The Indian Journal of Labour Economics
https://doi.org/10.1007/s41027-021-00304-2

RESEARCH NOTE

Returns to Computer Use in Bangladesh: An Econometric Analysis

Fahmida Khatun[1] · **Syed Yusuf Saadat**[1]

Accepted: 1 February 2021

Abstract

In the highly competitive labour market of Bangladesh, workers are unable to find the right jobs and employers are also unable to find the right workers. Such skills mismatch is a manifestation that education alone cannot bridge the gaps in the labour market. Computer skills, which can complement education, are becoming increasingly important. Nevertheless, there is a dearth of research regarding the potential benefit of possessing computer skills in the labour market of Bangladesh. This paper intends to fill in this knowledge gap by estimating the returns to computer use using a national labour force survey dataset. By utilising Heckman's two-step selection model, it was found that workers who could use computers earned 17% more than workers who could not use computers. Additionally, the returns to education and experience were estimated to be 3% and 2%, respectively. This implies that computer skills were as valuable as 5 years of education or 8 years of experience, assuming that the returns to all three were constant over time. The results of this study provide empirical evidence in favor of allocating government resources for computer training, and also advocate for individual investment towards learning computer skills.

Keywords Education · Skills · Computers · Bangladesh

JEL Classification I23 · I24 · I26 · O33

✉ Syed Yusuf Saadat
saadat@cpd.org.bd

Fahmida Khatun
fahmida@cpd.org.bd

[1] Centre for Policy Dialogue (CPD), House 40/C, Road No 11 (new), Dhanmondi, Dhaka 1209, Bangladesh

Published online: 08 March 2021

1 Introduction

Bangladesh is now at a crossroads of structural transformation and technological change. In 1991, 60% of workers were employed in the agriculture sector and 26% of workers were employed in the services sector. By 2017, the economy had undergone substantial structural change, and both agriculture and services sectors employed 40% of workers (World Bank, 2017). Such rapid changes mean that there is often a mismatch between the skills possessed by workers and the skills demanded by employers. Such a skills gap is the hallmark of structural unemployment, which is now set to become the next big development challenge for Bangladesh.

Bangladesh has made great strides in the attainment of Sustainable Development Goal (SDG) number 4, education for all. In 2016–2017, 90.6% of the rural population and 93.1% of the urban population aged between 15 and 29 years old were literate (BBS, 2018). However, literacy is merely the ability to read, write, and do arithmetic, which is by no means adequate to secure a job in the contemporary labour market. Only 2.3% of the rural and 8.6% of the urban population in Bangladesh had completed tertiary education in 2016–2017 (BBS, 2018). Thus, there still remains a long way to go before the nation can truly reap the dividends of higher education.

In the meantime, a spectre of impending pervasive structural unemployment is haunting the country. The labour market is changing at an incredible pace, and the employers are becoming increasingly demanding. In this backdrop, there are fears that many workers may be unable to catch up to the prevailing standards, especially because the education and training that they have received have crippled them with obsolete skills. According to the latest survey by the government, 14% of university graduates from rural areas, and 9% of university graduates from urban areas were found to be unemployed in 2016–2017 (BBS, 2018).

In this context, it is clear that the small proportion of workers who possess the right skills will be able to extract considerable rewards from the labour market. One fundamental skill that yields substantial returns is computer skill. Data from the Bangladesh ICT Use and Access Survey 2013 (BBS, 2015b) shows that only 4% of individuals who do not have computers in their household can use computers (Appendix Fig. 5). This means that access to computers is still limited to individuals whose families can afford to buy a computer. It also implies that educational institutions in the country are not adequately equipped with computers, and some students may graduate from school without developing computer skills. Additionally, the data also show that even in households that own a computer, only 33% of individuals can use a computer. Consequently, computer literacy is still very low in Bangladesh.

The hitherto research and existing literature has not addressed the returns to computer skills in Bangladesh. This paper intends to fill in this knowledge gap by calculating the returns to computer use in the context of Bangladesh. This overarching research question is answered using Heckman's two-step estimation technique, whose efficacy has not been fully utilised by other authors in the

past. The findings of this research will provide hard evidence of the importance of computer skills and elucidate the fact that education alone cannot bridge the gaps in the labour market. The remainder of this paper is structured as follows: Section 2 establishes the theoretical framework of the study, Section 3 contains a brief review of some of the relevant literature on the topic, Section 4 describes the data and sampling strategy, Section 5 explains the methodology behind the econometric estimation technique, Section 6 showcases the results of the study, and Section 7 ends the paper with some concluding remarks.

2 Theoretical Framework

The labour market in Bangladesh is driven by the complex and dynamic processes of job creation and job destruction (Fig. 1). When a firm with a vacant job starts looking for a worker for producing, a vacancy is created. On the other hand, when a worker joins a job, that particular job is filled in, and is out of the market. These processes take place in the backdrop of people quitting their jobs voluntarily. Vacancies sometimes lead to new hiring, and those who are unemployed may also be hired. However, the procedure of matching an unemployed worker with a vacancy may involve significant lags, especially if there is a mismatch between the skills of the unemployed worker and the skills required by the job vacancy. The rate at which new vacancies are filled in by unemployed workers can be thought of as a measure of the degree of the efficiency of the labour market. If new vacancies are filled in by unemployed workers relatively quickly, then the labour market is efficient and there is no skills gap. Conversely, if the new vacancies take a long time to be filled in by unemployed workers, it indicates that the labour market is not efficient and is a signal that there may be a skills gap. Such a skills gap is likely to exist in an economy like Bangladesh which is undergoing structural transformation. This is because structural transformation tends to alter the needs of employers, and these changes in job requirements may often take place at a faster rate than the pace at which workers can cope with. As a result, workers end up being

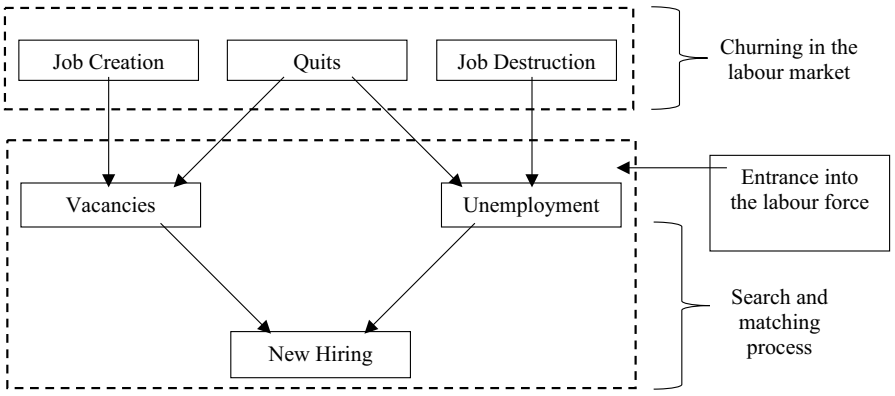

Fig. 1 Simple framework of unemployment and vacancies. Source: Author's illustration based on Bleakley and Fuhrer (1997)

inadequately prepared for the skills that are demanded by their prospective employers. This results in a prolonged job search and matching process whereby employers find it hard to get the right workers and workers find it hard to get the right employers. The unemployment-vacancy nexus is also additionally affected by new entrants into the labour market. If the new entrants possess the appropriate skills, then vacancies will be filled up quickly and the market will become more efficient. However, if the new entrants do not possess the appropriate skills, then the vacancies will persist and unemployment will increase.

