UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Econometric analysis of educational mismatch and earnings using survey data from Ghana

Darko, Christian; Abrokwa, Kennedy

DOI:
10.4337/9781788118774.00024

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Darko, C & Abrokwa, K 2019, Econometric analysis of educational mismatch and earnings using survey data from Ghana. in *Handbook of Research Methods on The Quality of Working Lives.* Handbooks of Research Methods in Management, Edward Elgar Publishing Ltd., pp. 235–252. https://doi.org/10.4337/9781788118774.00024

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
This is a draft chapter. The final version is available in Handbook of Research Methods on the Quality of Working Lives edited by Daniel Wheatley, published in 2019, Edward Elgar Publishing Ltd https://doi.org/10.4337/9781788118774

The material cannot be used for any other purpose without further permission of the publisher, and is for private use only.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 28. Apr. 2024

CHAPTER 17

Econometric Analysis of Educational Mismatch and Earnings using Survey Data from Ghana

Christian K. Darko, University of Birmingham, UK, and Kennedy K. Abrokwa, Ghana Institute of Management and Public Administration, Ghana

Christian K. Darko[a]

Christian Darko is a lecturer in Applied Business and Labour economics in the Department of Management at the University of Birmingham. His research interests include human capital development, intergenerational transmission of inequality, labour market outcomes, firm performance, and gender-related issues and their effects on socioeconomic outcomes.

Kennedy K. Abrokwa

Kennedy Kwabena Abrokwa is a lecturer in the School of Liberal Arts and Social Science at the Ghana Institute of Management and Public Administration (GIMPA). His research interest includes energy and economic growth modelling, labour market outcomes and inequality, structural transformations and financial deepening.

[a]*corresponding author: Birmingham Business School, University of Birmingham, Edgbaston Park Road, B15 2TT. Email:* c.k.darko@bham.ac.uk*; Tel: +44 (0) 121 414 7202.*

SUMMARY

Realized matches and objective methods are used to measure educational mismatch. There are observed differences in the incidences of over and undereducation depending on the method of measurement. Results from OLS selectivity-corrected regressions indicate a wage premium for workers with excess schooling and a penalty for workers with deficit schooling.

ABSTRACT

This chapter evaluates empirical methodologies that relate to the measurement of educational mismatch and estimation of earning regressions. The realised matches method and objective method are used to show evidence of the incidence of educational mismatch among wage workers in the Ghana labour market and their effects on earnings. We use micro data from the recent 2012/13 Ghana Living Standards Survey. The results indicate that each method produces different incidences of educational mismatch, although the associated returns to over and undereducation are similar. Our findings from OLS selectivity-corrected regressions highlight wage premium for workers with excess years of schooling than is required for their jobs, while

a wage penalty is observed for workers with fewer years of schooling than is required for their jobs.

Keywords: Overeducation; undereducation; earnings; wage work; sample selection.

INTRODUCTION

In his book 'the overeducated American', Freeman (1976) noted that a college diploma which was once valued, was no longer a guarantee to economic success. Returns from college education had declined considerably such that additional investment in college education yielded at best, very marginal returns. Since then, several studies have examined educational mismatch in the labour market and its implications on earnings (Diem, 2015; Groot and Van den Brink, 2000a; 2000b; Kiker et al., 1997; McGuinness, 2006; Peiró et al., 2010; Piper, 2015; Verhaest and Omey, 2006).

Obtaining reliable and consistent estimates of the effects of educational mismatch on earnings can be a challenging task. This is mainly due to how educational mismatch is measured and the estimation techniques used. Educational mismatch can be defined as either overeducation (where an individual has more years of education than what is required for his/her job), or undereducation (where the individual has fewer years of education than is required for the job). Three main methods are used to measure overeducation and undereducation in the literature: self-assessment (subjective method); job analysis (objective job evaluation method); and realized matches or empirical approach (Leuven and Oosterbeek, 2003; McGuiness et al., 2017).

This chapter reflects on the use of the realised matches and objective methods approaches of measuring educational mismatch to provide evidence on their effects on earnings using data from the 2012/2013 Ghana Living Standards Survey. The chapter addresses the issue of sample selection by using Heckman’s correction procedure (Heckman, 1979). The choice of Ghana is interesting for a number of reasons: over the last few decades, there have been significant improvements in educational investments that has led to increased educational attainment**.** For instance, the proportion of individuals without any educational attainment reduced from 30.8 percent in 2005/06 (Ghana Statistical Service, 2008) to 19.7 percent in 2012/13 (Ghana Statistical Service, 2015). Despite these improvements, access to jobs are limited and the labour market has been unable to absorb the increasing supply of labour.

Our results showed that each method produced different incidences of educational mismatch, but the earnings effects of over and undereducation were similar. OLS selectivity-corrected regressions indicated that there were wage premiums for individuals with excess years of schooling, while a wage penalty was observed for workers with fewer years of schooling than was required for their jobs.

