Dhaka Univ. J. Sci. 62(1): 11-16, 2014 (January)

# Determinants of Earnings and Selection into Labor Market by Controlling Selection Bias in Bangladesh

**Md. Abdus Salam Akanda[1], Maksuda Khanam[2] and Abu Zar Md. Shafiullah[1*]**

[1]*Department of Statistics, Biostatistics and Informatics, Dhaka University, Dhaka-1000, Bangladesh*

[2]*Mirza Abbas Mohila Degree College, Shahjahanpur, Dhaka-1217, Bangladesh*

(Received: 1 October 2012; Accepted for Publication: 17 September 2013)

## Abstract

This paper uses the data of Household Income and Expenditure Survey (HIES) 2005 to estimate the determinants of earnings and selection into labor market in Bangladesh. We deal with the selection bias in earnings by using a maximum likelihood system of equations, and a multinomial selection approach is used modeling for selection into the labor market. By instrumenting years of schooling in both the multinomial selection approach and the earnings equations, we deal with reverse causality between educational attainment and earnings. We find that the estimated parameters of the earnings equation under multinomial selection approach differs from ordinary least square (OLS) estimates and a binomial selection procedure. The estimated parameters that vary the most with those related variables have the strongest impact on multinomial selection into the different labor-market statuses. We also find that workers with higher educational attainment are more likely to search out a salaried employees' job, which non-salaried work is as another to inactivity.

**Keywords:** Employment, earnings, multinomial selection, endogeneity, selection bias, Bangladesh

## I. Introduction

The empirical literature detailing the mechanisms underlying schooling and earnings is limited in the context of Bangladesh, a country characterized by low enrollment rates and education levels, high illiteracy and an outsized inequality between male and female education. This paper utilized standard Mincerian earning methodology to measure the effect on earnings of individual characteristics, like age, academic attainment and marital status among others[1, 2] using the data of Household Income and Expenditure Survey (HIES) 2005. Many methodological extensions have been proposed to deal with the limitations of the conventional Mincerian model: as for examples, the choice of instruments[3] affects the estimated returns to education. Other literatures[4] discuss how well the schooling coefficients of standard Mincer equations approximate the rate of return to education. Empirical evidence is currently available for a number of developing and rising market countries, including Brazil[5] and Nepal[6].

An important extension to the empirical literature is the Heckman selection model, which deals with truncations within the earnings distribution. The data used in this paper are basically based on information on earnings just for salaried workers. Empirical evidence in economic theory shows that ordinary least square (OLS) estimates are inconsistent if the earnings distribution is truncated. The literature contains various methods for handling with multinomial selectivity, as in the case where labor market status cannot be described by simple two alternatives. Multinomial selection models are applied in numerous settings, including the study of self-selection into technical training[7], the size of firm wage differentials[8], and the estimation of demand for electricity[9]. We tend to follow their step modeling selection into labor market as a thrichotomous alternative, which we tend to estimate jointly with the wage equation in a full-information maximum likelihood setting. Our thrichotomous selection is analogous in spirit to existing method[2]. But the basic difference is they estimate two wage equations (for formal and informal workers respectively) with the aim of comparing two selectivity models: ordered probit and multinomial logit[10]. Our strategy also permits to deal with the reverse causality of education, which we assume to be endogenous in both the multinomial selection and the earning equations. Few methodological papers have dealt specifically with the issue of regressors' endogeneity in sample selection models[11] in a non-parametric context, and for a common endogenous dummy[12]. National Socio Economic Survey (Susenas) data were used[13] to estimate a wage equation and the returns to schooling for different groups. The approach is comparable to ours in the way that he acknowledges the matter of selectivity. However he deals with it on the idea of a dichotomous selection rule (i.e., individuals may work as wage earners or not), whereas we show like others[14] that a multinomial selection approach is more appropriate to estimate determinants of earnings and for describing selection into labor market by controlling selection bias.