The quintessential purpose of the labour market is to match unemployed workers with job vacancies. Following Pissarides (2000), the process of matching can be explained using a matching function defined as

$$mL = m(uL, vL) \tag{1}$$

where m = number of job matches occurring per unit of time; L = size of the labour force; u = unemployment rate; v = number of vacancies.

Alternatively, the matching function may be redefined in terms of the labour market efficiency or tightness, such as

$$q(\theta) = m\left(\frac{u}{v}, 1\right) \tag{2}$$

where θ = labour market efficiency or tightness.

Matching occurs according to a sequential stochastic process given by

$$\frac{m(uL, vL)}{uL} = (\theta)q(\theta). \tag{3}$$

If the rate of job destruction is given by λ and entry into the labour market is blocked, then the number of workers becoming unemployed is given by

$$(1 - u)\lambda L\delta t. \tag{4}$$

where δt = is an infinitesimally small amount of time.

Unemployed workers fill job vacancies and abandon their unemployed status, which can be given as

$$u\theta q(\theta)\lambda L\delta t. \tag{5}$$

From this, it follows that the change in unemployment is

$$\Delta u = (1 - u)\lambda L\delta t - u\theta q(\theta)L\delta t. \tag{6}$$

Since $\Delta u = 0$ for any given level of unemployment, we have

$$(1 - u)\lambda L\delta t = u\theta q(\theta)L\delta t. \tag{7}$$

This can be rewritten as

$$u = \frac{\lambda}{(\lambda + \theta q(\theta))}. \tag{8}$$

When this expression is plotted in an unemployment-vacancy coordinate space, the resulting function is known as the Beveridge curve (Fig. 2). The Beveridge curve is negatively sloped and convex to the origin and illustrates a nonlinear inverse relationship between the level of unemployment and the number of vacancies. The Beveridge curve is named after British Member of Parliament, William Beveridge, who discussed the nexus between unemployment and vacancies in his magnum opus (Beveridge, 1944). Following the line of reasoning established by Beveridge, several economists have worked on the unemployment–vacancy relationship and formalised their ideas (Blanchard et al., 1989; Bowden, 1980; Dow and Dicks-Mireaux, 1958; Hansen, 1970; Holt and David, 1966; Pissarides, 1986; Petrongolo and Pissarides, 2001; Yashiv, 2006).

Beveridge curves are useful for understanding the efficiency or tightness of a labour market. If a Beveridge curve in the unemployment–vacancy coordinate space is close to the origin, then it indicates that labour market efficiency or tightness is high. This is because for a given number of vacancies, V^*, the level of unemployment, U_1, is low. The pace of matching is relatively fast since the skills required by the job vacancies are similar to the skills of unemployed workers. Thus new job vacancies are quickly filled up by unemployed workers. This situation can be likened to the case of the market for unskilled labour who do not possess computer skills. On the other hand, if the Beveridge curve is far away from the origin, then it demonstrates that the level of labour market efficiency or tightness is low. Therefore, for the same corresponding number of vacancies, V^*, the level of unemployment, U_2, will be much higher. In this case, the speed of matching is relatively slow since the skills required by the job vacancies are not comparable to the skills of the unemployed workers. These circumstances represent the case of the market for skilled labour who possess computer skills.

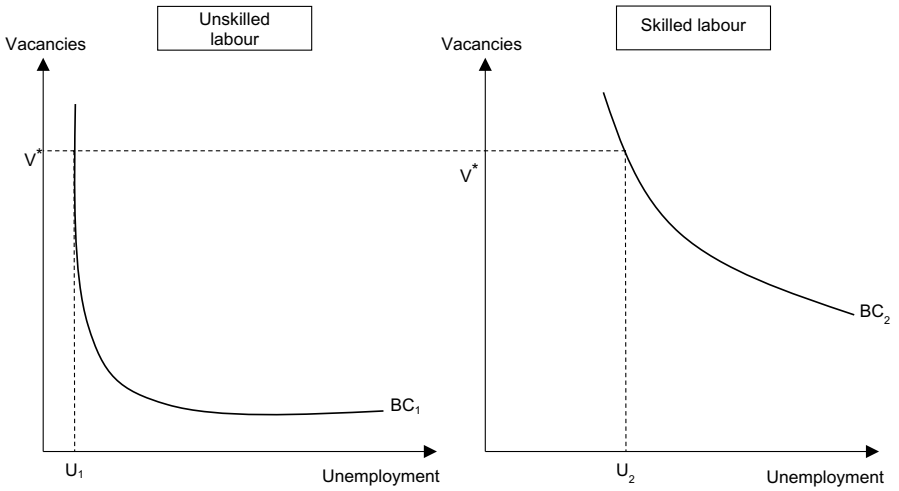

Fig. 2 Beveridge curves for unskilled and skilled labour markets. Source: Authors' illustration based on Beveridge (1944)

Since few workers possess the appropriate computer skills that are demanded by employers, the return to computer skills is high in countries such as Bangladesh. The high return to computer skills can be explained by the fact that the rate at which the demand for computer skilled workers is increasing is faster than the rate at which new computer skilled workers are entering the labour market. Since the increase in demand for computer skilled workers is higher than the increase in supply of computer skilled workers, the returns to computer skills are high (Fig. 3).

3 Literature Review

The pioneering study on the returns to computer use was carried out by Alan Krueger (Krueger, 1993). Krueger observed that the proliferation of computers was transforming the wage structure in the labour market in the USA. He hypothesised that computing technology is a complement to skilled workers, and therefore, the computer revolution has been instrumental in creating skill-based earnings differentials. He validated this claim by showing that increase in computer use at an occupational level is positively linked to increase in wages in the occupation. Using data of the USA from the Current Population Surveys of October 1984 and October 1989, Krueger found that workers who use computers at work earn 10 to 15% more than workers who do use computers at work. Krueger advocated that as much as 40% of the returns to education could be attributed to the propagation of computer use. Including a computer use dummy in the conventional Mincer-type wage model (Mincer, 1974) tends to reduce the returns to education. Additionally, using a probit

