative and statistically significant for males. It implies that the expected wage of male employees who select into employment, *ceteris paribus*, is lower than a man selected at random from the sample of employed and unemployed population with comparable observable characteristics.[24] On the other hand, the coefficient on λ is negative but not statistically different from zero at conventional levels of significance in the log-wage equations of females in all samples. What this implies is that working females, by contrast, earn no more than would have an average female in the entire population. The empirical literature on this issue (see Ermisch &

Table 3. Selectivity-adjusted estimates of wage equations: LFS 1999–2000

	Urban sample		Rural sample	
Variable	Males coef.	Females coef.	Males coef.	Females coef.
Constant	4.2340***	5.3172***	5.0638***	6.5943***
	(0.1740)	(0.2594)	(0.1648)	(0.1516)
LN_MHRS	−0.7536***	−0.7008***	−0.8569***	−0.9379***
	(0.0266)	(0.0335)	(0.0236)	(0.0162)
AGE	0.04739***	0.0173**	0.0277***	0.0058
	(0.0048)	(0.0072)	(0.0039)	(0.0036)
AGESQ	−0.0004***	−0.0002**	−0.0003***	−0.0001
	(0.0001)	(0.0001)	(0.0000)	(0.0000)
UNEMPS	0.0087	0.1123	0.0036	0.0016
	(0.0085)	(0.1498)	(0.0118)	(0.0122)
TRA_VOC	−0.033	0.1258	0.0323	0.0514
	(0.0851)	(0.1738)	(0.0905)	(0.1193)
TRA_GEN	0.1815***	0.4233***	−0.1743***	0.1851**
	(0.0596)	(0.1173)	(0.0664)	(0.0805)
TRA_TECH	0.1268**	0.4562***	−0.0271	−0.0615
	(0.0558)	(0.1650)	(0.1153)	(0.4610)
SELF_EMPD	1.1726***	0.1316**	1.2462***	0.3638***
	(0.0383)	(0.0583)	(0.0282)	(0.0294)
W_EMPD	0.9464***	0.3835***	0.9085***	0.0408
	(0.0410)	(0.0643)	(0.0377)	(0.0501)
C_WKR	0.5466***	0.1089*	0.5366***	0.1034***
	(0.0441)	(0.0635)	(0.0287)	(0.0304)
PRIM	0.1350***	0.1129***	0.1328***	0.0006
	(0.0251)	(0.0387)	(0.0200)	(0.0178)
SECOND	0.2891***	0.1205***	0.2617***	0.0211
	(0.0266)	(0.0461)	(0.0244)	(0.0263)
PSECON	0.5423***	0.1812***	0.4307***	0.1092**
	(0.0329)	(0.0679)	(0.0399)	(0.0544)
GRAD	0.8128***	0.6540***	0.6538***	0.2961**
	(0.0389)	(0.0935)	(0.0613)	(0.1172)
OCCUP_PROF	0.0167	0.4946*	0.1043	0.6014
	(0.0745)	(0.2829)	(0.2398)	(0.4717)
OCCUP_CLERIC	−0.1122**	−0.1033	0.1106	0.2937**
	(0.0505)	(0.1256)	(0.0858)	(0.1468)
OCCUP_SALES	−0.0369	−0.1578	0.1348	−0.4616***
	(0.0538)	(0.1496)	(0.0882)	(0.1516)
OCCUP_SER	−0.297***	−0.8834***	−0.0457	−0.7535***
	(0.0572)	(0.1074)	(0.0905)	(0.1139)
OCCUP_AGRI	−0.0507	−0.6039***	0.0106	−0.55***
	(0.0742)	(0.1375)	(0.0833)	(0.1030)
OCCUP_PROD	−0.1503***	−0.4872***	0.1627**	−0.7077***
	(0.0491)	(0.1009)	(0.0789)	(0.1121)
INDS_MANU	0.1520**	−0.1475	−0.1030*	0.1223*
	(0.0632)	(0.1133)	(0.0533)	(0.0724)
INDS_WHRET	0.0622	−0.1160	−0.1076*	−0.0648
	(0.0632)	(0.1498)	(0.0593)	(0.1171)
INDS_HOTEL	0.0614	0.2245	0.0692	−0.5543**
	(0.0808)	(0.1917)	(0.0938)	(0.2224)
INDS_COMMU	0.0614	0.4431**	−0.2330***	0.5512***
	(0.0642)	(0.1886)	(0.0600)	(0.1892)

Table 3. *Continued*

	Urban sample		Rural sample	
Variable	Males coef.	Females coef.	Males coef.	Females coef.
INDS_FINST	0.2859***	0.4176	0.2157	0.9923***
	(0.0900)	(0.2897)	(0.1588)	(0.3388)
INDS_REALEST	0.1622	0.4074	0.2508	1.2897***
	(0.1061)	(0.2487)	(0.1753)	(0.3349)
INDS_PUBADMN	0.1964***	0.4519***	0.2282**	1.4542***
	(0.0719)	(0.1518)	(0.0906)	(0.2803)
INDS_HLTH	0.09111	−0.1302	−0.0730	0.0951
	(0.0640)	. (0.1084)	(0.0646)	(0.0819)
INDS_EDU	−0.1298	−0.3965***	0.0646	−0.076
	(0.0849)	(0.1455)	(0.0930)	(0.1319)
SKIL	0.1553***	0.1788***	0.1799***	0.0595**
	(0.0190)	(0.0471)	(0.0171)	(0.0262)
MARRIED	0.0854**	−0.0765	0.0123	−0.0787***
	(0.0390)	(0.0470)	(0.0333)	(0.0296)
λ	−0.3724***	−0.0919	−0.6218***	−0.0206
	(0.1251)	(0.1824)	(0.1900)	(0.1349)
Observations	5654	2221	6265	4097
Wald chi-squared	5092.46	1456.52	6317.47	4362.34

Note: Dependent variable: natural logarithm of hourly wages. Standard errors in parentheses.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Wright, 1994) suggests that a negative coefficient on λ in the female sample is very plausible when a woman's reservation and offered wages are positively correlated. This is probably because women who are more productive in jobs also tend to be more productive in home activities. Using data from Pakistan, Aslam (2009) and Ashraf & Ashraf (1993) obtained similar negative values of λ. Given the comparable development levels of Bangladesh and Pakistan, the result is perhaps not surprising. The selection correction term λ is not, however, statistically significant, suggesting that selectivity-adjusted estimates for females should be interpreted cautiously.