## II. Problem Statement

To explore how individual characteristics, like age, place of residence and educational attainment have an effect on a worker's labor market status and earnings in a standard Mincerian setting. We face two main problems for dealing this issue.

i) The first is to deal with the fact that earnings data are available only for salaried workers in HIES, but not for the self-employed and household workers, who account for the majority of employment in Bangladesh. Hence, conventional estimation techniques, such as the Heckman[15] binomial selection procedure would be too simple to cover all relevant segmented labor market outcomes like labor market in Bangladesh.

ii) Another drawback is related to the endogeneity of educational attainment in the wage equation. Duflo[16] acknowledges the existence of a selection problem but she does not address the selection problem directly.

*Author for Correspondence. *e-mail*: shafi.ullah@du.ac.bd

To overcome the above mentioned problems, we deal with the selection bias by estimating a full-information maximum likelihood system of equations, wherever wages are observed for wage-earners (salaried employees), and selection into different labor-market statuses is modeled in a multinomial choice setting.

**III. Data**

The cross-sectional data used in this analysis come from the Household Income and Expenditure Survey (HIES) which was administered by Bangladesh Bureau of Statistics (BBS) from January 2005 to December 2005 in Bangladesh[17]. A two stage stratified random sampling technique was followed in drawing samples for HIES 2005 under the framework of Integrated Multi-Purpose Sampling Design (IMPS). The survey included responses from 10,080 households and 49,969 individuals. Details of the data collection process have been described in BBS[17]. Data on earnings and employment are reported in HIES as follows. Each family member belonging to the working age population (those aged 15 years and above) is classified as employed or unemployed depending on his/her activities during the previous month. Employed individuals are classified as wage-earners (salaried workers), self-employed (with or without assistance) or unpaid/family/casual workers. While the HIES data are overall considered to be of good quality, earnings data are collected for employees only, hence excluding the huge number of workers.

**IV. Theoretical Background**

Since earnings data are available only for wage-earners, estimation of a Mincerian equation[18] by OLS would produce biased estimators if, as expected, selection into different job market statuses are correlated with potential determinants of earnings. In an influential paper[1], Heckman proposes a two-step statistical approach, offers a means of correcting for non-randomly selected samples. In the first stage, a model is formulated based on economic theory for the probability of working. The canonical specification for this relationship is a probit regression of the form

$$Prob(D = 1|Z) = \Phi(Z\gamma)$$

where $D$ indicates employment ( $D = 1$ if the respondent is employed and D = 0 otherwise), Z is a vector of explanatory variables, $\gamma$ is a vector of unknown parameters, and $\Phi$ is the cumulative distribution function of the standard normal distribution. Estimation of the model yields results that can be used to predict this employment probability for each individual.

In the second stage, a correction is made for self-selection by incorporating a transformation of these predicted individual probabilities as an additional explanatory variable. The wage equation may be specified by,

$$\omega^* = X\beta + u$$

where $\omega^*$ denotes an underlying wage offer, which is not observed if the respondent does not work. The conditional expectation of wages given the person works is then

$$E[\omega|X, D = 1] = X\beta + E[u|X, D = 1].$$

Under the assumption that the error terms are jointly normal, we have

$$E[\omega|X, D = 1] = X\beta + \rho\sigma_u\lambda(Z\gamma).$$

Where $\rho$ is the correlation between unobserved determinants of propensity to work $\varepsilon$ and unobserved determinants of wage offers u, $\sigma_u$ , is the standard deviation of u, and $\lambda$ is the inverse Mills ratio evaluated at $Z\gamma$. This equation demonstrates Heckman's insight that sample selection can be viewed as a form of omitted-variables bias, as conditional on both X and on $\lambda$ it is as if the sample is randomly selected. The wage equation can be estimated by replacing $\gamma$ with probit estimates from the first stage, constructing the $\lambda$ term, and including it as an additional explanatory variable in linear regression estimation of the wage equation. Since $\sigma_u > 0$ , the coefficient on $\lambda$ can only be zero if $\rho > 0$, so testing the null hypothesis that the coefficient on $\lambda$ is zero is equivalent to testing for sample selectivity.