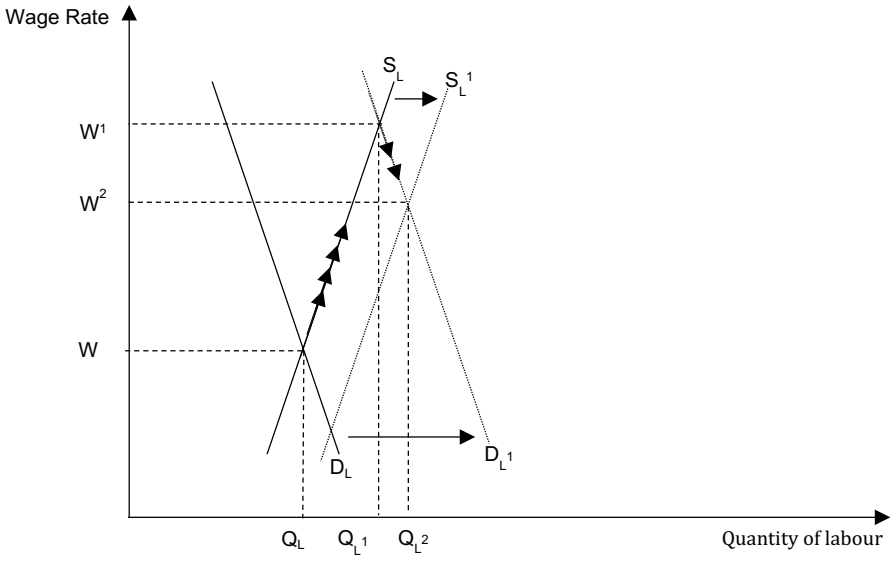

Fig. 3 Market for workers with computer skills. Source: Authors' illustration

model, Krueger found that the probability of using computers at work increases with education, which meant that workers who were relatively more educated were more likely to use computers. On the other hand, the relationship between experience and computer use was found to be nonlinear, with the probability of computer use increasing with experience initially, but then decreasing after reaching 15 years of experience. Therefore, it may be implied that computer skills are linked with education and experience, and collectively all of these three factors have an influence on wages.

There may be speculation as to whether wage differentials among differently skilled workers were due to computer use per se, or simply due to the fact that more skilled workers were more likely to use computers. In order to address this issue, Krueger meticulously checked the robustness of his claims using several different approaches. His investigation suggested that computer use at work, and not simply general computer use, was responsible for the wage differential. From this, it can be implied that employers value computer skills which are pertinent to specific occupations. In fact, Krueger's own informal surveys revealed that employers were willing to pay more for workers with certain computer skills, and that secretaries obtained salary increments after attaining word processing skills. Krueger followed up on his research three years later with another paper and discovered that the returns to computer use in the USA was 17% in 1984, 19% 1989 and 20% in 1993 (Autor et al., 1996).

Following in the footsteps of Alan Krueger, other economists have attempted to estimate the returns to computer use. The most significant of such studies was carried out by John E. DiNardo and Jörn-Steffen Pischke. Using cross-sectional data of the Qualification and Career Survey of West Germany from 1979, 1985–1986, and 1991–1992, DiNardo and Pischke found that the returns to computer use were 11% in 1979, 16% in 1985–1986, and 17% in 1991–1992 (DiNardo and Pischke, 1997). These results were similar to Krueger's original estimates. However, the returns to computer use in Germany were lower than the 17% to 20% returns to computer use in the USA during 1984 to 1993. Nonetheless, DiNardo and Pischke were not content with simply estimating the returns to computer use. They took full advantage of the high-quality German dataset at their disposal and estimated the returns to 'white collar tools' such as pens, pencils, and desk calculators, as well as 'blue collar tools' such as hammers, screwdrivers, and paint brushes. Their empirical investigation revealed that workers who used white collar tools and did sedentary tasks earned 9% to 14% more than workers who did not use such tools, or worked on their feet. On the other hand, workers who used blue collar tools earned 9% to 11% less than workers who did not use such tools. Hence, whilst the use of white collar tools yielded a wage premium, on the contrary, the use of blue collar tools resulted in a wage penalty (DiNardo and Pischke, 1997).

Using their findings as a premise, DiNardo and Pischke argued that there is a possibility of significant selection bias, since white collar tools were only used by workers who were highly paid. Such workers may have been more skilled in the first place, which was why they were rewarded with high wages. Therefore, white collar tools may not accrue returns on their own accord, but rather represent the dexterity of their adept users.

Subsequent studies in the literature have attempted to incorporate some of the important lessons from DiNardo and Pischke's research. Through the use of an endogenous switching model, it was found that computer skills, such as familiarity with software packages and programming languages, were rewarded with a 13% to 25% return in the labour market in the USA in 1980 (Hamilton, 1997).

Studies on the returns to computer use in other countries have found results which are similar to Krueger's original findings. One such interesting study used data from the National Child Development Study of the UK, which included information on British birth cohorts born in 1958 (Dolton and Makepeace, 2004). This dataset included information on computer use over a decade which had particularly experienced a widespread proliferation of computers. Using a variety of methods, Dolton and Makepeace found that the returns to computer use during the period 1991–2000 were 14% for men and 9% for women. Men who used computers only in the year 2000 earned 5% more, whilst women who used computers only in the year 2000 earned 14% more.

In a study employing panel data from the Canadian Workplace and Employee Survey from 1999 to 2002, it was found that the returns to computer use were 4% (Dostie et al., 2006). This finding is substantially lower than previous studies in the literature, since this research corrects for worker selection bias and workplace unobserved heterogeneity.

Data from the Labour Force Survey 2013 and ICT Use and Access 2013 shows that individuals who could use computers earned more than individuals who could not (BBS, 2015gb, c). Average monthly wage for computer literate individuals was higher in every division as also on national level. The wage differential was the highest in Barishal, where computer literate individuals earned 38.68% higher monthly wages on average than computer illiterate individuals (CPD, 2018). On the other hand, the wage differential was the lowest in Khulna where computer literate individuals earned 22.53% higher monthly wages on average than computer illiterate individuals (Table 1) (CPD, 2018).

These findings provide ample evidence to the effect that computer skills are a necessary complement to general education. Such skill could play a key role in bridging the supply–demand gaps in the Bangladesh labour market.

4 Data and Sampling Strategy

This study is based on the cross-sectional data collected by the Bangladesh Bureau of Statistics as part of its Labour Force Survey (LFS) 2013 (BBS, 2015c). All the variables are drawn from this dataset, except the data for the variable CPI, which was collected from the Annual Report 2013 of Bangladesh Bank (Bangladesh Bank, 2013). The LFS 2013 dataset was chosen since it is the last national dataset having information on computer use by workers.