The rest of the chapter is structured as follows. The next section provides a review of the literature on how educational mismatch is measured and some of the econometric issues associated with estimating earnings regressions. Section three provides a description of the data and variables. The estimation techniques are discussed in section four and section five presents a discussion of the methods, including key strengths and limitations.

MEASURING EDUCATIONAL MISMATCH

This section examines the different approaches to measuring educational mismatch and a brief discussion of some of the econometric issues associated with estimating earnings regressions.

Self-assessment (Subjective Method)

This measure of mismatch focuses on the schooling requirements of the worker's job. The researcher questions the individual to assess their level of education and gauge it with the required level of qualification needed for the type of job. The measure can either be direct self-assessment (DSA) or indirect self-assessment (ISA). The DSA requires individuals to confirm if their job matches their level of education (Chevalier, 2003; and Verhaest and Omey, 2006) while the ISA asks the individuals the educational requirements needed for their current job (Battu et al., 2000; Green and Zhu, 2010). The subjective nature of this approach has the potential to generate biased estimates and may imply different interpretations (McGuiness et al., 2017). For instance, a respondent could overstate or understate their requirements, thus producing conflicting matching for the same job (Hartog, 2000; Leuven and Oosterbeek, 2003).

Objective Method (Job analysis)

The objective method, sometimes referred to as job analysis, requires evaluation of job requirements by professional job analysts. Such information is often provided in the US Dictionary of Occupational Titles (DOT). Although this approach is often perceived to be more accurate (McGuinness et al., 2017), it has a number of downsides. First, gathering this information is very expensive and is therefore not widely available at the national level for most surveys. There is also the tendency for data to be outdated (Hartog, 2000). Secondly, measurement may contain some level of subjectivity since it involves the opinions of experts (McGuinness et al., 2018).

Realized Matches

The realised matches method uses the distribution of workers' educational level within each occupation to identify the level of education required for a job. Two main approaches are adopted when using realised matches: the mean years of education within each occupation (Verdugo and Verdugo, 1989; Hung, 2008); and the modal level of education (Kiker et al., 1997; de Oliveira et al., 2000) among workers within each occupation. Workers with acquired education above the average level are considered overeducated and those below are considered to be undereducated. The realised matches method is appealing mainly because it can be applied to most micro datasets that contain information on occupation and educational attainment (Hung, 2008; Flisi et al., 2017). A disadvantage with this method is the lack of information on the actual skill requirement of the job. This makes it difficult to appreciate actual variations in required schooling for particular jobs (Verhaest and Omey, 2010). The lack of information on skill requirement in the data that we use is a limitation of this study. The mean approach relative to the mode approach may also be sensitive to outliers. For instance, older workers with longer tenure are likely to affect the occupational averages and the measurement may not reflect current requirements (Leuven and Oosterbeek, 2003; McGuinness et al., 2018). This makes the modal approach more appealing (Kleinbrink, 2013).

Regardless of method and specifications used to analyse earnings effects of educational mismatch, the general finding has been that overeducated workers suffer a wage penalty (Nieto and Ramos, 2017; Kleibrink, 2015; Mateos-Romero and Salinas-Jiménez, 2017). For instance, Mateos-Romero and Salinas-Jiménez (2017) found an estimated penalty in the range of 15% to 23% for a sample of Spanish workers with higher level of education.

Dealing with Econometric Issues

When estimating earnings regressions, a number of potential econometric issues need to be considered. Two main potential issues often encountered are sample selection, mainly because wages are observed only for employed individuals, an indication that the sample of employed workers may not be a random subset of the total population; and potential endogeneity of the education variables which arises when the education variable is correlated with the error term (Harmon et al., 2003; Leuven and Oosterbeek, 2011). The technique often used to address the former is the Heckman selection procedure (Heckman, 1979), while instrumental variables are used for the latter.

Numerous studies find OLS estimates to be biased downwards when sample bias is not addressed. Dealing with sample selection requires identifying variables that significantly affect participation into employment or wage jobs, while at the same time these variable(s) do not directly affect earnings. Variables used as exclusion restrictions include number of children (Barone and Ortiz, 2011; Borden, 1996; Herrera and Merceron, 2013; Cutillo and Di Pietro, 2006; and Fersterer and Winter-Ebmer, 2002); parental education (Kucel and Vilalta-Bufi, 2016); and employment status of spouse. For further details see Mulligan and Rubinstein (2008) and Huber and Mellace (2014).

In addition, education is endogenous due to certain unobserved factors that influences the individual's ability to acquire more years of schooling while these same factors do not have an independent effect on the individual's earnings (Leuven and Oosterbeek, 2011). Previous studies address this issue by using instrumental variables, such as place of residence during childhood, economic problems and disruptions in the family of origin, number of siblings, and family background (Dearden, 1999; Dolton and Silles, 2008; Harmon et al., 2003; and Korpi and Tahlin, 2009; Robst, 1994)[1]. A recent study by Sellami et al. (2017) use direct self-assessment of overeducation and indirect self-assessment as instruments to address measurement error in the objective method. In the current study, we are unable to address the issue of endogeneity. No convincing instrument was found in the data that we use.