Rates of return have been estimated using the Mincerian equation[18], which establishes a relationship between the logarithm of wages and age, age square, female, married, female× married, dependency ratio, female× dependency ratio, years of education, household education. The description of the variables follows next section. Under certain assumptions, the parameter linked to years of schooling may be regarded as the rate of return to an additional year of schooling. The return to education may as well be derived from the computation of the internal rate of return from investments in education. To ensure comparability, a full-information maximum likelihood (FIML) technique is used to estimate three models: a single continuous-variable earnings equation; a two-equation system for the binomial selection model, including a wage equation with a continuous censored dependent variable and a selection equation with a binomial dependent variable; and a multiple-equation system for the multinomial model, including a wage equation with a continuous censored dependent variable and separate equations for each alternative labor-market status.

The multinomial selection model, where individuals can chose among M alternatives, can be defined as:

$$y_1 = x\beta_1 + \varepsilon_1 \quad (1)$$

$$y_s^* = z_s\gamma_s + v_s \quad (2)$$

where $s = 1,2,.........,M$ and the wage outcome $y_1^*$ is observed if and only if $y_1^* = max_{j\neq 1} y_j^*$, so that category 1 (salaried work) is chosen. Hence, under the Independence of Irrelevant Alternatives (IIA) hypothesis, Equation (1) reduces to a multinomial logit model. We estimate the two equations for $y_1$ and $y_s^*$ jointly to take account for the correlation between the error terms, which is equivalent to estimating a recursive system of generalized linear models (GLM) framework with a Gaussian error distribution: in the ML-based seemingly unrelated regression model (SUR), all equations are independent, but the underlying errors are jointly normally distributed. In the multinomial selection equation, each choice other than the base alternative $s = 2,3,....,M$ is represented separately by an equation. Since multinomial choice depends on the same set of regressors for all alternatives, we have to impose the IIA condition through constraints on the covariance among the errors of the M-1 equations representing the selection alternatives. In some specifications, we also address the problem of endogeneity of educational attainment by instrumenting the individual's years of schooling. For doing this, we add a reduced-form equation to the FIML system(s) with years of schooling as the dependent variable and all exogenous variables and the instrument as regressors. Since the instrumentation strategy imposes recursiveness, in this case only the second-step coefficients are structural (limited-information maximum likelihood).

## V. Variables and their Measurements

Under binomial selection, individuals are either employed as salaried workers (and hence we can observe their wages), or not salaried workers. In the multinomial selection framework, we assume that workers can select themselves into three labor-market statuses: inactivity, employment as a wage-earner and non-salaried work. The set of exogenous explanatory variables is the same for both selection rules (binomial and multinomial) and includes: age, age squared, an area of residence dummy (rural), a gender dummy (female) and a marital status dummy (married). We also include an interaction term (female$\times$married) and the dependency ratio (computed as the number of household members who are younger than 15 or older than 65 divided by the number of household members aged 15-65) on its own and interacted with gender (female$\times$dependency ratio). Educational attainment is measured as years of education. Finally, we control for the average years of schooling of the other adult household members, which proxies for an individual's socio-economic background. Divisional dummies (the omitted division is Barishal) are included in all regressions. The set of regressors is the same in the wage and selection equations, with the exception of the dependency ratio and its interaction with the gender dummy, which are omitted from the wage equation to fulfill the exclusion restrictions. We aggregate all sectors under four macro labels: "agriculture" (agriculture, fishing), "manufacturing" (mining, manufacturing, electricity, and construction), "trade" (trade, hotels, transports) and "services" (finance, real estate, government, education, health, and other services). As an additional robustness check, we control for the worker's sector of activity (agriculture, manufacturing or services, with trade as the omitted category). To deal with the endogeneity of educational attainment, we instrument years of education, measured as the intensity of school construction in an individual's district of birth and his/her age within this survey. Our instrument survey exposure is equal to survey intensity in the individual's district of birth if he/she was aged 11 or less in 2001, and zero otherwise. The descriptive statistics are reported in Table 1.