In order to form the sampling frame for the survey, 1512 primary sampling units (primary sampling units) or enumeration areas (EAs) were created. Each PSU/EA was defined as a geographical area of contiguous non-overlapping land with identifiable boundaries. The country was divided into regions based on seven divisions:

Table 1 Average monthly wage by computer use (in BDT)

Area	Never used computer (in BDT)	Used computer (in BDT)	Wage differential (in %)
Barishal	11,343	15,731	38.68
Chattogram	11,217	14,358	28.00
Dhaka	10,881	14,788	35.91
Khulna	10,709	13,122	22.53
Rajshahi	10,620	13,425	26.41
Rangpur	10,264	13,551	32.02
Sylhet	10,320	13,401	29.85
Rural	10,287	12,925	25.64
Urban	11,254	14,434	28.26
National	10,812	14,162	30.98

Source: Authors' calculations based on Labour Force 2013 and ICT Use and Access Survey data (BBS, 2015b, c)

Wage differential is calculated as the percentage difference in the average monthly wages of individuals who have used computers and the average monthly wage of individuals who have never used computers; This table also appears in (CPD, 2018) where the authors of the present study were contributors

Barisal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur, and Sylhet. Each division was further disaggregated into three kinds of localities: City Corporation, Urban, and Rural. Thus, the country was stratified into 21 different strata. Households were classified as "strong", "semi-strong", and "not strong", based on the quality of the house building materials. Each PSU/EA had approximately 80–120 households, so that the total sample included roughly 36,242 households. The PSUs/EAs were scattered across every corner of the country, and encompassed every socio-economic class of people in order to obtain a representative sample of the entire population of Bangladesh.

The sampling strategy used was multi-stage cluster random sampling. The method was as follows. In stage 1, there is random selection of 1512 PSUs/EAs from all of the 64 districts and 21 regional strata, such that households residing in all three categories of houses are covered. In stage 2, there is systematic random sampling of clusters of 24 households from each of the 1512 PSUs/EAs with strict prohibition of replacement of non-responding households. A total of exactly 36,242 households are selected at this stage.

The sample size for sub-populations was calculated using the formula

$$n = \left[\frac{(1-p)}{p} * \left(\frac{z\left(\frac{\alpha}{2}\right)}{r} \right)^2 \right] * \text{deff.} \tag{9}$$

where p = apriori proportion of the required characteristics in the population; $z(\alpha/2)$ = value of the standard normal variate allowing 100(1-α)%p confidence;

r = rate of allowable margin of error; N = population size; deff = design effect used in complex surveys using multistage cluster sampling; and assuming, $\alpha = 0.005$; deff = 2; $p = 0.046$ (from Labour Force Survey, 2010).

The main dependent variable used in the model was the natural log of weekly wages in cash and kind from both primary and secondary jobs. Education, measured by years of schooling, experience, measured as the age minus education minus 6, and experience squared, measured as experience multiplied by experience, were used as independent variables. Computer skills are difficult to measure, and no such data exists for the labour market of Bangladesh. Therefore, a computer-use dummy variable was used a proxy for computer skills. This dummy variable had a value of 1 if the worker ever used a computer in their lifetime, and 0 otherwise. Use of such a proxy is undoubtedly a big compromise. However, based on the data available, there is no better proxy of computer skills. Hours of work, assets, marital status, number of children, and a vector of goods prices are included as additional independent variables in labour force participation equation to correct for selection bias (Table 2).

Table 2 Variables used in the econometric analysis

Variable	Definition	
Lnwage	Natural log of weekly wage in cash and kind from both primary and secondary job	
Education	Years of schooling up to 12 years	
	1 = class-I	7 = class-VII
	2 = class-II	8 = class-VIII
	3 = class-III	9 = class-IX
	4 = class-IV	10 = class X
	5 = class-V	11 = SSC
	6 = class-VI	12 = HSC
Experience	Potential experience; (experience = [age] – [education] – [6])	
Experience^2	Squared experience term; (experience^2 = experience*experience)	
Computer	Computer use dummy;	
	Computer = 1 if ever used computer	
	Computer = 0 if never used computer	
Hours	Total number of hours worked per week at both primary and secondary job	
Assets	Total amount of land owned by households, measured in acres	
Married	Marital status dummy;	
	Married = 1 if currently married	
	Married = 0 if currently not married	
Children	Number of children aged less than 6 years	
CPI*	Consumer price index (CPI);	
	CPI = 183.90 if rural, CPI = 177.71 if urban	

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

*Data from Bangladesh Bank

5 Methodology

The econometric estimation techniques used in this paper are based on the work of James Heckman (Heckman, 1974) (Heckman, 1976) (Heckman, 1979). Hence the following exposition draws heavily from Heckman's papers, with minor changes that have been highlighted. For a more detailed analysis, refer to Heckman's original papers.

For estimating the model, the following functional form is used:

$$\ln(W_i^*) = \beta_0 + \beta_1 H_i + \beta_2 A_i + \beta_3 M_i + \beta_4 K_i + \beta_5 P_i + u_i \tag{10}$$

$$\ln(W_i) = b_0 + b_1 S_i + b_2 E_i + b_3 E_i^2 + b_4 C_i + \varepsilon_i \tag{11}$$

where $\ln(W_i)$ = natural log of weekly wage; $\ln(W_i^*)$ = latent wage or employment indicator, $\ln(W_i^*) = 1$ if $\ln(W_i) > 0$; S_i = years of schooling; E_i = years of potential experience; E_i^2 = years of potential experience squared; C_i = computer use dummy, equal to 1 if ever used computer and 0 otherwise; H_i = total hours worked per week; A_i = total amount of land owned by household, measured in acres; M_i = marital status dummy, equal to 1 if currently married and 0 otherwise; K_i = number of children aged less than 6 years; P_i = consumer price index, CPI = 183.90 if rural and CPI = 177.71 if urban.

It is assumed that ε_i and u_i are jointly normally distributed, each with zero mean, and that there is correlation between ε_i and u_i. It is also assumed that the error terms are not correlated with lagged values of themselves, so there is no autocorrelation. Equation (11) suffers from unobserved heterogeneity or the problem of omitted variables. This means that there are unobserved omitted variables which affect the independent variables of these equations. The effect of these unobserved variables is captured through the error terms, and so the errors of Eq. (11) are correlated with the independent variables. The underlying reason behind this unobserved heterogeneity or omitted variable bias is the fact that the samples used for estimating these equations were not randomly selected. Market wages are only observed for workers who are working. However, the worker's decision to work is not a random decision. It is a rational decision and a conscious choice. Hence, workers who choose to work self-select themselves into the sample. The distribution of market wages is censored at the reservation wage threshold. This can be illustrated using a simple example. Suppose education is the only variable that determines offered wage or market wage. Assume that education has no effect on reservation wage. If we now plot a graph of market wage against education, then we get the diagram in Fig. 4.