4.2 Decomposition of the Wage Differential

Next we turn to the decomposition results (see Tables 4 and 5). First, we report decompositions of the wage differential, unadjusted for sample selection bias. The decomposition of the wage differential, adjusted for sample selection bias, is reported next.

The decomposition of unadjusted estimates (Table 4) reveals that the wage differential between males and females in the urban sample when males are the baseline (equation (2)) is 67.03 percentage points (column 1). The decomposition of this gender gap shows that the explained proportion is smaller in magnitude than the discriminatory proportion of the gender wage differential. After accounting for differences in productive characteristics (the explained proportion), the adjusted wage gap (or the discriminatory component) remains 48.51 percentage points, or 72%. The difference in the explained component is 18.52 percentage points in favour of males. In other words, 28% of the differential is due to the superior endowments of the typical male. However, using the female wage structure

Table 4. Decomposition of the gender wage differential (urban sample)

	OLS			Selectivity corrected	
	(1)	(2)	(3)	(4)	(5)
Differences in observed wages	0.6703	0.6703	0.6703	0.6703	0.6703
Differences in adjusted wages	0.4851	0.5914	0.5148	na	na
Differences in selection bias	na	na	na	0.0206	0.0206
Differences in offered wages	na	na	na	0.6909	0.6909
Contribution of characteristics					
LN_MHRS	−0.4024	−0.3744	−0.3946	−0.4029	−0.3747
AGE	0.2092	0.0785	0.1726	0.1939	0.0709
AGESQ	−0.1475	−0.0629	−0.1238	−0.1358	−0.0577
UNEMPS	0.0003	0.0033	0.0011	0.0003	0.0034
EDUCATION	0.1093	0.0659	0.0971	0.1162	0.0683
SKIL	0.0352	0.0403	0.0366	0.0346	0.0399
TRAINING	0.0030	0.0105	0.0051	0.0030	0.0105
MARRIED	0.0016	−0.0006	0.0009	0.0009	−0.0008
JOB STATUS	0.3313	0.0381	0.2493	0.3313	0.0381
OCCUPATION	0.0204	0.1802	0.0652	0.0201	0.1800
INDUSTRY	0.0248	0.1000	0.0460	0.0228	0.0999
Total	0.1852 (28%)	0.0789 (12%)	0.1555 (23%)	0.1844 (27%)	0.0778 (11%)
Discrimination	0.4851 (72%)	0.5914 (88%)	0.5148 (77%)	0.5065 (73%)	0.6131 (89%)
Male advantage			0.4973		
Female disadvantage			1.2020		
Differences in constants			−1.1845		

Note: Dependent variable: natural logarithm of hourly wages.
Columns 1 and 4 show the result from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 5 use the female wage structure as the non-discriminatory following the Blinder approach. Column 3 describes the decomposition using the Cotton approach. Total discriminatory component in the Cotton approach is the sum of the components attributable to the differences in constant terms, the male advantage and the female disadvantage. The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na: not available.

Table 5. Decomposition of the gender wage differential (rural sample)

	OLS			Selectivity corrected	
	(1)	(2)	(3)	(4)	(5)
Differences in observed wages	0.3551	0.3551	0.3551	0.3551	0.3551
Differences in adjusted wages	0.1532	0.7452	0.3901	na	na
Differences in selection bias	na	na	na	0.0304	0.0304
Differences in offered wages	na	na	na	0.3855	0.3855
Contribution of characteristics					
LN_MHRS	− 0.6015	− 0.6609	− 0.6252	− 0.6039	− 0.661
AGE	0.1076	0.0199	0.0725	0.0901	0.0189
AGESQ	− 0.0863	− 0.0182	− 0.059	− 0.0719	− 0.0172
UNEMPS	3.60E-05	1.73E-07	2.85E-07	3.80E-08	1.71E-09
EDUCATION	0.0471	0.0112	0.0327	0.0524	0.0115
SKIL	0.0505	0.0166	0.0369	0.0501	0.0166
TRAINING	− 0.0008	0.0008	− 0.0002	− 0.0009	0.0009
MARRIED	− 0.0055	0.0063	− 0.0008	− 0.001	0.0065
JOB STATUS	0.6796	0.173	0.4769	0.6752	0.1727
OCCUPATION	0.0302	0.0272	0.029	0.0298	0.0273
INDUSTRY	− 0.019	0.034	0.0022	− 0.0193	0.0339
Total	0.2019 (57%)	− 0.3901 (− 109%)	− 0.0350 (− 0.09%)	0.2006 (52%)	− 0.3899 (− 101%)
Discrimination	0.1532 (43%)	0.7452 (209%)	0.3901 (109%)	0.1849 (48%)	0.7754 (201%)
Male advantage			0.9852		
Female disadvantage			1.1226		
Differences in constants			− 1.7177		

Note: Dependent variable: natural logarithm of hourly wages.
Columns (1) and (4) show the result from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 5 use the female wage structure as the non-discriminatory following the Blinder approach. Column 3 describes the decomposition using the Cotton approach. Total discriminatory component in the Cotton approach is the sum of the components attributable to the differences in constant terms, the male advantage and the female disadvantage.
The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na: not available.

(column 2) as the non-discriminatory norm (equation (3)) results in a discriminatory component of 59.14 percentage points, or 88%, leaving an explained component of 7.89 percentage points, or 12%. The Blinder (1973) decomposition results obtained from the two baselines in fact suggest that a large fraction of the wage gap between men and women is not explained by differences in the accumulation of productive differences and discrimination is even larger when the female rather than the male structure is used as the competitive standard.[25] The Cotton (1988) decomposition analysis (equation (4)) provides similar results, with discrimination and productivity differences (explained component) accounting, respectively, for 77% (i.e. $(0.4973 + 1.2020\ (1.1845) = 0.5148$ or 51.48 percentage points) and 23% of the differential (column 3). The implication here is that discrimination is playing a role in the determination of wage rates and the major contribution is attributed to female disadvantages in the urban labour market.