DATA

Data was drawn from the sixth round of the Ghana Living Standards Survey conducted in 2012/13. The survey is nationally representative and was initiated in 1980 by the Policy Research Division of the World Bank. The sixth round of the survey covered 1,200 enumeration areas across the ten administrative regions and covered a representative sample of 18,000 households. 16,772 households were successfully enumerated, leading to a response rate of 93.2 percent. The survey collects detailed information on individual characteristics including education and employment, and household characteristics.

The sample used consists of individuals aged 25 to 60 years and in wage jobs. Individuals in this age group are in the prime of their working lives (OECD, 2016), and are not likely to be participating in any schooling activities, thus making them suitable for our analysis. The upper limit of 60 years reflects the official retirement age in Ghana. The sample analysed excluded self-employed individuals. Self-employment is a heterogeneous group and often difficult to measure their earnings and working hours due to the flexibility associated with this form of employment (see Fields (2014) for a discussion on self-employment in developing countries). As such, it is best to exclude from our analysis. Table 1 presents summary description of the variables.

[Insert Table 1. Here]

Figure 1 illustrates the extent of educational mismatch in years by occupation (major groups). For both the mean and mode approach, individuals in more skilled occupations such as managers and professionals tend to have fewer years of education. This could be related to the age composition of workers in these occupations. Older generations of workers generally have fewer years of education and are in slightly better-paid occupations than younger cohorts, mainly due to seniority. On the contrary, workers with fewer years of education than is required for the occupation tend to be in relatively low skilled occupations such as craft and related trades, and plant and machine operators.

[Insert Figure 1. Here]

Quinn and Rubb (2006) argued in their study that the required level of education for occupations may be dynamic due to technological changes and educational quality. They therefore allowed required education to vary by age. In results not presented, we examined this possibility for different age groups and found that undereducated young workers (24-34 years) suffered a higher wage penalty than prime aged (35-44 years) and older workers (45-60 years).

Measuring Overeducation and Undereducation

The main method we rely on is the realised matches method. The availability of information on education and occupation in the data makes it appealing for our analysis. Two definitions of required education for each occupation based on the unit groups level according to the International Standard Classification of Occupations (ISCO) are used: the first assumes the required level of education for each unit group occupation to be the modal level of schooling among individuals within each occupation (Kiker et al., 1997). The second relies on the mean years of educational attainment for each unit group occupation. The required amount of schooling is inferred from the mean of completed years of schooling of all workers in the same unit group occupation. Using these approaches, we derive required education (E^r), overeducation (E^o), and undereducation (E^u), all in years. Educational mismatch is then derived by decomposing the total years of education into E^r, E^o, E^u. The actual years of education (E^a) is therefore given by;

$$E^a = E^r + E^o - E^u \qquad (1)$$

with $E^o = E^a - E^r,\ if\ E^a > E^r$ and $E^u = E^r - E^a,\ if\ E^a < E^r$.

Our preferred measure of overeducation and undereducation is the mode approach, mainly because the mode approach reduces sensitivity to outliers and provides a more accurate measure of adequate education. We also present results that uses the mean approach.

We also measured over and undereducation based on the objective method by using information provided by the International Labour Organisation. This allows us to generate a match between educational attainment levels and occupational groups according to the International Standard Classification of Education (ISCED) and ISCO, as done elsewhere in the literature (Mateos-Romero & Salinas-Jimenez, 2017).

Incidence of Over and Undereducation

The incidences of overeducation and undereducation can be determined based on those individuals that deviate from the required level of education, E^r- an individual is overeducated if the years of education is above the years of education for the required level of that occupation, and undereducated if his/her education is below the required educational attainment for that occupation. For the objective approach, an individual is overeducated if their educational attainment based on ISCED is higher than the required education necessary for that occupation, and vice versa for undereducation.

Figure 2 shows the incidence of overeducation and undereducation. The proportion of individuals that are undereducated are similar using the mean and mode measure. The overeducated are substantially larger using the mean measure, but similar when using the mode and also the objective method. This illustration suggests that different measurements of over and undereducation lead to sizeable differences in the incidence of over and undereducation.

[Insert Figure 2. Here]

ESTIMATION TECHNIQUES

The returns to overeducation and undereducation are examined using an augmented Mincer-type earnings regression (Mincer, 1958; 1974). Mincer's model of earnings is a framework used to estimate the effects of schooling, schooling quality and experience on earnings and has been applied extensively in the literature. See Heckman (2003) for a discussion on Mincer earning regressions. All estimations account for selection into wage employment using Heckman's correction procedure (Heckman, 1979). Since wages are observed only for employed individuals, there is the possibility that our sample of employed workers is not a random subset of the total population and failure to take this into account could bias our estimates of the returns to education. We therefore proceed by firstly deriving the Inverse Mills Ratio (IMR) from the selection equation (wage or self-employment). IMR is the ratio of the probability density function over the cumulative distribution function of a distribution (Heckman, 1979). The selection equation is estimated using a probit model and is of the form:

$$PR(Wemp_i = 1) = x_i\beta + g_i\delta + \varepsilon_i, \qquad (2)$$

where $Wemp_i$ reflects wage employment of a worker i. x_i is a vector of individual characteristics including years of education, age and gender, and g_i is a vector of variables not included in the earnings equation and therefore satisfying the exclusion restrictions; number of young children (less than six years old) in the household and the number of people in the household that are in paid work. Our rationale for including these variables are that, young children may impose a time constraint and can therefore affect the individual's participation in wage work (Mulligan and Rubinstein, 2008; and Huber and Mellace, 2014) as opposed to self-employment which is more flexible in terms of working hours. Similarly, individuals in households with more people in work will be less likely to find wage jobs (Pagan, 2002). These can have direct effects on the decision to find wage jobs but will not directly affect earnings.

Different specifications have been proposed when estimating the effect of educational mismatch on wages - the ORU specification by Duncan and Hoffman (1981, henceforth DH), and Verdugo and Verdugo (1989, henceforth VV) specification. The first earnings equation that we estimate is based on the approach proposed by DH and is of the form:

$$lny_i = \alpha + \beta_1 E_i^r + \beta_2 E_i^o + \beta_3 E_i^u + \beta_4 X_i + \beta_5 IMR + \varepsilon_i \tag{3}$$

where lny_i denotes the log of weekly earnings in Ghana Cedis for individual i. E_i^r, E_i^o, and E_i^u denotes required education, overeducation, and undereducation respectively. These are all measured in years. X_i is a vector of individual characteristics including gender, tenure, and firm level characteristics such as occupation and sector dummies. Region and locational dummies are also included to capture regional and development policies. IMR is the selection term derived from equation 2.

The second earnings estimation is based on VV. This specification uses dummy variables related to overeducation and undereducation and is of the form:

$$lny_i = \sigma + \theta_1 E_i^a + \theta_2 OE_i + \theta_3 UE_i + \theta_4 X_i + \theta_5 IMR + \varepsilon_i \tag{4}$$

where OE and UE are dummies for overeducation and undereducation respectively. These take a value of 1 if the individual is overeducated or undereducated for their current job and 0 if otherwise. E_i^a is the attained years of educational attainment measured as the highest grade completed. All other variables are as previously defined.

While we acknowledge the need to take into account endogeneity, a key challenge we face with our data is that, there are no retrospective information or other convincing instruments that could be used as instruments. We therefore proceed with OLS estimations which we believe are reasonable estimates of the true returns to over and undereducation once selectivity bias and other covariates are considered.

DISCUSSION

Table 2 presents results from the estimations specified in equation 3. For brevity, we comment only on the coefficients for the main variables of concern – education. The results of all the other covariates are in line with findings from the literature (Nieto & Ramos, 2017; Rubb, 2014).

[Insert Table 2. Here]

Across both specifications, the selection term is positive and significant, indicating that an individual with the sample average characteristics who selects into wage employment gets higher earnings than an individual drawn randomly from the population with comparable characteristics.

The coefficients of required education, overeducation and undereducation are statistically significant and have the expected signs. The highest returns to an additional year of schooling are received by those with the required level of education for their jobs, while the lowest returns are associated with those that have deficit schooling. Specifically, each extra year of schooling beyond the required schooling for that occupation generates additional earnings of 6.1 and 7.6% for the mode and mean method respectively. These are similar to what Herrera & Merceron (2013) find for a sample of West African countries. A penalty of almost 10% is associated with undereducation. Despite receiving a wage premium, workers with surplus schooling still earn lower than those with the required years of education, but more than those with fewer years of schooling. These results are also consistent with the literature (Mehta et al, 2011; Nieto & Ramos, 2017; Rubb, 2014). Particularly for those with deficit schooling, the preceding results suggest why they receive a penalty – undereducated workers are more likely to work in relatively skilled jobs. Their lack of adequate educational attainment thus reduces their bargaining power in these jobs (Katz et al., 2015), thus forcing their earnings downwards.

Overall, these findings show that despite the observed differences in the incidence of over and undereducation, the differences in terms of their impact on the earnings functions are minimal once selection bias and other variables are controlled for.

Table 3 reports results based on VV. In all three specifications, average returns to educational attainment (actual years of schooling) are about 8%. Overeducated workers relative to exactly-educated workers suffer a penalty (estimated in the range between 7 and 16%), although this is only significant for the mode measure. The negative coefficient for undereducated relative to exactly-educated individuals is in line with Groot & Brink (2000) for developed countries.

The objective method shows a larger penalty for undereducated workers compared to the realised matches method.

[Insert Table 3. Here]

Table A1 and A2 in the appendix show results from estimations that do not account for selectivity bias. The results suggest that ignoring selection into wage employment is likely to underestimate the full returns to over and undereducation among workers. This is consistent with parts of the literature (Cutillo and Di Pietro, 2006; Dolton and Silles, 2008).