The vertical axis represents wage, and the horizontal axis represents level of education. The horizontal line dividing the plot space is the reservation wage. The three curves represent the distributions of market wages, at different levels of education. Note that the curve on the left represents workers with a comparatively low level of education and is thus skewed in a manner that most workers earn a relatively low wage. Alternatively, the curve on the right represents workers with comparatively high level of education and is thus skewed in a way that most workers earn a relatively high wage. Note also that the distribution of market wage runs from the

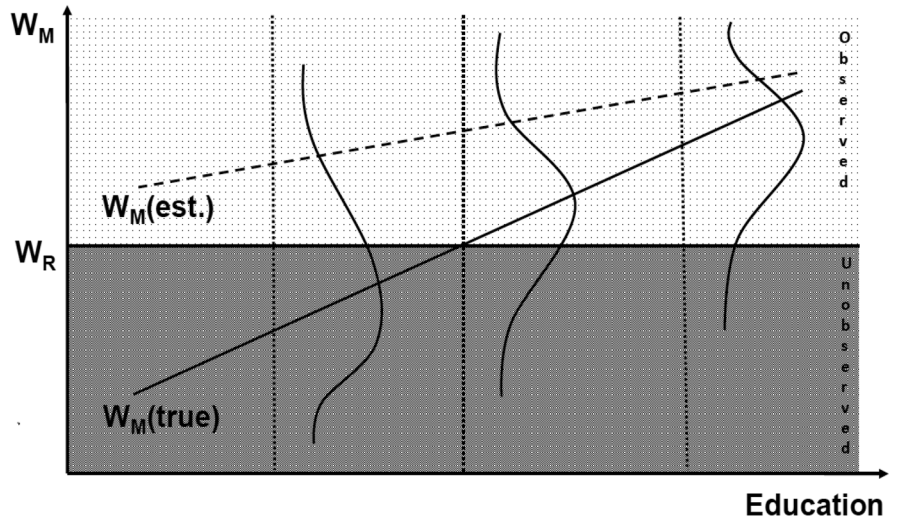

Fig. 4 Sample selection bias. Source: *Author's illustration based on Smith (2014) and Begum et al., (2018)

top to the bottom, going across the reservation wage boundary in the middle. Since market wages are only observed for workers who are working, we can only see the part of the market wage distribution that lies above the horizontal reservation wage line. Thus, our sample is a censored sample, with the censoring occurring at the reservation wage. Any analysis which fails to account for the unobserved part of the market wage distribution is bound to be biased. In the diagram below, we can see that the estimated market wage line W_M(est) is much flatter than the true market wage line W_M(true). This means that the coefficients from an ordinary least squares regression model will be lower than the coefficients from a Heckman selection bias corrected model. Hence, the rate of return to education will be underestimated in a model which is not adjusted for sample selection bias.

Hence, it is not possible to obtain unbiased or consistent estimates of these equations using ordinary least squares (OLS) estimation techniques. This is because the strict exogeneity assumption of the OLS model states that the errors should be exogenous of the independent variables. Violation of the strict exogeneity assumption implies that the unconditional mean of the error term is not zero, the independent variables are not orthogonal to the errors for all observations, and the independent variables and errors are not uncorrelated for all observations.

Therefore, if the strict exogeneity assumption is violated, then OLS will produce biased estimates. However, under such circumstances, estimating Eq. (11) is still possible. Heckman suggested that statistical models which feature truncation, sample selection, or limited dependent variables share a common structure which makes them vulnerable to unobserved heterogeneity or an omitted variable bias (Heckman, 1976). If the unobserved heterogeneity can be modelled separately, and the resulting information can be incorporated into the main model, then the problem can be resolved. Subsequently, Heckman (1979) proposed that selection bias due to using non-randomly

selected samples could be viewed as an ordinary specification error. Heckman proposed that the specification of the original biased model could be improved by using the estimated values of the omitted variables as additional regressors. By doing so, the model could be estimated using ordinary least squares, without violating the strict exogeneity assumption. Heckman outlined an ingenious two-step estimation technique to correct sample selection bias (Heckman, 1979). This is the technique used for econometric estimation in this paper.

In the first step, the factors that influence a worker's decision to work is modelled using a probit model. The general form of the sample likelihood function for this probit analysis is:

$$\mathcal{L} = \prod_{i=1}^{T} \left[F\left(\phi_i\right)\right]^{1-d_i} [1 - F\left(\phi_i\right)] d_i. \tag{12}$$

where d is a random variable, which is equal to one if the dependent variable is observed and equal to zero if the dependent variable is not observed. Suppose we use a sample of T workers, K of who work and T-K who do not work. Then, in the case of our model, the aforementioned likelihood function becomes:

$$\mathcal{L} = \prod_{i=1}^{K} j(h_i, \ln(W_i)|\left(W_i > W_i^*\right)_{h=0}).pr\left(\left[W_i > W_i^*\right]_{h=0}\right) \times \prod_{i=K+1}^{T} pr\left(\left[W_i < W_i^*\right]_{h=0}\right). \tag{13}$$

From this, we may extract an important ratio called the inverse Mills ratio, or the hazard function, denoted as λ. The inverse Mills ratio is the ratio of the standard normal probability distribution function of the selection equation to the standard normal cumulative distribution function of the selection equation.

$$\lambda_i = \frac{f\left(\phi_i\right)}{1 - F\left(\phi_i\right)}. \tag{14}$$

where f is the standard normal probability distribution function of the selection equation and F is the standard normal cumulative distribution function of the selection equation. The inverse Mills ratio is a monotone decreasing function of the probability that an observation is selected into the sample. The denominator of the inverse Mills ratio is the probability that an observation has data for the dependent variable. The lower the probability that an observation has data for the dependent variable, the greater the value of the inverse Mills ratio for that observation. We can therefore say that,

$$\frac{\partial \lambda}{\partial \phi_i} > 0$$

and

$$\lim_{\phi_i \to \infty} \lambda_i = \infty, \lim_{\phi_i \to -\infty} \lambda_i = 0$$

For our model, the inverse Mills ratio can be defined as:

$$\lambda = j(h_i, \ln(W_i)|\left(W_i^* < W_i\right)_{h=0} = \frac{n(h_i, \ln(W_i))}{pr\left(\left[W_i > W_i^*\right]_{h=0}\right)} \because \varepsilon_i, u_i \sim N(0). \tag{15}$$

By using this inverse Mills ratio in our original likelihood function, we can further simplify it to:

$$\mathcal{L} = \prod_{i=1}^{K} n(h_i, \ln\left(W_i\right)) \prod_{i=K+1}^{T} pr\left(\left[W_i < W_i^*\right]_{h=0}\right). \tag{16}$$

We now maximise this likelihood function with respect to the parameters of the model, including the variances and covariances of the errors in Eqs. (10) and (11) to get consistent, asymptotically unbiased, and efficient parameter estimates which are asymptotically normally distributed. Thus, our selection bias corrected model now becomes:

$$\ln(W_i) = b_0 + b_1 S_i + b_2 E_i + b_3 \lambda_i + u_i. \tag{17}$$

Augmenting the basic wage model with the squared experience term and the computer use dummy variable, we get

$$\ln(W_i) = b_0 + b_1 S_i + b_2 E_i + b_3 E_i^2 + b_4 C_i + b_5 \lambda_i + u_i. \tag{18}$$

where λ is the inverse Mills ratio.