In the rural sample (Table 5), when the male wage structure is the baseline (equation (2)), the explained and discriminatory components of the gender wage differential are 20.19 percentage points, or 57%, and 15.32 percentage points, or 43%, respectively (column 1). So there is a large difference between men and women in terms of the components of characteristics (explained proportion). While using the female wage structure as the standard baseline (equation (3)), the decomposition reveals that the differences in productivity characteristics serve to reduce rather than widen the wage gap (column 2). However, after controlling for productivity-related factors, the adjusted wage gap remains 74.52 percentage points (i.e. $(0.3551 + 0.3901) = 0.7452$). The large unexplained wage gaps in rural areas, ranging from 43 to 209%, might reflect unobserved differences between men and women that affect earnings. This may also imply that males are over-rewarded for observed characteristics compared with females in the rural labour market. Finally, using the Cotton (1988) approach (column 3), we find that discrimination against women (combining the "male advantage", the "female disadvantage" and the "differences in constants") is relatively low (39.01 percentage points).

Tables 4 and 5 imply that a large proportion of the gender wage differential in both urban and rural areas is attributable to differences in constants. The constant term can be considered as an indicator of base level wage. It appears that women in Bangladesh receive substantially lower base level wages than males and a significant portion of the gender wage differential can be attributed to the lower basic wage received by women.[26] The results are consistent with other studies (see, e.g. Gregory, 1999; Eastough & Miller, 2004)

Controlling for sample selection factor in wage equations for rural and urban samples reveals that the wage offer differential is higher than the observed wage differential because of the difference in the average selectivity bias between males and females (Tables 4 and 5, columns 4 and 5). Although the percentage contribution of the difference in the selectivity bias is very small, the addition of the term to the observed wage gap indicates that if the wage equation is not adjusted for sample selection bias, the extent of the wage differential will be underestimated.

In the urban sample, irrespective of weights, the results provide evidence of significant discrimination that accounts for between 73 and 89% of the wage offer differential. In the rural sample except for the male baseline, the male–female wage offer differential is largely attributable to discrimination rather than any endowment effect (Table 5, column 5). However, it is worth noting that the wage gap due to differences in productivity (explained component) is negative when the female wage structure is used as the baseline (equation (10)). It implies that if female characteristics are comparable to the male characteristics but rewarded according to the female wage structure in the labour market, then the wage gap is reduced significantly.

4.3 Robustness of Our Results

In this section we examine the robustness of our results to alternative measures of wages. First, we examine what happens if we restrict ourselves to the sample of full-time workers aged 15 years and higher both for the urban sample and for the rural sample. This would allow us to avoid potential biases resulting from different wage structures faced by full-time and part-time workers. Unfortunately, the LFS 1999–2000 did not collect direct information on whether an individual is employed on a full-time or part-time basis. This needs to be computed. We do so, assuming that any individual working 48 hours per week is a full-time worker.[27] If we restrict ourselves to the set of full-time workers, we are left with a sample of 12 582 individuals, of whom 5131 are from urban areas (61% males) and 7451 from rural areas (46% males). The difference between this specification with the initial one (equation (1)) is that here we excluded the log of monthly hours worked (*LN_MHRS*) and the number of months unemployed (*UNEMPS*) from the set of explanatory variables.

In the results presented in Table 3, we included the log of monthly hours worked (*LN_MHRS*) as an additional explanatory variable because we were interested in seeing whether the work duration has any effect on the wage structure. Alternatively, we could have included only a full-time dummy—that would in some sense tell the same story. In unreported regressions we re-estimated the model but included a full-time dummy and not the log of monthly hours worked. Hourly wages are generally significantly lower for full-time workers. When we restrict ourselves to full-time workers we do not have sufficient variability in the log of monthly hours worked (*LN_MHRS*) to obtain sensible estimates. Therefore, in this specification we dropped this particular variable from the set of explanatory variables. Additionally, one could expect that if any individual has been unemployed (*UNEMPS*) in the near past, he/she is more likely to be employed as a part-time worker when he/she is finally employed. As we are restricting ourselves to full-time workers, we chose to exclude this variable from the set of explanatory variables.

Second, instead of the hourly wage as the variable of interest, we re-estimated equation (1) but this time considered the log of monthly wages as the dependent variable. In this specification, we included all variables except workers' monthly hours of work (*LN_MHRS*) and the number of months unemployed (*UNEMPS*). We dropped the number of months unemployed (*UNEMPS*) because it is difficult to predict *a priori* how or why this variable should affect monthly wages for full-time workers. We continued to restrict ourselves to full-time workers.

Results for the decomposition analysis are presented in Tables 6–9. We limit our discussion to the Blinder decomposition. The full set of results and those corresponding to the Cotton decomposition method are available on request.

In Tables 6 and 7 we present the corresponding decomposition analysis where the dependent variable is the log of hourly wage (the sample is restricted to full-time workers only). When we compare these results with those presented in Tables 4 and 5, although the results are qualitatively similar, the discrimination term is now smaller in magnitude. For example, in the urban sample when we compare the results in columns 1–2 in Tables 4 and 6 we find that the effect of discrimination decreases from 72 to 37% when we use males as the base (column 1 of Tables 4 and 6) and decreases from 88 to 67% when we use females

Table 6. Decomposition of the gender wage differential (urban sample: full-time workers)

	OLS		Selectivity corrected	
	(1)	(2)	(3)	(4)
Differences in observed wages	0.7664	0.7664	0.7664	0.7664
Differences in adjusted wages	0.2831	0.5116	na	na
Differences in selection bias	na	na	0.2582	0.2582
Differences in offered wages	na	na	1.0246	1.0246
Contribution of characteristics				
AGE	0.1478	0.0698	0.1382	0.0706
AGESQ	−0.0945	−0.0598	−0.0870	−0.0605
EDUCATION	0.0905	0.0662	0.0961	0.0660
SKIL	0.0403	0.0115	0.0399	0.0116
TRAINING	0.0022	0.0066	0.0021	0.0066
MARRIED	−0.0059	0.0004	−0.0032	0.0004
JOB STATUS	0.2852	−0.0367	0.2859	−0.0367
OCCUPATION	−0.0128	0.1463	−0.0126	0.1463
INDUSTRY	0.0305	0.0505	0.0285	0.0506
Total	0.4833 (63%)	0.2548 (33%)	0.4879 (48%)	0.2549 (25%)
Discrimination	0.2831 (37%)	0.5116 (67%)	0.5367 (52%)	0.7697 (75%)

Note: Dependent variable: natural logarithm of hourly wages.
Columns 1 and 3 show the result from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 4 use the female wage structure as the non-discriminatory following the Blinder approach.
The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na: not available.

as the base (column 2 of Tables 4 and 6). The corresponding figures for the rural sample are 43 and −37% when we use males as the base (column 1 of Tables 5 and 7)[28] and 209 and 81% when we use females as the base (column 2 of Tables 5 and 7).