The results from this paper show largely that, workers with surplus schooling receive a premium, although lower than those with the required level of education. Workers with deficit schooling on the other hand receive a penalty for working in occupations that are higher than their educational attainment. In terms of the incidences of over and undereducation, there are variations in the share of under and overeducated workers, although, the estimated impact on earnings are similar for both estimations that use the mode and mean measures. The differences in incidences of mismatch between the mean and mode approaches reflect the sensitivity of the latter to outliers. Reliance on the mean method may for instance, be influenced by older workers with longer tenure which are more likely to affect the occupational averages. As such, measurement of mismatch using the mean method may not reflect current requirements (Leuven and Oosterbeek, 2003; McGuinness et al., 2018). A further advantage of the realised matches approach is that the educational standard of the status-specific occupational benchmark is updated regularly to capture changing educational requirements overtime. This is not necessarily the case for other measures such as self-assessment which suffers from subjectivity bias (Leuven and Oosterbeek, 2011; Hartog, 2000; Borghans & de Grip, 2000).

Unlike the realised matches method which can be computed easily from most surveys, a major drawback of the objective method is that it is costly to collect information on occupation and educational attainment on a regular basis. Although a direct measure for the objective method would have been a preferred option, data necessary for this computation is not available in the survey data used in this paper. The indirect measure we employ however show similar effects on earnings as did the realised matches method. This again seems to suggest that differences in measurement methods mainly relate to incidence of over and undereducation rather than impact on earnings. Results from our analyses should be treated with caution as these are only suggestive of correlations rather than causation. To further understand the causal mechanisms of educational mismatch and earnings, it is important for developing countries to collect detailed information on for instance, retrospective data on family background and economic shocks, skill requirements for jobs and other cognitive and non-cognitive measures. Such information can allow researchers to use other estimation techniques such as instrumental variable techniques.

REFERENCES

Barone, C. and Luis Ortiz, L., (2011). Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation, *High Education*, 61:325–337.

Battu, H., Belfield, C. R., & Sloane, P. J. (2000). How well can we measure graduate over-education and its effects? *National Institute Economic Review*, 171(1), 82–93.

Borden, R. J., (1996). Gender and Self-Employment Selection: An Empirical Assessment**,** *Journal of Socio-Economics*, 25 (6), 671-682**.**

Borghans, L. & de Grip A. (2000). The debate in economics about skill utilization, in: Borghans, L., A. de Grip (eds.), The overeducated worker? The economics of skill utilization. Edward Elgar, Cheltenham/UK, 3-23.

Chevalier, A. (2003). Measuring over-education. *Economica*, 70(279), 509–531.

Chevalier, A. and Lindley, J. (2009). Overeducation and the skills of UK graduates. Journal of the Royal Statistical Society 172: 307–337.

Cutillo, A. and Di Pietro, G. (2006). The effects of overeducation on wages in Italy: A bivariate selectivity approach. International Journal of Manpower 27: 143–168.

de Oliveira, M. M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review,* 19(2), 199-206

Dearden, L. (1999). The effects of families and ability on men's education and earnings in Britain. *Labour Economics*, 6, 551–567.

Diem, A. (2015). Overeducation among graduates from universities of applied sciences: Determinants and consequences. *Journal of Economic & Financial Studies,* 3: 63–77.

Dolton, P. and Silles, M. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27(2): 125–139.

Duncan, J. and Hoffman, S. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1): 75–86.

Fersterer, J. & Winter-Ebmer, R. (2002). Are Austrian returns to education falling over time? *Labour Economics*, 10, 73–89.

Fields, Gary. (2014). Self-employment and poverty in developing countries. *IZA World of Labor 2014*: 60. doi: 10.15185/izawol.60.

Flisi, S., Goglio, V., Meroni, E., Rodrigues M., and Vera-Toscano, E. (2017). Measuring Occupational Mismatch: Overeducation and Overskill in Europe—Evidence from PIAAC. *Soc Indic Res*, 131:1211–1249.

Freeman, R. (1976). *The Overeducated American*. New York: Academic Press.

Ghana Statistical Service (2015). 'Ghana Living Standards Survey round 6.' Main report. Available at: http://www.statsghana.gov.gh/docfiles/glss6/GLSS6_Main%20Report.pdf

Ghana Statistical Service. (2008). "Ghana Living Standards Survey, Report of the fifth round." Available at: http://www.statsghana.gov.gh/docfiles/glss5_report.pdf

Green, F., & Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62(4), 740–763.

Groot, W. and van den Brink, H. (2000). Skill mismatches in the Dutch labour market. *International Journal of Manpower*, 21(8): 584–595.

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The returns to education: Microeconomics. *Journal of economic surveys*, 17(2), 115-156.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go?. *Economics of education review*, 19(2), 131-147.

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

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2003). Fifty years of Mincer earnings regressions (No. w9732). *National Bureau of Economic Research*.

Herrera, J. & Merceron, S. "Underemployment and Job Mismatch in Sub-Saharan Africa. In: De Vreyer, Philippe, and François Roubaud, Editors. 2013. *Urban Labor Markets in Sub-Saharan Africa*. Africa Development Forum series. Washington, DC: World Bank. doi:10.1596/978-0-8213-9781-7.