6 Results

The empirical analysis is commenced by estimating Eq. (11), both with and without the computer dummy, using ordinary least squares. The results of the ordinary least squares estimation show that, on average, workers who used computers earned 19% more than workers who did not use computers, provided all other factors were held constant (Table 3). Interestingly, inclusion of the computer use dummy reduced the returns to education, indicating that a portion of the returns to education were actually due to the returns to computer use. Nevertheless, both models estimated with ordinary least squares exhibited very low values of adjusted R squared. For example, in the model with the computer use dummy, the adjusted R squared was only 15%. This means that only 15% of the variation in the dependent variable could be explained by the variation in the independent variables of the model. This implies that there are other variables that affect the dependent variable which we failed to include in our model. As a result of this, the effect of the unobserved omitted variables is being captured through the error term. Owing to such a poor goodness of fit, the ordinary least squares estimates were put to a series of post-estimation diagnostic tests to investigate the root cause of the problem.

To check if the model is correctly specified, a Ramsey Regression Specification Error Test (RESET) was conducted (Table 4). Note that RESET does not

Table 3 Results of ordinary least squares model estimation

	Model without computer	Model with computer
Variable	(Natural log of) Wage	(Natural log of) Wage
Education	0.0345818***	0.0296701***
	(0.0008388)	(0.00087)
Experience	0.020699***	0.0213602***
	(0.0006296)	(0.0006244)
Experience^2	-0.0003208***	-0.0003276***
	(0.0000124)	(0.0000123)
Computer		0.1928568***
		(0.010153)
Constant	7.29877***	7.312575***
	(0.0095537)	(0.0094888)
F statistic	928.38	800.22
Probability > F	0.0000	0.0000
R-squared	0.1324	0.1492
Adjusted R-squared	0.1322	0.1490

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

(i) Standard errors in parentheses; (ii) ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table 4 Results of Ramsey reset specification test

Test Statistic	Model without computer	Model with computer
F statistic	588.80	516.56
	(0.0000)	0.0000

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

H_0: there is no omitted variable, H_A: there is at least one omitted variable; Probability values in parentheses

conclusively indicate omitted variable bias, but simply shows that model is misspecified. Since the probability > F is less than 0.05, the null hypothesis that the model is correctly specified is rejected.

The investigation of the problem of omitted variables is continued using the Link test. Link test is based on the idea that if a regression is properly specified, one should not be able to find any additional independent variables that are significant, except by chance. Link test creates two new variables, the variable of prediction, and the variable of squared prediction. The model is then refit using these two variables as predictors. The prediction variable should be significant since it is the predicted value. On the other hand, squared prediction should not be significant, because if the model is specified correctly, the squared predictions should not have much explanatory power. That is, the squared prediction should

not be a significant predictor if the model is specified correctly. Since the squared prediction does have explanatory power in this case, the model specification is not appropriate (Table 5).

Now that a case in favor of unobserved heterogeneity or omitted variable bias has been developed, the analysis is carried forward by checking if the model can satisfy some of the other assumptions of the ordinary least squares model. In this regard, the variance inflation factor (VIF) of the dependent variables of the model is checked to see if there are any perfect linear relationships between them. VIF measures the linear association between an independent variable and all other independent variables. The decision rule for VIF is as follows: i) if VIF > 10: perfect multicollinearity is highly likely, ii) if 5 < VIF < 10: perfect multicollinearity is somewhat likely, iii) if 0 < VIF < 5: perfect multicollinearity is unlikely.

Multicollinearity is not a problem of existence, but a problem of severity. Since the VIF for experience and experience squared are close to 10, it can be concluded that perfect multicollinearity is somewhat likely in this model (Table 6). Next, it is checked if the variance of the error term in our model is constant or homoscedastic.

Heteroscedasticity is tested formally using the Breusch–Pagan and Cook–Weisberg test for heteroscedasticity (Table 7). If the $p < 0.05$, then the null hypothesis of no heteroscedasticity is rejected. In this case, since $p < 0.05$, the null hypothesis of no heteroscedasticity is rejected. The Breusch–Pagan and Cook–Weisberg test for heteroscedasticity assumes that the heteroscedasticity is a linear function of the independent variables. Failing to find heteroscedasticity with the Breusch–Pagan and Cook–Weisberg test does not rule out the possibility of heteroscedasticity. A more robust test for heteroscedasticity is the White test (Table 8). The White test allows the heteroscedasticity process to be a function of one or more independent variables. It allows the independent variable to have a nonlinear and interactive effect on the error variance.

Table 5 Results of link test

	Model without computer	Model with computer
Variable	(Natural log of) Wage	(Natural log of) Wage
Prediction	51.53402***	32.39009***
	(1.67844)	(1.341881)
Squared prediction	−3.249214***	−2.013056***
	(0.107913)	(0.086048)
Constant	−196.4245***	−122.3259***
	(6.525249)	(5.230572)
F statistic	1915.11	1922.25
Probability > F	0.0000	0.0000
R-squared	0.1734	0.1740
Adjusted R-squared	0.1733	0.1739

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

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

Table 6 Results of variance inflation factor test

	Model without computer		Model with computer	
Variable	VIF	1/VIF	VIF	1/VIF
Education	1.02	0.985135	1.11	0.898111
Experience	10.00	0.100022	10.03	0.099712
Experience^2	9.97	0.100328	9.98	0.100242
Computer			1.11	0.897545
Mean VIF	6.99	5.56		

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

VIF > 10: perfect multicollinearity is highly likely; 5 < VIF < 10: perfect multicollinearity is somewhat likely; 0 < VIF < 5: perfect multicollinearity is unlikely

Table 7 Results from Breusch-Pagan and Cook-Weisberg test

Test Statistic	Model without computer	Model with computer
Chi-square statistic	107.87 (0.0000)	118.51 (0.0000)
F statistic	55.13 (0.0000)	60.59 (0.0000)

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

H_0: errors have are homoscedastic, H_A: errors are not homoscedastic; Probability values in parentheses

Table 8 Results from White's test

Test Statistic	Model without computer	Model with computer
Chi-square statistic	958.62 (0.0000)	962.88 (0.0000)

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

H_0: errors have are homoscedastic, H_A: errors are not homoscedastic; Probability values in parentheses

If the $p < 0.05$, then the null hypothesis of no heteroscedasticity is rejected. In this case, since $p < 0.05$, the null hypothesis of no heteroscedasticity is rejected. Thus, there is unrestricted heteroscedasticity in the model estimated.

If the conditional distribution of the errors is jointly normal, then OLS becomes the best unbiased estimator among both linear and nonlinear estimators. Thus, it is of interest to check if the errors of an estimated model are jointly normally distributed. Therefore, the Shapiro–Wilk test for normality of data is conducted (Table 9). If p value < 0.05, then the null hypothesis that errors are normal is rejected. If p value > 0.05, then the null hypothesis that errors are normal cannot be rejected. Since the p value < 0.05, it can be concluded that the errors are not normally distributed.