Tables 8 and 9 (log of monthly wage is the dependent variable, sample is restricted to full-time workers) show that total log wage differential between full-time males and females is 1.2391 percentage points in urban areas and 1.0256 percentage points in rural areas; this in turn implies that the average monthly wage for males is 245% higher than that of females in urban areas and 179% higher in rural areas. When we compare these results with those presented in Tables 4 and 5, we find that restricting the sample to full-time workers only results in a significant increase of wage differentials in both urban and rural areas. Looking first at the results for urban areas (columns 1–2 of Table 8), the difference in productive characteristics (the explained component) is in favour of males and is now even greater than the contribution of the explained component presented in columns 1–2 in Table 4. The dominant factors favouring males (using either male or female wage structures) are age (*AGE*) and job status, both of which contribute to the widening of the wage gap. Although the discrimination component is now smaller in magnitude compared with those presented in columns 1–2 of Table 4, the major component of the wage gap is unexplained. After accounting for differences in

Table 7. Decomposition of the gender wage differential (rural sample: full-time workers)

	OLS		Selectivity corrected	
	(1)	(2)	(3)	(4)
Differences in observed wages	0.4899	0.4899	0.4899	0.4899
Differences in adjusted wages	−0.1824	0.3962	na	na
Differences in selection bias	na	na	0.2755	0.2755
Differences in offered wages	na	na	0.7654	0.7654
Contribution of characteristics				
AGE	0.1199	0.0102	0.1139	0.0133
AGESQ	−0.0976	−0.0039	−0.0920	−0.0071
EDUCATION	0.0589	0.0105	0.0610	0.0089
SKIL	0.0475	−0.0165	0.0474	−0.0164
TRAINING	−0.0010	0.0016	−0.0010	0.0016
MARRIED	0.0001	0.0161	0.0024	0.0146
JOB STATUS	0.5429	−0.0118	0.5422	−0.0108
OCCUPATION	0.0219	0.0539	0.0217	0.0537
INDUSTRY	−0.0203	0.0335	−0.0204	0.0339
Total	0.6723 (137%)	0.0937 (19%)	0.6752 (88%)	0.0917 (12%)
Discrimination	−0.1824 (−37%)	0.3962 (81%)	0.0902 (12%)	0.6737 (88%)

Note: Dependent variable: natural logarithm of hourly wages.
Columns 1 and 3 show the result from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 4 use the female wage structure as the non-discriminatory following the Blinder approach.
The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na = not available.

Table 8. Decomposition of the gender wage differential (urban sample: full-time workers)

	OLS		Selectivity corrected	
	(1)	(2)	(3)	(4)
Differences in observed wages	1.2391	1.2391	1.2391	1.2391
Differences in adjusted wages	0.6511	0.7092	na	na
Differences in selection bias	na	na	0.1823	0.1823
Differences in offered wages	na	na	1.4214	1.4214
Contribution of characteristics				
AGE	0.2239	0.0834	0.2147	0.0795
AGESQ	−0.1744	−0.0695	−0.1671	−0.0667
EDUCATION	0.0996	0.0714	0.1049	0.0726
SKIL	0.0375	0.0224	0.0371	0.0221
TRAINING	0.0026	0.0084	0.0025	0.0084
MARRIED	−0.0069	0.0023	−0.0044	0.0028
JOB STATUS	0.3831	0.1161	0.3838	0.1159
OCCUPATION	−0.0029	0.1655	−0.0027	0.1656
INDUSTRY	0.0256	0.1299	0.0237	0.1298
Total	0.5880 (47%)	0.5299 (43%)	0.5925 (42%)	0.5298 (37%)
Discrimination	0.6511 (53%)	0.7092 (57%)	0.8289 (58%)	0.8916 (63%)

Note: Dependent variable: natural logarithm of monthly wages.
Columns 1 and 3 show the result from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 4 use the female wage structure as the non-discriminatory following the Blinder approach.
The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na: not available.

characteristics, the adjusted wage gap is 65.11 percentage points or 53% (using male wage structure) and 70.92 percentage points or 57% (using female wage structure), respectively.

For the rural sample, the results obtained using the male wage structure (column 1 of Table 9) are qualitatively similar to those presented in column 1 of Table 5. The majority of the wage gap is due to differences in characteristics (the explained component). While using the female wage structure as the standard baseline (column 2 of Table 9), the decomposition reveals that the differences in productivity characteristics are now in favour of males that widen the wage gap. After controlling for productivity-related factors, the adjusted wage gap remains 72.91 percentage points, which is the major component of the gap. The decompositions for the selectivity-adjusted estimates (columns 3–4 of Table 9) exhibit a different pattern of results. When the male wage structure is used as the competitive standard (column 3 of Table 9), the explained and the unexplained (discrimination) components contribute equally in percentage terms to the wage gap. The decomposition results obtained with the female wage structure, however, suggest that a large fraction of the wage gap is unexplained, but it is now smaller in magnitude compared with the results presented in column 4 of Table 5. It is also worth noting that the differences in the selectivity bias are very large (see columns 3–4 of Table 9) and the differences in offered wages (which is the sum of the differences in observed wages and the differences in selection bias) are much higher for full-time workers.