Huber, M., & Mellace, G. (2014). Testing exclusion restrictions and additive separability in sample selection models. *Empirical Economics*, 47(1), 75-92.

Hung, C. (2008). Overeducation and undereducation in Taiwan. *Journal of Asian Economics* 19: 125–137.

Katz, Harry C., & Kochan, T. A., & Colvin, A. J. S. (2015). The role of the economic, technological, and demographic environments [Electronic version]. In *Labor relations in a globalizing world* (pp. 79-101). Ithaca, NY: ILR Press, an imprint of Cornell University Press. Accessed [07 April 2018], from Cornell University, ILR: http://digitalcommons.ilr.cornell.edu/articles/1041

Kiker, B. F., Santos, M. C., & De Oliveira, M. M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2), 111-125.

Korpi T. and Tåhlin M. (2009). 'Educational Mismatch, Wages, and Wage Growth: Overeducation in Sweden, 1974–2000', *Labour Economics*, 16(2): 183–193.

Kucel, A. & Vilalta-Bufi, M. (2016). "Entrepreneurial skills and wage employment", *International Journal of Manpower*, 37 (3), 556-588, https://doi.org/10.1108/

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, 4, 283-326.

Mateos-Romero, L., & del Mar Salinas-Jiménez, M. (2017). Skills heterogeneity among graduate workers: real and apparent overeducation in the Spanish labor market. *Social Indicators Research*, 132(3), 1247-1264.

McGuinness, S., Pouliakas, K. and Redmond, P. (2018). Skills Mismatch: Concepts, Measurement and Policy Approaches. *Journal of Economic Surveys*, 00 (0), 1–31, doi:10.1111/joes.12254

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20: 387–418.

Mehta, A., Felipe, J., Quising, P., & Camingue, S. (2011). Overeducation in developing economies: How can we test for it, and what does it mean?. *Economics of Education Review*, 30(6), 1334-1347.

Mincer, J. (1958). Investment in human capital and personal income distribution. *The journal of political economy*, 66 (4): 281-302

Mincer, J. (1974). Schooling, Experience, and Earnings. *Human Behavior & Social Institutions* No. 2.

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

Nieto, S., & Ramos, R. (2017). Overeducation, skills and wage penalty: Evidence for Spain using PIAAC data. *Social Indicators Research*, 134(1), 219-236.

OECD (2016). Employment rate by age group (indicator). [online] doi: 10.1787/084f32c7-en [Accessed on 03 April 2018].

Pagán, J. A. (2002). Gender differences in labor market decisions in rural Guatemala. *Review of Development Economics*, *6*(3), 428-441.

Peiró, J. M., Agut, S., & Grau, R. (2010). The relationship between overeducation and job satisfaction among young Spanish workers: The role of salary, contract of employment, and work experience. *Journal of Applied Social Psychology*, *40*(3), 666-689.

Piper, A. (2015). Heaven knows I'm miserable now: Overeducation and reduced life satisfaction. *Education Economics*, 23: 677–692.

Robst, J. (1994). Measurement Error and the Returns to Excess Schooling. *Appl. Econ. Lett. 1*, 142–144.

Rubb, S. (2014). Overeducation and earnings within an occupation: controlling for occupational heterogeneity of nurses. *Education Economics*, *22*(3), 257-269.

Sellami, S., Verhaest, D., Nonneman, W. and Van Trier, W. (2017). The impact of educational mismatches on wages: The influence of measurement error and unobserved heterogeneity. *The BE Journal of Economic Analysis and Policy*. 17(1).

Silles, M. and Dolton, P. (2002). The Effects of Over-education on Earnings in the Graduate Labour Market, Department of Economics Discussion Paper Series 126, University of Oxford, Oxford.

Verhaest, D. and Omey, E. (2012). Overeducation, undereducation and earnings: Further evidence on the importance of ability and measurement error bias. *Journal of Labor Research*, 33: 76–90.

Verhaest, D. and Omey, E. (2010). The determinants of overeducation: Different measures, different outcomes? *International Journal of Manpower*, 31(6): 608–625.

Verdugo, R. R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, 629-643.

ANNOTATED FURTHER READING

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press**.**

The fourth chapter of this book discusses instrumental variable techniques in a very 'friendly' way. Examples from published peer-reviewed articles are used and careful interpretation of results provided. The book offers extremely useful introduction and guides the reader through the many issues that researchers are confronted with in empirical work.

Handel, M. J., Valerio, A., & Sánchez Puerta, M. L. (2016). *Accounting for Mismatch in Low- and Middle-Income Countries: Measurement, Magnitudes, and Explanations*. The World Bank.

This book offers a useful understanding of the concept of mismatch both in terms of education and skills in twelve low and middle-income countries. Although intended primarily at policy-making audiences, it provides a clear discussion on defining and measuring mismatch and detailed analysis across different sub-groups.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, 4, 283-326.