**Table 1. Descriptive Statistics**

| Variable | Obs. | Mean | Min. | Max. | S.D. |
|---|---|---|---|---|---|
| log. hourly wage | 24345 | 8.53 | 4.29 | 14 | 0.84 |
| rural | 24345 | 0.53 | 0 | 1 | 0.49 |
| age | 24345 | 37 | 15 | 69 | 13.26 |
| age squared | 24345 | 1324 | 225 | 4761 | 816.05 |
| female | 24345 | 0.5 | 0 | 1 | 0.5 |
| married | 24345 | 0.67 | 0 | 1 | 0.43 |
| dependency ratio | 24345 | 0.32 | 0 | 5 | 0.28 |
| years of education | 24345 | 6.37 | 0 | 16 | 3.94 |
| household education | 24345 | 6.18 | 0 | 16 | 3.31 |
| sector: agriculture | 24345 | 0.42 | 0 | 1 | 0.47 |
| sector: manufacture | 24345 | 0.16 | 0 | 1 | 0.39 |
| sector: services | 24345 | 0.11 | 0 | 1 | 0.26 |
| survey intensity | 24345 | 2.17 | 0.73 | 8.14 | 1.22 |
| survey exposure | 24345 | 1.93 | 0 | 8.14 | 1.34 |

## VI. Determinants of Earnings

The results of the estimation of a Mincerian wage equation done by Stata[19] that are reported in Table 2. The sample includes all individuals aged 15-65 years who worked at least one hour as salaried workers in the previous week. The logarithm of hourly wages is treated as the dependent variable. Nine different specifications are reported: educational attainment is treated as exogenous in the first set of results (Table 2, columns 1 to 3) and is instrumented by program exposure in the second set of results (columns 4 to 6). We control for workers' sector of activity in the third set of results (columns 7 to 9). For each set of results, three specifications are presented: OLS, which ignores the selection bias (column 1, 4 and 7); binomial selection, where inactivity and non-salaried work fall in the same category (column 2, 5 and 8); and the multinomial selection process described above with three different outcomes: salaried work, non-salaried work and inactivity (column 3, 6, 9).

A few parameter estimates differ a great deal across specifications. For instance, the rural dummy is positive signed or insignificant in the OLS and binomial selection specifications, while it is negative and highly significant under multinomial selection, which takes into account for the fact that salaried work, is very infrequent in rural areas. Likewise, the interaction female$\times$married is insignificant under binomial selection, but positive and significant under multinomial selection. The magnitude of the estimated coefficient on the interaction term suggests that being married, which yields a wage premium, offsets in part the negative effect of being female, which is probably related to the fact that very few married women work as salaried workers. It is worth noticing that the regressors whose estimated effects on wage vary the most across specifications are the ones that have the strongest impact on multinomial selection into the labor market: rural, married and female×married. These findings suggest that a binomial rule is too crude for describing selection into the Bangladeshi labor market.

All other coefficients are comparable in sign and magnitude across specifications. For instance, wages rise with educational attainment and age (albeit for age in a nonlinear manner), and women are paid less than men. Socio-economic background, proxied by the average years of schooling of all other adult household members, is positively signed and significant, as expected. Moreover, all else equal, workers in trade are paid less than in the other sectors, while the highest wages are in manufacturing. Our estimate for the returns to education ranges from 9.8 to 11.1%. The estimated coefficients do not change significantly whether educational attainment is instrumented or not, which underscores Duflo's finding that OLS coefficients do not appear to be biased upwards.

### VII. Selection into Labor Market

The results of the selection equation(s) are reported in Table 3. The estimations carried out under binomial selection are reported in column (1), where the estimates refer to the probability of non-salaried work or inactivity (salaried work is the omitted category). Columns (2) and (3) report the multinomial selection results: column (2) refers to the probability of non-salaried work, and column (3) refers to the probability of inactivity (salaried work is the omitted category). In columns (4) to (6) educational attainment is instrumented as described above. Again, column (4) reports the binomial selection coefficients, while columns (5) and (6) refer to the multinomial selection equations. The estimation results shed some light on the differences between non-salaried work and inactivity. The rural dummy is always positive in columns (1) to (3), but the magnitude of the effect is much bigger for non-salaried workers. This suggests that individuals who living in rural areas are on average less educated but have a higher participation rate, are more likely to work in non-salaried jobs than being inactive and to work as salaried employees. This effect is not captured by the binomial selection rule, which averages out non-salaried and inactive workers. However, when educational attainment is instrumented, the rural dummy for inactive workers under multinomial selection is not significant. The effect of age on labor-market status is, as expected, nonlinear. Older workers are more experienced and therefore more likely to work as salaried employees, although the effect is counterbalanced by a quadratic term, which is positively signed.