Table 9. Decomposition of the gender wage differential (rural sample: full-time workers)

	OLS		Selectivity corrected	
	(1)	(2)	(3)	(4)
Differences in observed wages	1.0256	1.0256	1.0256	1.0256
Differences in adjusted wages	0.2466	0.7291	na	na
Differences in selection bias	na	na	0.5509	0.5509
Differences in offered wages	na	na	1.5765	1.5765
Contribution of characteristics				
AGE	0.1334	0.0173	0.1147	0.0158
AGESQ	−0.1239	−0.0164	−0.1066	−0.0149
EDUCATION	0.0558	0.0155	0.0622	0.0162
SKIL	0.0539	0.0094	0.0535	0.0094
TRAINING	−0.0019	0.0006	−0.0019	0.0006
MARRIED	−0.0063	0.0119	0.0007	0.0126
JOB STATUS	0.6619	0.1564	0.6598	0.1559
OCCUPATION	0.0229	0.0585	0.0225	0.0586
INDUSTRY	−0.0168	0.0434	−0.0171	0.0432
Total	0.7790	0.2965	0.7877	0.2965
	(76%)	(29%)	(50%)	(19%)
Discrimination	0.2466	0.7291	0.7888	1.2790
	(24%)	(71%)	(50%)	(81%)

Note: Dependent variable: natural logarithm of monthly wages.
Columns 1 and 3 show the results from the Blinder approach and use the male wage structure as the non-discriminatory norm, whereas columns 2 and 4 use the female wage structure as the non-discriminatory following the Blinder approach.
The following explanatory variables are included in each group. Education: *PRIM*, *SECOND*, *PSECON*, *GRAD*. Training: *TRA_VOC*, *TRA_GEN*, *TRA_TECH*. Job status: *SELF_EMPD*, *W_EMPD*, *C_WKR*. Occupation: *OCCUP_PROF*, *OCCUP_CLERIC*, *OCCUP_SALES*, *OCCUP_SER*, *OCCUP_AGRI*, *OCCUP_PROD*. Industry: *INDS_MANU*, *INDS_WHRET*, *INDS_HOTEL*, *INDS_COMMU*, *INDS_FINST*, *INDS_REALEST*, *INDS_PUBADMN*, *INDS_HLTH*, *INDS_EDU*.
na: not available.

5. Conclusions and Policy Implications

This paper has analysed the determinants of employment decisions and (log) of hourly wages (adjusted and unadjusted for selectivity bias), and decomposed the gender wage gap in rural and urban labour markets of Bangladesh for the individual level unit record data from the LFS 1999–2000. We found that there exists a large gap between male and female wage rates. This is more so in urban areas than rural areas. We argue that much of the wage gap in the urban areas is directly attributable to discrimination using either the male or the female wage structure as the discriminatory norm. However, it needs to be noted that the wage gap in rural areas varies much more compared with that in the urban areas, indicating possibly that the results for the rural sample are less robust compared with those in the urban areas. Our results also show that the discrimination component is larger and the productivity difference component is smaller in magnitude across all samples when the female rather than the male wage structure is used as the competitive standard: a clear example of the “index number problem”. On the other hand, the results using the Cotton (1988) decomposition method suggest that the largest component of the unexplained wage gap in both the rural and the urban samples (Tables 4 and 5) comes from females being disadvantaged. Therefore, discrimination against women is more prevalent than nepotism towards men in explaining the wage gap. Correcting for sample selection bias leads to an increase in the gender wage gap in both rural and urban areas, and is an important contributor to the total discrimination component. The implication of this result is that the extent of the male–female wage gap is likely to be understated if selection bias effects are ignored in the wage equations.

The main policy implication of the analysis is that the most fruitful approach could be to call for affirmative action policies that persuade firms to employ women on comparable pay and working conditions. Additionally, in order to implement legislative and public policy changes there appears to be a strong need for gender-conscious planning.

Notes

[1] Kapsos (2008) in his analysis uses data from the Bangladesh national occupational wage survey, which was conducted by the Bangladesh Bureau of Statistics in 2007. That survey collected data on wages and hours of work from establishment surveys and establishments in the survey were not selected randomly.

[2] 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.

[3] A similar method has been used by Rodgers (2004).

[4] The selection equation represents the probability that an individual is employed at a given time conditional on a set of personal characteristics.

[5] In other words, a significant estimate of θ_j, the coefficient on the selectivity term, indicates the presence of selectivity.

[6] See, for example, Creedy *et al.* (2000) and Kidd & Viney (1991).

[7] This part may be viewed as the differences in unobservables, which influence wages.

[8] See also Duncan & Leigh (1980), Boymond *et al.* (1994) and Reimers (1983).

[9] The lower boundary of the working age group is 15, that is, we do not account for child labour. The upper boundary of the working age group is higher than the conventional retirement age of 60–65 years. This is due to the inclusion of the rural areas where individuals tend to work beyond their conventional retirement age.

[10] A casual worker refers to a wage worker whose services are solicited only for a periodic time intervals during the reference period (i.e. the week preceding the day of the survey).

[11] An unpaid family worker is a person who works at least 1 hour in the reference period (other than household work) without pay or profit in a family-operated farm or in a business owned/operated by the household head or other members of the household to whom he/she is related by kinship, marriage, adoption or dependency. Unpaid family workers who worked at least 1 hour or more during the reference period are considered as a part of the labour force.

[12] As we use an extended definition of the labour force, persons with non-market activities such as unpaid family workers are also included.

[13] The detailed definitions of the explanatory variables are presented in the Appendix.

[14] We do not have any information on actual labour market experience. Age is used as the approximate variable for general labour market experience. Moreover, as age increases, productivity and wage rates tend to rise; but further increases in age may lead to a decline in wage rates and productivity because of diminishing marginal returns. To capture the concavity of the wage profile a quadratic age term is included.

[15] This variable measures the gross elasticity of hours worked per month with respect to wage rates. Ajwad & Kurukulasuraiya (2002) also employ this variable in their study on ethnic and gender wage disparities in Sri Lanka. The authors found that the log of hours worked has a negative and significant impact on wage rates.

[16] It has been hypothesized that time out of the labour force can result in depreciation of human capital and depress wage rates (Mincer & Polachek, 1974).

[17] Here skill refers to occupational skills.