This chapter presents a comprehensive discussion of mismatch in the labour market. The different facets of mismatch are carefully and thoroughly discussed. Readers with limited knowledge of related labour market theories will find the sixth section particularly important. Issues of omitted variables and measurement error often faced by researchers when estimating earnings regressions are also presented.

TABLES

Table 1. Summary statistics

	Mean	SD
Log weekly earnings (in Ghana Cedis)	4.57	1.02
Attained education (in years)	12.71	5.03
Experience in job (in years)	8.91	8.45
Experience squared	150.72	295.95
Hours worked per week	48.04	19.20
Female	0.29	0.45
Rural location	0.25	0.43
Observations	3,115	

Table 2. Overeducation, undereducation and earnings - DH specification

	(1) Mode	(2) Mean
Required education	0.0860***	0.114***
	(0.0121)	(0.0145)
Overeducation	0.0612***	0.0762***
	(0.0116)	(0.0108)
Undereducation	-0.104***	-0.0948***
	(0.0101)	(0.0133)
Experience	0.0329***	0.0337***
	(0.0045)	(0.004)
Experience squared	-0.0005***	-0.0005***
	(0.0001)	(0.0001)
Hours worked per week	0.0027***	0.0028***
	(0.001)	(0.001)
Female	-0.458***	-0.443***
	(0.0789)	(0.0786)
IMR	0.282**	0.263**
	(0.134)	(0.134)
Constant	3.838***	3.370***
	(0.266)	(0.286)
Region	Yes	Yes
Location	Yes	Yes
Occupation	Yes	Yes
Employment sector	Yes	Yes
Observations	3,115	3,115
R-squared	0.296	0.294

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

Table 3. Overeducation, undereducation and earnings - VV

	(1)	(2)	(3)
	Realised matches		
	Mode	Mean	Objective method
Education attained	0.0886***	0.0829***	0.0784***
	(0.0100)	(0.0107)	(0.0101)
Overeducated	-0.161***	-0.0721	-0.0736
	(0.0418)	(0.0500)	(0.0511)
Undereducated	-0.0822*	-0.118**	-0.154***
	(0.0492)	(0.0587)	(0.0562)
Experience	0.0344***	0.0332***	0.0326***
	(0.0045)	(0.0044)	(0.0044)
Experience squared	-0.0005***	-0.0005***	-0.0005***
	(0.0001)	(0.0001)	(0.0001)
Hours worked per week	0.0028***	0.0029***	0.00282***
	(0.0009)	(0.0009)	(0.0009)
Female	-0.434***	-0.444***	-0.426***
	(0.0791)	(0.0783)	(0.0791)
IMR	0.246*	0.259*	0.229*
	(0.135)	(0.133)	(0.135)
Constant	3.776***	3.943***	3.994***
	(0.231)	(0.235)	(0.230)
Region	Yes	Yes	Yes
Location	Yes	Yes	Yes
Occupation	Yes	Yes	Yes
Employment sector	Yes	Yes	Yes
Observations	3,115	3,115	3,115
R-squared	0.295	0.293	0.295