Now that it has been conclusively established that there are significant problems in the OLS model, the model is estimated using the selection bias-corrected Heckman model (Table 10). The Heckman model is estimated in two steps. In the

Table 9 Results from Shapiro–Wilk test for normal data

	Model without computer	Model with computer
Variable	Residual	Residual
Shapiro–Wilk test statistic (W)	0.98491	0.98579
Covariance matrix (V)	125.125	117.833
Z statistic	13.125	12.962
Probability > z	0.00000	0.00000

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

H_0: errors are normally distributed, H_A: errors are not normally distributed

first step, a probit model of observing the dependent variable is estimated. In this case, this is the probability of a worker working. This is estimated using Eq. (10) which is the labour force participation equation. In other words, this is the selection equation which models the probability of workers self-selecting themselves in the sample. Using the estimates from the probit model, the inverse Mills ratio is calculated. This inverse Mills ratio is used as an additional regressor in Eq. (18) to estimate the selection bias-corrected Heckman model.

The results from the Heckman two-step selection model estimation show that, on average, workers who had used computers earned 17% more than workers who did not use computers, provided all other factors are held constant. Moreover, it is found that, on average, for every additional year of education, wage increases by 3%, provided all other factors are held constant. The coefficient for experience is positive whilst the coefficient for experience squared is negative. This indicates that at low levels of experience, the relationship between experience and earnings is positive, but at high levels of experience, the relationship between experience and earnings is negative. The threshold level of experience that maximises earnings can be calculated as follows.

Using the value of the coefficients from the Heckman model with computer use, the peak of the wage–experience curve can be calculated with simple differential calculus. The basic idea is to find the level of experience which maximises wage. Let us recall that our model specification was as follows:

$$\ln(W_i) = b_0 + b_1 S_i + b_2 E_i + b_3 E_i^2 + b_4 C_i + b_5 \lambda_i + u_i. \tag{18}$$

Substituting the coefficients from the Heckman model with computer use, we get

$$\widehat{\ln\left(\mathrm{W_i}\right)} = 7.4678 + 0.0303\mathrm{S_i} + 0.0203\mathrm{E_i} + (-0.0003)\mathrm{E_i^2} + 0.1703\mathrm{C_i} + (-0.1559)\lambda_{\mathrm{i}} + \mathrm{u_i}$$

Differentiating the equation with respect to experience, we get

$$\frac{\partial W}{\partial E} = 0.0203 - 0.0006E$$

At the turning point, the first derivative is zero, so we get

Table 10 Results of Heckman two-step selection model estimation

	Model without computer		Model with computer	
Regression	Selection equation	Main equation	Selection equation	Main equation
Variable	(Natural log of) Wage	(Natural log of) Wage	(Natural log of) Wage	(Natural log of) Wage
Education		0.0343627***		0.0302509***
		(0.0010397)		(0.0010727)
Experience		0.0198216***		0.0203444***
		(0.0007809)		(0.0007734)
$Experience^2$		−0.0002861***		−0.0002924***
		(0.0000151)		(0.000015)
Computer				0.1702658***
				(0.0123785)
Hours	0.0107224***		0.0107224***	
	(0.0005879)		(0.0005879)	
Assets	−0.0003627***		−0.0003627***	
	(0.0000306)		(0.0000306)	
Married	−0.3196878***		−0.3196878***	
	(0.0171752)		(0.0171752)	
Children	−0.0234653**		−0.0234653**	
	(0.00963)		(0.00963)	
CPI	−0.041563***		6.955875***	
	(0.00241)		(0.4395531)	
Lambda		−0.1805037		−0.1558748
		(0.0128445)		(0.0133838)
Constant	6.955875***	7.480404***	6.955875***	7.467818***
	(0.4395531)	(0.0181366)	(0.4395531)	(0.0182261)

Table 10 (continued)

Regression	Model without computer		Model with computer	
	Selection equation	Main equation	Selection equation	Main equation
LR Chi-squared	1730.34		1730.34	
Wald Chi-squared		1794.51		2030.21
Probability > Chi-squared	0.0000	0.0000	0.0000	0.0000

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

Standard errors in parentheses; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$; lambda refers to the inverse Mills ratio

$$0.0203 - 0.0006E = 0$$
$$-0.0006E = -0.0203$$
$$E = \frac{0.0203}{0.0006}$$
$$E = 34.7886$$

Thus, wages are maximised at 34 years of potential experience. The second derivative is negative, further confirming the inverted U-shaped nature of the relationship.

All the results from the Heckman two-step estimation are statistically significant at the one percent level. Furthermore, the probability of the Chi-square test statistic also indicates that overall, the variables used in the models are also jointly significant. It is observed once again that including the computer use dummy reduces the value of the education coefficient. This further consolidates the notion that estimating returns to education without considering the returns to computer use tends to produce biased results. This is because a part of the returns to education are actually due to computer skills and not due to education.

A comparison of the estimated returns to education, experience, and computer use reveals that computer skills are equivalent in value to roughly 5 years of education or 8 years of experience, if the returns to both education and experience are assumed to remain constant over a lifetime (Table 11).

7 Conclusion

The gaps in the labour market of Bangladesh need to be bridged urgently, if the country is keen on enjoying the greatest benefit from its demographic dividend. The findings of this research show that workers who have used computers earn 17% more than workers who did not use computers. At the same time, it was found that each additional year of schooling and potential labour market experience increased wages by 3% and 2%, respectively. However, the returns to experience peaked at 34 years, and subsequently declined. Ability to use a computer was found to be equally rewarding as 5 years of education or 8 years of potential labour market experience, provided that the returns to education and experience remained constant over a worker's lifetime. These results are comparable to the findings of previous studies in the literature conducted in several other countries. This study provides empirical evidence that computer skills are complementary to education and can play an instrumental role in bridging the gaps in the labour market in Bangladesh. The results of this paper are an eye-opener for policy makers to allocate resources for computer training, and a wake-up call for workers to invest their time in learning computer skills.

Appendix

See Fig. 5.