[18] One possible test is presented in Bryan (2007). To gain confidence that the variables included in Z, but not included in X are not actually incorrectly excluded from X, following Bryan (2007) we regressed

the residuals from the selection-corrected wage equation on all the exogenous variables and then calculated the test statistic NR^2, where N is the sample size and the R^2 is from this supplementary regression. The statistic was then compared with the appropriate critical value from the $x^2(k-1)$ distribution, where k is the number of excluded variables. Under the null hypothesis, the instruments are uncorrelated with the error term, and excluded instruments are truly exogenous to the main (wage) equation. In both the rural and urban samples, the p-values (0.10 for the urban sample and 0.33 for the rural sample) of the test confirmed the validity of excluded instruments used. We conducted another test: we re-ran the regressions but this time we included all of the excluded variables in the wage equation and examined: (1) whether the remaining coefficients are sensitive to the inclusion of these additional variables; and (2) whether these additional variables are statistically significant in the wage equation. The results indicate that including the set of identification variables in the wage equation does not alter our initial estimates. These excluded variables are not statistically significant in the wage equations, implying that these variables do not have a direct effect on wage rates.

[19] One of the major drawbacks in our data set is that we do not have information on pre-school-age children, which would not only be a deterrent to women for being employed but also influence their engagement with the labour market.

[20] Differences in the natural logarithm of hourly wages can be converted to percentage wage differences using the formula 100[exp (difference) − 1]. Hence, the difference between the mean of the natural logarithm of hourly wages for males and females in the urban sample in Table 1 of 0.6703 yields a 100[exp (0.6703) − 1] = 95% wage differential. It is also worth noting that the gender wage differential in our urban data is significantly larger than that found using similar methodologies in other studies: Loureiro *et al.* (2004) found a 20 and 61% gap in the urban labour market of Brazil in 1992 and 1998, respectively, while Ashraf & Ashraf (1993) found a 61.14% gap in Rawalpindi City, Pakistan.

[21] 100[exp(0.3551) − 1] = 43%.

[22] SSC: Secondary School Certificate; HSC: Higher Secondary School Certificate.

[23] The results indicate that additional income from these assets increases the male's probability of employment compared with a situation in which he has "no asset" (the omitted category). One explanation could be that in the absence of the value of the assets (which are not provided in the sample), one cannot gauge the importance of feedback effects of these assets on the probability of employment.

[24] A number of studies have used the Heckman (1979) estimator, but the estimates of the coefficient on the λ variable in the male wage equation differ considerably: Miller & Rummery (1991) reported a significant, negative coefficient, Ashraf & Ashraf (1993) a positive, significant coefficient.

[25] Oaxaca & Ransom (1994) found similar results.

[26] This might indicate that the adverse consequences of a deregulated labour market are more pronounced for those women at the lower end of the wage distribution.

[27] In Bangladesh, the full work week is 48 hours.

[28] Notice that the discrimination component of column of 1 of Table 7 is negative when we use male wage structure as the non-discriminatory norm. In the Blinder decomposition, discrimination is defined as $(\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f + (\hat{\alpha}_m - \hat{\alpha}_f)$. Now when we consider the sample of full-time workers the value for $(\hat{\alpha}_m - \hat{\alpha}_f)$ becomes large (negative), resulting in the discrimination term becoming negative. Interestingly, this is not a new result. Using a different set of data, Akter (1999) found a similar result (see table 8, p. 130) for the rural labour market of Bangladesh when she used male wage structure as the non-discriminatory norm. Unfortunately, we do not have a proper explanation for this negative discrimination component. All we can say is that the results for the rural sample are quite sensitive to the specification used.

References

Adamchik, V. & Bedi, A. (2003) Gender pay differentials during the transition in Poland, *Economics of Transition*, 11(4), pp. 697–726.

Ajwad, M. & Kurukulasuraiya, P. (2002) *Ethnic and Gender Wage Disparities in Sri Lanka*, Policy Research Working Paper No. 2859 (Washington, DC: The World Bank).

Akter, S. (1999) Gender wage differentials in rural employment in Bangladesh, *Journal of Quantative Economics*, 15(1), pp. 115–143.

Akter, S. (2005) Occupational segregation, and wage discrimination, and impact on poverty in rural Bangladesh, *The Journal of Development Areas*, 39(1), pp. 15–39.

Appleton, S., Hoddinott, J. & Krishnan, P. (1999) The gender wage gap in three African countries, *Economic Development and Cultural Change*, 47(2), pp. 289–312.

Arrow, K. (1972) The models of job discrimination, in: A. H. Pascal (Ed.) *Racial Discrimination in Economic Life* (Lexington, MA: Lexington Books).

Ashraf, B. & Ashraf, J. (1993) Estimating the gender wage gap in Rawalpindi City, *Journal of Development Studies*, 24(2), pp. 365–376.

Aslam, M. (2009) Education gender gaps in Pakistan: is the labor market to blame? *Economic Development and Cultural Change*, 57(4), pp. 747–784.

Becker, G. (1957) *The Economics of Discrimination* (Chicago: University of Chicago Press).

Blinder, A. (1973) Wage discrimination: reduced form and structural estimates, *The Journal of Human Resources*, 8(4), pp. 436–455.

Boymond, M., Fluckiger, Y. & Silber, J. (1994) Wage discrimination and occupational segregation by gender: some evidence from Swiss data. Mimeo, Geneva.

Brainerd, E. (1998) Winners and losers in Russia's economic transition, *American Economic Review*, 88(5), pp. 1094–1116.

Bryan, L. M. (2007) Free to choose? Differences in the hours determination of constrained and unconstrained workers, *Oxford Economic Papers*, 59(2), pp. 226–252.

Cotton, J. (1988) On the decomposition of wage differentials, *The Review of Economics and Statistics*, 70(2), pp. 236–243.

Creedy, J., Duncan, A. & Harris, M. (2000) Wage functions for demographic groups in Australia, *Australian Journal of Labour Economics*, 4(4), pp. 296–316.

Duncan, G. & Leigh, D. (1980) Wage determination in the union and non-union sectors: a sample selectivity approach, *Industrial and Labour Relations Review*, 34(1), pp. 24–34.

Eastough, K. & Miller, P. (2004) The gender wage gap on paid and self-employment in Australia, *Australian Economic Papers*, 63(3), pp. 50–69.

Ermisch, J. & Wright, R. (1994) Interpretation of negative sample selection effects in wage offer equations, *Applied Economics Letters*, 1(11), pp. 187–189.

Glick, P. & Sahn, D. E. (1997) Gender and education impacts on employment and earnings in West Africa: evidence from Guinea, *Economic Development and Cultural Change*, 45(4), pp. 793–823.

Gregory, R. G. (1999) The gender wage gap in paid and self-employment in Australia, *Australian Economic Papers*, 43(3), pp. 273–278.