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

FIGURES

Figure 1. Over and undereducation by occupation

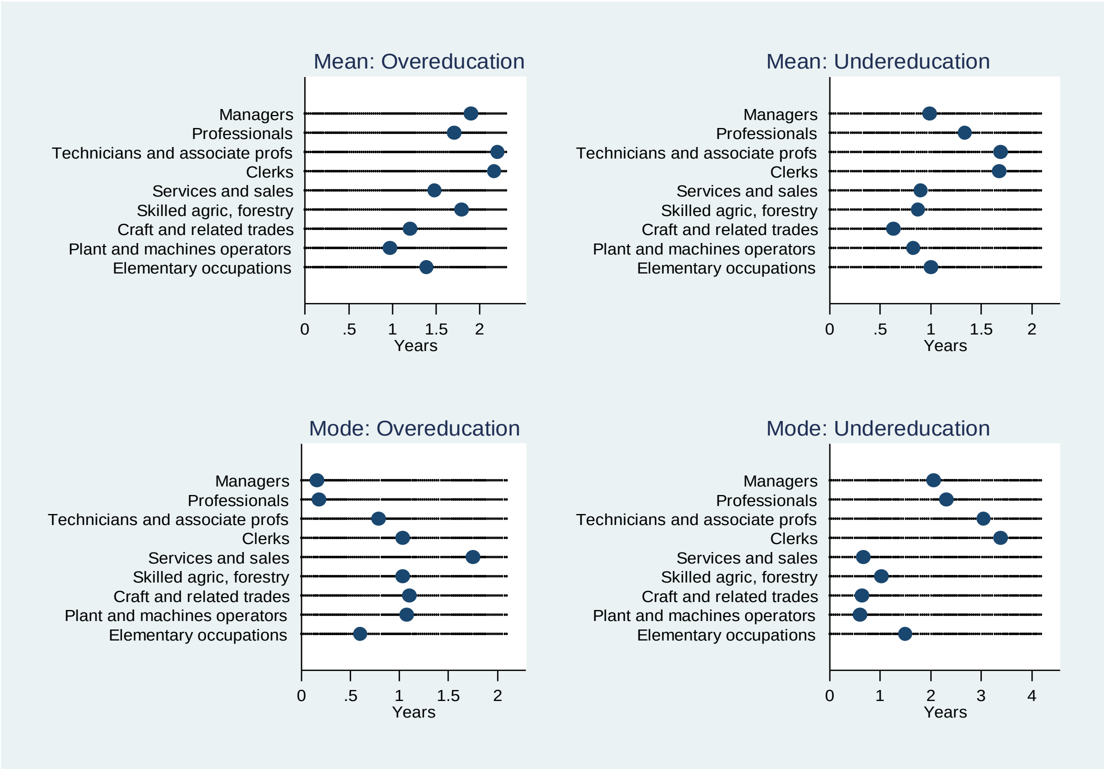
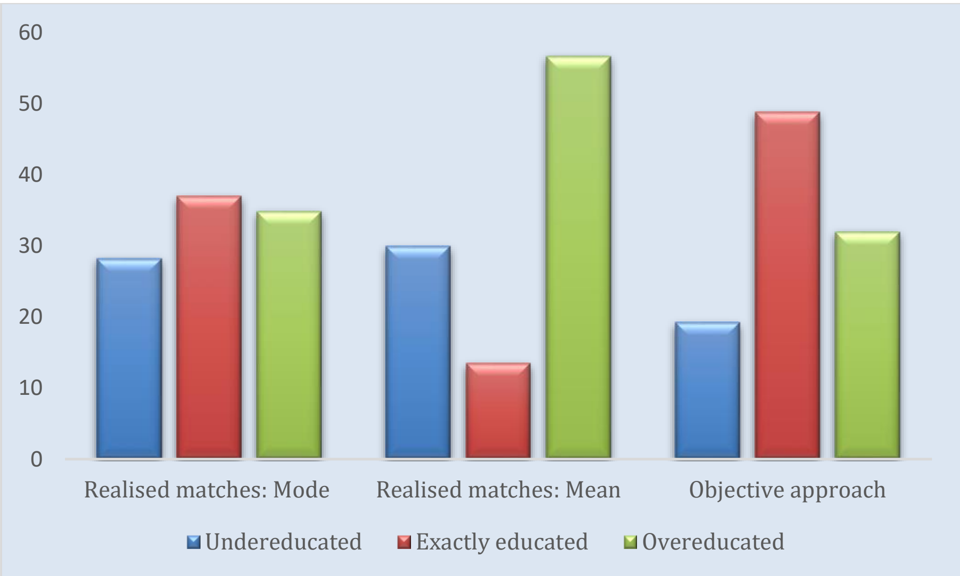


Figure 2. Incidence of educational mismatch



APPENDIX

Table A1. Overeducation, undereducation and earnings - DH specification with no selectivity correction

	(1) Mode	(2) Mean
Required education	0.0698***	0.0987***
	(0.0094)	(0.0122)
Overeducation	0.0443***	0.0622***
	(0.0084)	(0.0081)
Undereducation	-0.0871***	-0.0764***
	(0.0058)	(0.0094)
Experience	0.0347***	0.0354***
	(0.0044)	(0.0044)
Experience squared	-0.00052***	-0.00055***
	(0.0001)	(0.0001)
Hours worked per week	0.0026***	0.00268***
	(0.0009)	(0.0009)
Female	-0.312***	-0.306***
	(0.0372)	(0.0367)
Constant	4.180***	3.689***
	(0.211)	(0.235)
Region	Yes	Yes
Location	Yes	Yes
Occupation	Yes	Yes
Employment sector	Yes	Yes
Observations	3,115	3,115
R-squared	0.29	0.29

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

Table A2. Overeducation, undereducation and earnings – VV with no selectivity correction

	(1)	(2)	(3)
	Realised matches		
	Mode	Mean	Objective method
Education attained	0.0747***	0.0681***	0.0662***
	(0.00647)	(0.0075)	(0.0071)
Overeducated	-0.167***	-0.0716	-0.0904*
	(0.0417)	(0.0501)	(0.0501)
Undereducated	-0.0737	-0.109*	-0.144***
	(0.0490)	(0.0585)	(0.0559)
Experience	0.0360***	0.0349***	0.0340***
	(0.0044)	(0.0044)	(0.0044)
Experience squared	-0.00055***	-0.0005***	-0.0005***

	(0.0001)	(0.0001)	(0.0001)
Hours worked per week	0.0027***	0.0028***	0.0027***
	(0.0009)	(0.0009)	(0.0009)
Female	-0.307***	-0.309***	-0.307***
	(0.0371)	(0.0368)	(0.0367)
Constant	4.071***	4.259***	4.255***
	(0.164)	(0.170)	(0.171)
Region	Yes	Yes	Yes
Location	Yes	Yes	Yes
Occupation	Yes	Yes	Yes
Employment sector	Yes	Yes	Yes
Observations	3,115	3,115	3,115
R-squared	0.29	0.29	0.29

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