Table 11 Returns to education, experience, and computer use

Years	Returns to education	Cumulative education differential	Returns to experience	Cumulative experience differential	Returns to computer use
0	0.030	0.00	0.020	0.00	0.170
1	0.030	3.03	0.020	2.03	0.170
2	0.030	6.14	0.020	4.11	0.170
3	0.030	9.35	0.020	6.23	0.170
4	0.030	12.66	0.020	8.39	0.170
5	0.030	16.07	0.020	10.59	0.170
6	0.030	19.58	0.020	12.84	0.170
7	0.030	23.20	0.020	15.14	0.170
8	0.030	26.92	0.020	17.48	0.170
9	0.030	30.76	0.020	19.87	0.170
10	0.030	34.72	0.020	22.31	0.170
11	0.030	38.79	0.020	24.80	0.170
12	0.030	42.99	0.020	27.34	0.170
13	0.030	47.32	0.020	29.93	0.170
14	0.030	51.78	0.020	32.57	0.170
15	0.030	56.37	0.020	35.27	0.170
16	0.030	61.10	0.020	38.02	0.170
17	0.030	65.97	0.020	40.83	0.170
18	0.030	70.99	0.020	43.70	0.170

Source: Authors' calculations based on Labour Force 2013 data (BBS, 2015c)

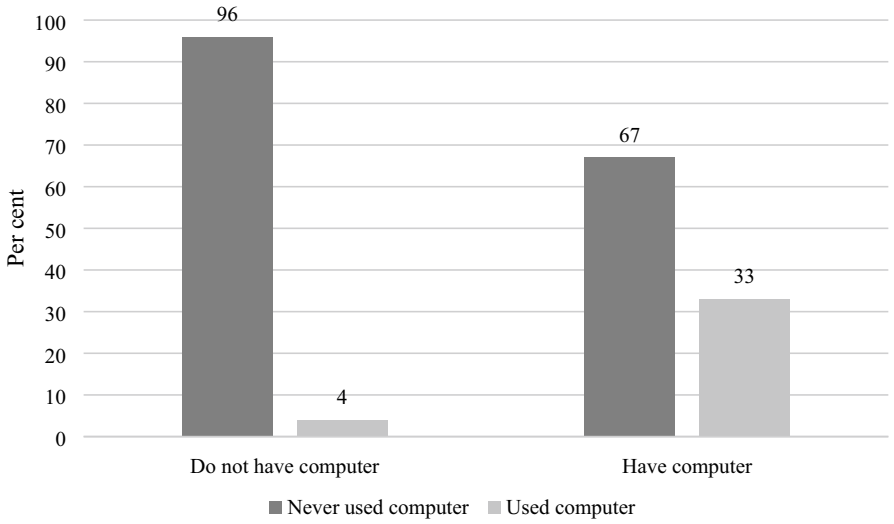

Fig. 5 Computer ownership and computer use (in percentage). *Note*: Computer ownership is shown on household level and computer use is shown on individual level. Source: Authors' calculations based on Labour Force 2013 and ICT Use and Access Survey data (BBS, 2015a)

References

Autor, D., L. F. Katz, A. B. Krueger. 1996. *Computing Inequality: Have Computers Changed the Labor Market?*

Bangladesh Bank. 2013. *Annual Report.* Dhaka: Bangladesh Bank.

BBS. 2015a. *Report on Labour Force Survey (LFS) Bangladesh 2013.* Dhaka: Bangladesh Bureau of Statistics (BBS).

BBS. 2015b. *ICT Use and Access by Individuals and Households Bangladesh 2013.* Dhaka: Bangladesh Bureau of Statistics (BBS).

BBS. 2015c. *Report on Labour Force Survey (LFS) Bangladesh 2013.* Dhaka: Bangladesh Bureau of Statistics (BBS).

BBS. 2018. *Quarterly Labor Force Survey 2016–2017.* Dhaka: Bangladesh Bureau of Statistics (BBS).

Begum, S., S. Y. Saadat. 2018. The Value of a Bangladeshi Woman's Time: An Econometric Analysis. *Journal of Accounting, Finance and Economics, 8*(2), 20–39. Retrieved from https://zantworldpress.com/wp-content/uploads/2018/09/2.-Salma-and-Yusuf.pdf

Beveridge, W.H. 1944. *Full Employment in a Free Society*. London: Allen and Unwin.

Blanchard, O. J., P. Diamond, R. E. Hall, J. Yellen. 1989. The Beveridge curve. *Brookings papers on economic activity*, 1–76.

Bleakley, H., J. C. Fuhrer. 1997. Shifts in the Beveridge Curve, Job Matching, and Labor Market Dynamics. *New England Economic Review*.

Bowden, R.J. 1980. On the existence and secular stability of uv loci. *Economica* 47 (18): 35–50.

CPD. 2018. *State of Bangladesh Economy and Upcoming National Elections.* Dhaka: Centre for Policy Dialogue (CPD). Retrieved from https://cpd.org.bd/wp-content/uploads/2018/12/IRBD-Paper-State-of-Bangladesh-Economy-and-Upcoming-National-Elections.pdf

DiNardo, J.E., and J.-S. Pischke. 1997. The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too? *The Quarterly Journal of Economics* 112 (1): 291–303.

Dolton, P., and G. Makepeace. 2004. Computer use and earnings in Britain. *The Economic Journal* 114 (494): C117–C129.

Dostie, B., R. Jayaraman, M. Trépanier. 2006. *The Returns to Computer Use Revisited, Again. IZA Discussion Papers, No. 2080*. Bonn: Institute for the Study of Labour (IZA).

Dow, J.C., and L.A. Dicks-Mireaux. 1958. The excess demand for labour A study of conditions in Great Britain. *Oxford Economic Papers* 10 (1): 1–33.

Hamilton, B.H. 1997. *Returns to computer skills and black-white wage differentials*. Olin School of Business, Mimeo: John M.

Hansen, B. 1970. Excess demand, unemployment, vacancies, and wages. *The Quarterly Journal of Economics* 84 (1): 1–23.

Heckman, J. 1974. Shadow Prices, Market Wages, and Labor Supply. *Econometrica* 42 (4): 679–693.

Heckman, J. 1976. The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and A Simple Estimator for such Models. *Annals of Economic and Social Measurement* 5 (4): 475–491.

Heckman, J. 1979. Sample Selection as a Specification Error. *Econometrica* 47 (1): 153–161.

Holt, C., and David, M. (1966). The concept of job vacancies in a dynamic theory of the labor market. In *The measurement and interpretation of job vacancies* (pp. 73–110). National Bureau of Economic Research (NBER).

Krueger, A.B. 1993. How computers have changed the wage structure: evidence from microdata 1984–1989. *The Quarterly Journal of Economics* 108: 33–60.

Mincer, J.A. 1974. *Schooling, Experience, and Earnings: Human Behavior & Social Institutions No.2*. New York: National Bureau of Economic Research.

Petrongolo, B., and C.A. Pissarides. 2001. Looking into the black box: A survey of the matching function. *Journal of Economic Literature* 39: 390–431.

Pissarides, C. 1986. Unemployment and vacancies in Britain. *Economic Policy* 1: 499–559.

Pissarides, C.A. 2000. *Equilibrium Unemployment Theory*. Cambridge: MIT Press.

Smith, J.P. 2014. *Female labor supply: theory and estimation*. New Jersey: Princeton University Press.

World Bank. 2017. *World Development Indicators*. Retrieved from Databank: http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators

Yashiv, E. 2006. *Search and matching in macroeconomics*.