Heckman, J. (1979) Sample selection bias as a specification error, *Econometrica*, 47(1), pp. 153–161.

Horton, S. (Ed.) (1996) *Women and Industrialization in Asia* (London and New York: Routledge).

Jacob, M. (2006) Changes in the wage gap of gender and caste groups in India, unpublished thesis.

Jurajda, S. (2001) *Gender Wage Gap and Segregation in Late Transition*, Centre for Economic Policy Research Discussion Paper Series.

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

Kidd, M. & Viney, R. (1991) Sex discrimination and non-random sampling in the Australian labour market, *Australian Economic Papers*, 30(56), pp. 28–49.

Knight, J. & Sabot, R. (1982) Labour market discrimination in a poor urban economy, *Journal of Development Studies*, 19(1), pp. 67–87.

Loureiro, P., Carneiro, F. & Sachsida, A. (2004) Race and gender discrimination in the labor market: an urban and rural sector analysis for Brazil, *Journal of Economic Studies*, 31(2), pp. 129–143.

Majumder, P. & Mahmud, S. (1994) Barriers to female employment in urban Bangladesh, Dhaka. Mimeo, Bangladesh Institute of Development Studies (BIDS).

Majumder, P. & Zohir, S. (1993) *Employment and Occupational Mobility among Women in Manufacturing Industries of Dhaka City: Bangladesh, Findings from a Survey of Employees* (New Delhi: ILO/ARTEP).

Miller, P. & Rummery, S. (1991) Male–female wage differentials in Australia: a reassessment, *Australian Economic Papers*, 50(56), pp. 50–69.

Mincer, J. & Polachek, S. (1974) Family investments in human capital: earnings of women, *Journal of Political Economy*, 82(2), pp. 76–108.

Neumark, D. (1988) Employers' discriminatory behaviour and the estimation of wage discrimination, *Journal of Human Resources*, 23(3), pp. 279–295.

Oaxaca, R. (1973) Male–female wage differentials in urban labour markets, *International Economic Review*, 114(3), pp. 693–709.

Oaxaca, R. & Ransom, M. (1994) On discrimination and the decomposition of wage differentials, *Journal of Econometrics*, 61(1), pp. 5–21.

Págán, J. (2002) Gender differences in labor market decisions in rural Guatemala, *Review of Development Economics*, 6(3), pp. 428–441.

Phelps, E. (1972) The statistical theory of racism and sexism, *American Economic Review*, 62(4), pp. 659–661.

Psacharpoulos, G. & Tzannatos, Z. (Eds) (1992) *Case Studies on Women's Employment and Pay in Latin America* (Washington, DC: World Bank).

Reilly, B. & Newell, A. (2000) *The Gender Wage Gap in Transition from Communism: Some Empirical Evidence*, William Davidson Working Paper No. 305.

Reimers, C. (1983) Labor market discrimination against Hispanic and Black men, *Review of Economics and Statistics*, 65(4), pp. 570–579.

Rodgers, R. J. (2004) Hourly wages of full-time and part-time employees in Australia, *Australian Journal of Labour Economics*, 7(2), pp. 231–254.

Vella, F. (1998) Estimating models with sample selection bias, *Journal of Human Resources*, 33(1), pp. 129–169.

Zohir, S. (1998) Gender implications of industrial reforms and adjustment in the manufacturing sector of Bangladesh, unpublished thesis.

Appendix

Table A1. Definition of variables used in the estimation of the wage and selection equations

Dependent variables	
LNHRSINC	= log of hourly wages
PART	= 1 if individual is employed; = 0 otherwise
Independent variables	
LN_MHRS	= log of monthly hours worked
Age	
AGE	= Age of individual measured in years
AGESQ	= Age of individual squared
Education	
PRIM	= 1 if individual is between levels 1 and 5; = 0 otherwise
SECOND	= 1 if individual is between levels 6 and 10; = 0 otherwise
*PSECO*N	= 1 if individual attains SSC/HSC[a] and equivalent; = 0 otherwise
GRAD	= 1 if individual attains graduate or higher degree; = 0 otherwise
Marital status	
MARRIED	= 1 if individual is married; = 0 otherwise
Occupation	
OCCUP_PROF	= 1 if occupation category is professional/technical; = 0 otherwise
OCCUP_CLERIC	= 1 if occupation category is clerical; = 0 otherwise
OCCUP_SALES	= 1 if occupation category is sales; = 0 otherwise
OCCUP_SER	= 1 if occupation category is service; = 0 otherwise
OCCUP_AGRI	= 1 if occupation category is agriculture; = 0 otherwise
OCCUP_PROD	= 1 if occupation category is production, transport labourers and others; = 0 otherwise
Industry	
INDS_MANU	= 1 if industry category is manufacturing (including mining and quarrying; electricity, gas and water supply and construction); = 0 otherwise
INDS_HLTH	= 1 if industry category is health and social work; = 0 otherwise
INDS_PUBADMN	= 1 if industry category is public administration; = 0 otherwise
INDS_COMMU	= 1 if industry category is transport, storage and other social services; = 0 otherwise
INDS_FINST	= 1 if industry category is bank, insurance and other financial institutions; = 0 otherwise
INDS_REALEST	= 1 if industry category is real estate, rental and other business activities; = 0 otherwise
INDS_HOTEL	= 1 if industry category is hotel and other business; = 0 otherwise
INDS_WHRET	= 1 if industry category is wholesale/retail; = 0 otherwise
INDS_EDU	= 1 if industry category is education; = 0 otherwise
Training	
TRA_VOC	= 1 if individual has vocational training; = 0 otherwise
TRA_GEN	= 1 if individual has general training; = 0 otherwise
TRA_TECH	= 1 if individual has technical training; = 0 otherwise
Skill	
SKIL	= 1 if individual is skilled; = 0 otherwise
Job status	
W_EMPD	= 1 if individual is wage employed; = 0 otherwise
SELF_EMPD	= 1 if individual is self-employed; = 0 otherwise

C_WKR	= 1 if individual is a casual worker or day labourer; = 0 otherwise
UNEMPS	= The number of months unemployed
Family status	
FEMALE_HHEAD	= No. of female household heads
MALE_HHEAD	= No. of male household heads
NMEN	= No. of males aged between 15 and 64
NFEM	= No. of females aged between 15 and 64
OLD01	= No. of males aged 65 and above
OLD02	= No. of females aged 65 and above
Asset ownership	
ASST01	= 1 if individual has shop/business; = 0 otherwise
ASST02	= 1 if individual has rickshaw/van/pushcart/boat; = 0 otherwise
ASST03	= 1 if individual has sewing machine/shallow machine/tractor; = 0 otherwise
Home ownership	
HRENT	= 1 if household is renting accommodation; = 0 otherwise
HFREE	= 1 if household pays no rent; = 0 otherwise
λ	= Inverse Mills ratio

[a] SSC: Secondary School Certificate; HSC: Higher Secondary School Certificate.