**Table 2. Wage equation (Dependent Variable: Logarithm of Hourly Wage)**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| **Variables** | **Mincerian** | **Heckman Selection** | **Multinomial Selection** | **Mincerian** | **Heckman Selection** | **Multinomial Selection** | **Mincerian** | **Heckman Selection** | **Multinomial Selection** |
| rural | 0.0327*** | -0.0101 | -0.0268*** | 0.0252** | -0.0112 | -0.0365*** | 0.0247** | -0.0241* | -0.0435*** |
| | (0.007) | (0.010) | (0.008) | (0.011) | (0.013) | (0.012) | (0.009) | (0.013) | (0.012) |
| age | 0.0461*** | 0.0527*** | 0.0385** | 0.0494*** | 0.0559*** | 0.0397*** | 0.0499*** | 0.0565*** | 0.0380*** |
| | (0.016) | (0.018) | (0.019) | (0.019) | (0.018) | (0.014) | (0.019) | (0.018) | (0.014) |
| age square | -0.0002*** | -0.0004*** | -0.0003*** | -0.0003*** | -0.0004*** | -0.0003*** | -0.0003*** | -0.0004*** | -0.0003*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| female | -0.1514*** | -0.1843*** | -0.1451*** | -0.1844*** | -0.1897*** | -0.1402*** | -0.1819*** | -0.1852*** | -0.1376*** |
| | (0.004) | (0.005) | (0.007) | (0.007) | (0.006) | (0.007) | (0.004) | (0.006) | (0.005) |
| married | 0.0757*** | 0.0822*** | 0.0613*** | 0.0731*** | 0.0811*** | 0.0602*** | 0.0637*** | 0.0745*** | 0.0533*** |
| | (0.011) | (0.012) | (0.014) | (0.015) | (0.011) | (0.013) | (0.015) | (0.011) | (0.013) |
| female×married | 0.0591*** | -0.0163 | 0.0734*** | 0.0637*** | -0.0281 | 0.0708*** | 0.0601*** | -0.0394* | 0.0682** |
| | (0.017) | (0.023) | (0.025) | (0.015) | (0.021) | (0.026) | (0.015) | (0.023) | (0.031) |
| years of education | 0.1163*** | 0.1196*** | 0.1171*** | 0.1009*** | 0.1104*** | 0.1133*** | 0.9819*** | 0.1014*** | 0.1111*** |
| | (0.001) | (0.002) | (0.002) | (0.011) | (0.007) | (0.006) | (0.010) | (0.007) | (0.007) |
| household education | 0.0077*** | 0.0083*** | 0.0101*** | 0.0086* | 0.0138*** | 0.0180*** | 0.0099** | 0.0138*** | 0.0180*** |
| | (0.002) | (0.002) | (0.002) | (0.005) | (0.003) | (0.003) | (0.005) | (0.003) | (0.003) |
| sector: agriculture | | | | | | | 0.0547*** | 0.0536*** | 0.0504*** |
| | | | | | | | (0.016) | (0.015) | (0.016) |
| sector: manufacture | | | | | | | 0.1583*** | 0.1571*** | 0.1598*** |
| | | | | | | | (0.006) | (0.005) | (0.007) |
| sector: services | | | | | | | 0.0744*** | 0.0705*** | 0.0801 *** |
| | | | | | | | (0.007) | (0.010) | (0.008) |
| constant | 5.0954*** | 4.4631*** | 4.9142*** | 5.1061*** | 4.4952*** | 5.0097*** | 5.0022*** | 4.4841*** | 4.5468*** |
| | (0.058) | (0.087) | (0.105) | (0.062) | (0.097) | (0.114) | (0.057) | (0.102) | (0.119) |
| divisional dummies | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| education instrumented | NO | NO | NO | YES | YES | YES | YES | YES | YES |