Table A2. OLS wage equation estimates: LFS 1999–2000

	Urban sample		Rural sample	
Variable	Males coef.	Females coef.	Males coef.	Females coef.
Constant	4.0738***	5.2583***	4.8668***	6.5845***
	(0.3440)	(0.3000)	(0.1700)	(0.1853)
LN_MHRS	− 0.7526***	− 0.7003***	− 0.8534***	− 0.9378***
	(0.0673)	(0.0560)	(0.0263)	(0.0212)
AGE	0.0511***	0.0192***	0.0331***	0.0061**
	(0.0050)	(0.0052)	(0.0034)	(0.0029)
AGESQ	− 0.0005***	− 0.0002***	− 0.0003***	− 0.0001*
	(0.0001)	(0.0001)	(0.0000)	(0.0000)
UNEMPS	0.0094**	0.1079	0.0034	0.0016
	(0.0045)	(0.2245)	(0.0153)	(0.0013)
TRA_VOC	− 0.0335	0.1262	0.0347	0.0516
	(0.0803)	(0.1856)	(0.0936)	(0.0590)
TRA_GEN	0.1828***	0.4231***	− 0.1678**	0.1850**
	(0.0529)	(0.1125)	(0.0680)	(0.0861)
TRA_TECH	0.1313***	0.4574***	− 0.0301	− 0.0624
	(0.0504)	(0.1405)	(0.1068)	(0.0700)
SELF_EMPD	1.1733***	0.1315**	1.2529***	0.3642***
	(0.0569)	(0.0619)	(0.0262)	(0.0468)
W_EMPD	0.9483***	0.3849***	0.9156***	0.0413
	(0.0591)	(0.0717)	(0.0384)	(0.0832)
C_WKR	0.5454***	0.1088**	0.5425***	0.1039***
	(0.0616)	(0.0486)	(0.0228)	(0.0325)
PRIM	0.1282***	0.1086***	0.1205***	0.0002
	(0.0241)	(0.0322)	(0.0174)	(0.0177)
SECOND	0.2640***	0.1103**	0.2390***	0.0198

(*Continues*)

Table A2. *Continued*

	Urban sample		Rural sample	
Variable	Males coef.	Females coef.	Males coef.	Females coef.
	(0.0252)	(0.0439)	(0.0205)	(0.0263)
PSECON	0.5030***	0.1677**	0.3731***	0.1054**
	(0.0297)	(0.0756)	(0.0348)	(0.0474)
GRAD	0.7762***	0.6417***	0.6021***	0.2913**
	(0.0433)	(0.1101)	(0.0659)	(0.1427)
OCCUP_PROF	0.0165	0.4935**	0.1023	0.5967***
	(0.0769)	(0.2308)	(0.2250)	(0.1660)
OCCUP_CLERIC	− 0.1115**	− 0.1018	0.1034	0.2932
	(0.0483)	(0.1588)	(0.0996)	(0.2297)
OCCUP_SALES	− 0.0344	− 0.1572	0.1271	− 0.4614**
	(0.0608)	(0.1910)	(0.1082)	(0.2278)
OCCUP_SER	− 0.2963***	− 0.8829***	− 0.0532	− 0.7535***
	(0.0578)	(0.1412)	(0.1026)	(0.1870)
OCCUP_AGRI	− 0.0496	− 0.6038***	0.0015	− 0.5497***
	(0.0763)	(0.1463)	(0.0984)	(0.1544)
OCCUP_PROD	− 0.1514***	− 0.4871***	0.1542*	− 0.7074***
	(0.0496)	(0.1458)	(0.0922)	(0.1818)
INDS_MANU	0.1607**	− 0.1455*	− 0.1038*	0.1221
	(0.0661)	(0.0848)	(0.0559)	(0.0867)
INDS_WHRET	0.0680	− 0.1148	− 0.1060*	− 0.0648
	(0.0697)	(0.1287)	(0.0638)	(0.1658)
INDS_HOTEL	0.0646	0.2264	0.0702	− 0.5538**
	(0.0835)	(0.2848)	(0.0889)	(0.2407)
INDS_COMMU	0.0702	0.4432**	− 0.2298***	0.5509
	(0.0659)	(0.2195)	(0.0606)	(0.4341)
INDS_FINST	0.2963***	0.4208**	0.2250*	0.9905**
	(0.0915)	(0.1779)	(0.1352)	(0.4068)
INDS_REALEST	0.1691	0.4106	0.2448*	1.2897***
	(0.1304)	(0.3341)	(0.1256)	(0.1463)
INDS_PUBADMN	0.2083***	0.4535***	0.2329***	1.4552***
	(0.0720)	(0.1213)	(0.0727)	(0.2091)
INDS_HLTH	0.0964	− 0.1280*	− 0.0703	0.0953
	(0.0669)	(0.0732)	(0.0651)	(0.1031)
INDS_EDU	− 0.1298	− 0.3967**	0.0653	− 0.0753
	(0.0856)	(0.1545)	(0.1094)	(0.2220)
SKIL	0.1579***	0.1808***	0.1811***	0.0595
	(0.0196)	(0.0481)	(0.0157)	(0.0395)
MARRIED	0.1571***	− 0.0628	0.0667***	− 0.0758***
	(0.0304)	(0.0395)	(0.0250)	(0.0221)
Observations	5654	2221	6265	4097
R^2	0.479	0.671	0.545	0.501

Note: Dependent variable: natural logarithm of hourly wages.
Standard errors reported in parentheses and are computed robustly to account for heteroskedasticity.
*Significant at 10%; **significant at 5%; ***significant at 1%.