**Table 3. Selection Equation**

| | Heckman Selection | Multinomial Selection | | Heckman Selection | Multinomial Selection | |
|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Variables** | **Non-salaried inactive workers** | **Non-salaried workers** | **Inactive workers** | **Non-salaried inactive workers** | **Non-salaried workers** | **Inactive workers** |
| rural | 0.0827*** | 0.2318*** | 0.0367*** | 0.0801*** | 0.2276*** | 0.0249*** |
| | (0.002) | (0.005) | (0.006) | (0.003) | (0.006) | (0.008) |
| age | -0.0149*** | -0.0092*** | -0.1237*** | -0.0131*** | -0.0084*** | -0.1136*** |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| age square | 0.0005*** | 0.0003*** | 0.0014*** | 0.0005*** | 0.0003*** | 0.0012*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| female | -0.0016 | -0.1291*** | 0.0673*** | -0.0019 | -0.1324*** | 0.0611*** |
| | (0.002) | (0.005) | (0.007) | (0.002) | (0.006) | (0.008) |
| married | -0.0142*** | 0.0594*** | -0.3187*** | -0.0145*** | 0.0583*** | -0.3194*** |
| | (0.004) | (0.007) | (0.007) | (0.004) | (0.006) | (0.007) |
| female × married | 0.1737*** | 0.1849*** | 0.6186*** | 0.1709*** | 0.1815*** | 0.6111*** |
| | (0.002) | (0.007) | (0.006) | (0.002) | (0.007) | (0.006) |
| dependency ratio | 0.0394*** | 0.0313*** | -0.0825*** | 0.0381*** | 0.0316*** | -0.0821*** |
| | (0.006) | (0.008) | (0.017) | (0.006) | (0.008) | (0.017) |
| female × dependency ratio | 0.0112* | 0.0483*** | 0.2591*** | 0.0131** | 0.0495*** | 0.2599*** |
| | (0.006) | (0.012) | (0.019) | (0.006) | (0.014) | (0.021) |
| years of education | -0.0315*** | -0.0498*** | -0.0618*** | -0.0378*** | -0.0583*** | -0.0694*** |
| | (0.001) | (0.002) | (0.003) | (0.002) | (0.005) | (0.007) |
| household education | -0.0003 | -0.0147*** | 0.0178*** | 0.0018* | -0.0012 | 0.0211*** |
| | (0.000) | (0.001) | (0.002) | (0.001) | (0.002) | (0.004) |
| divisional dummies | YES | YES | YES | YES | YES | YES |
| education instrumented | NO | NO | NO | YES | YES | YES |

**Note**: ***p<0.01, **p<0.05, *p<0.1, standard errors in parenthesis

Under the multinomial approach, the female dummy is negatively signed for non-salaried workers but positively signed for inactive individuals, and women tend to choose inactivity much more frequently than men. Marital status also matters. The married dummy is negatively signed under binominal selection, although married individuals are more likely to have non-salaried jobs and less likely to be inactive than single individuals under the multinomial rule. The combined sign and magnitude of the interaction terms suggests that married women have a slightly higher probability of having a non-salaried job than working as salaried workers and a much higher probability of being inactive.

Under multinomial selection, a higher dependency ratio seems to discourage workers from remaining inactive and to push them into non-salaried jobs, while the distinction is not captured under binomial selection. As for the interaction female× dependency ratio, females living in a household with a high dependency ratio are less likely to work as a salaried employee and more likely to be inactive than those living in a low dependency household. The effect is overall positive, but greater in magnitude for non-participants under multinomial selection. The finding is robust to instrumentation of years of schooling.

Educational attainment seems to be a powerful predictor of labor-market outcomes: an additional year of education decreases the probability of non-salaried work and inactivity with respect to salaried work across all specification, and the negative effect is more pronounced when educational attainment is instrumented (columns 4 to 6). Finally, the average years of schooling of the individual's household raises his/her probability to be inactive relative to having a salaried or non-salaried job. This seems to suggest that members of highly educated households tend not to accept low quality non-salaried jobs. The effect is stronger once the endogeneity of educational attainment is taken into account.

### IIX. Conclusion

Our findings ensure the conventional economic theory that investment in education behaves more or less in a similar manner as investment in physical capita. A comparison has been made among the resulting estimates obtained under different selection methods such as multinomial selection approach, standard OLS which ignores the selection bias, and a binomial Heckman selection procedure. We find that estimated parameters obtained under multinomial selection approach differ from OLS and Heckman selection procedure in the estimation of the wage equation. In addition, the estimated effects that vary the most with those related variables have the strongest impact on multinomial selection into different labor market statuses. Our findings also suggest that workers with higher educational attainment are more likely to find a job as salaried employees when the endogeneity of education attainment is taken into account. Overall, our findings throw doubt upon the binomial selection procedure such as OLS and Heckman selection procedure, and suggest that multinomial section approach is more appropriate to estimate determinants of earnings and for describing selection into the labor market in Bangladesh by controlling selection bias.

### References

1. Heckman, J., 1979. Sample Selection Bias as a Specification Error. *Econometrica*, **47(1)**,153-61.
2. Pradhan, M. and A. van Soest, 1995. Formal and informal sector employment in urban areas of Bolivia. *Labour Economics*, **2(3)**, 275-297.
3. Ichino, A. and R. Winter Ebmer, 1999. Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments. *European Economic Review*, **43(4-6)**, 889-901.
4. Bjorklund, A. and C. Kjellstrom, 2002. Estimating the Return to Investments in Education: How Useful is the Standard Mincer Equation? *Economics of Education Review*, **21(3)**, 195-210.
5. Dickerson, A., F. Green and J.S. Arbache, 2001. *Trade Liberalization and the Returns to Education: A Pseudo panel Approach*. Working Paper, No. 0114, University of Kent, 17 Canterbury.
6. Akanda, M. A. S., 2010. Returns to Education in Nepal: Evidence from Living Standard Survey. *The Dhaka University Journal of Science*, **58(2)**, 257-264.
7. Trost, R. and L. Lee, 1984. Technical Training and Earnings: A Polychotomous Choice Model with Selectivity. *Review of Economics and Statistics*, **66(1)**, 151-156.
8. Bunello, G. and A. Colusso, 1998. The employer size-wage effect: evidence from Italy. *Labour Economics*, **5(2)**, 217-230.
9. Dubin, J.A. and D.L. McFadden, 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, **52(2)**, 345-62.
10. Lee, L., 1983. Generalized Econometric Models with Selectivity. *Econometrica*, **51(2)**, 507-12.
11. Das, M., W. Newey and F. Vella, 2003. Nonparametric Estimation of Sample Selection Models. *Review of Economic Studies*, **70(1)**, 33-58.
12. Kim, K., 2006. Sample Selection Models with a Common Dummy Endogenous Regressor in Simultaneous Equations: A Simple Two-Step Estimation. *Economic Letters*, **91(2)**, 80-86.
13. Deolalikar, A., 1993. Gender Differences in the Returns to Schooling and in School Enrolment Rates in Indonesia. *Journal of Human Resources*, **28(4)**, 899-932.
14. Comola, M. and L. de Mello, 2010. *Educational attainment and selection into the labour market: The determinants of employment and earnings in Indonesia*, Working Paper No 2010 – 06.
15. Heckman, J., J. Lance, J. Lochner and P. Todd, 2003. *Fifty Years of Mincer Earnings Regressions*. NBER Working Paper, No. 9732, NBER, Cambridge, MA.
16. Duflo, E., 2001. Schooling and Labour Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, **91(4)**, 795-813.
17. BBS, 2007. *Report of Household Income and Expenditure Survey 2005*. Bangladesh Bureau of Statistics, Planning Division, Ministry of Planning, Government of the People’s Republic of Bangladesh.
18. Mincer, J., 1974. *Schooling, Experience, and Earnings*. NBER, Columbia University Press, New York, N.Y.
19. StataCorp, 2010. *Stata Statistical Software: Release 11*. StataCorp LP, College Station, TX. URL http://www.stata.com/.