Jurnal Ekonomi Malaysia 51(1) 2017 41 - 56
http://dx.doi.org/10.17576/JEM-2017-5001-4

The Incidence and the Effect of Overskilling on Individuals' Wages in Malaysia: A Quantile Regression Approach

(Insiden dan Kesan Terlebih Kemahiran ke atas Upah Individu di Malaysia: Satu Pendekatan Regresi Kuantil)

Zainizam Zakariya
Norasibah Abdul Jalil
Khoo Yin Yin
Universiti Pendidikan Sultan Idris

ABSTRACT

This paper examines the incidence and the effect of overskilling on wages by taking individuals' unobserved heterogeneity in ability using quantile regression (QR) method. Using data from the second Malaysia Productivity and Investment Climate Survey (PICS-2), the incidence of overskilling was reported around 31 percent - for which moderately overskilled accounted for 23 percent and severely overskilled accounted for 8 percent. Preliminary analysis revealed that overskilling was found to be heavily concentrated within low-ability segments of the workers' conditional wage distributions. Using quantile regression (QR) method, the results revealed that although being overskilled resulted in wage penalty, the penalty, however, was heterogeneous across the entire workers' conditional wages distribution. Indeed, the penalty for moderately overskilled was greater at the lower deciles and became smaller or even disappears as one moved up the wages distribution. This may be consistent with the view that the overskilled workers are likely amongst the low-ability workers. By contrast, the penalty for severely overskilled, in particular women was evident all the way through the conditional wage distribution. This perhaps suggests that unobserved heterogeneity unable to explain the wages penalty for mismatched women. Nevertheless, this study may suggest the importance of including explicit controls for individuals' unobserved ability where possible, as a mean to avoid bias estimation of the wage impacts of the overskilling.

Keywords: Overskilling; quantile regression; unobserved ability; wages

ABSTRAK

Artikel ini mengkaji insiden dan kesan terlebih kemahiran ke atas upah individu dengan mengambil kira keheteregonan yang tidak dicerap di kalangan pekerja menggunakan kaedah regresi kuantil (QR). Berdasarkan data dari Tinjauan Iklim Pelaburan dan Produktiviti (PICS), insiden terlebih pendidikan dilaporkan sekitar 31 percent - terlebih kemahiran yang sederhana menyumbang 23 percent dan terlebih kemahiran yang serius mewakili 8 peratus. Analisis awal mendapati bahawa insiden terlebih kemahiran lebih banyak tertumpu pada segmen bawah taburan gaji bersyarat pekerja. Menggunakan kaedah quantile regresi (QR), dapatan kajian mendedahkan bahawa walaupun pekerja terlebih kemahiran mengakibatkan upah penalti, namun penalti tersebut adalah heterogen di sepanjang taburan upah tersebut. Malah, upah penalti bagi pekerja terlebih kemahiran yang sederhana adalah lebih besar di desil lebih rendah dan menjadi lebih kecil atau hilang apabila bergerak ke desil yang tinggi di atas taburan upah bersyarat pekerja. Keadaan ini mungkin konsisten dengan pandangan bahawa pekerja terlebih kemahiran terdiri daripada mereka yang mempunyai keupayaan yang rendah. Sebaliknya, upah penalti bagi pekerja terlebih kemahiran yang serius didapati wujud di sepanjang taburan upah bersyarat terutamanya wanita. Ini mungkin menunjukkan bahawa keheteregonan yang tidak dicerap di kalangan individu tidak dapat menjelaskan penalti upah bagi wanita. Kajian ini mungkin mencadangkan bahawa adalah penting mengawal kepelbagaian keupayaan individu yang tersembunyi seboleh mungkin bagi mengelakkan ketidaksamaan anggaran kesan insiden terlebih kemahiran ke atas upah.

Kata kunci: Terlebih kemahiran; regresi quantil; keheteregonan yang tidak dicerap; upah

INTRODUCTION

Over the last ten years, the study of over-education in Malaysia has been increased in particular (Zulkifly, Yussof, & Abu Hassan 2010; Lim 2011; 2013; Zakariya 2014a, 2014b; Zakariya & Battu 2013; Zakariya & Mohd. Noor 2014). These studies found that overeducated workers accounted for over one-third of the sample surveyed, and the incidence resulted in lower job satisfaction (Lim 2013; Zakariya & Battu 2013) and greater earnings loss (Zakariya 2014a, 2014b; Zulkifly et al. 2010). However, it is broadly accepted in the literature that over-education is a far from perfect indicator of mismatch in the labour market, especially related to unobserved heterogeneity/ability.[1] Instead, overskilling is much relevant on the context employment since it is capturing the skills available and skills required for which could not be captured by formal educational achievements due to it makes explicit reference to the respondents' abilities to perform in a job (McGuinness & Sloane 2011; Mavromaras, Mcguinness, O'Leary, Sloane & Fok 2010;).[2]

Recent research, however, revealed that overskilling measurement might be insufficient to capture the real effect of unobserved ability especially in the wage regression. Mavromaras, Mcguinness and Fok (2009) and McGuinness and Sloane (2011) revealed even after eliminating as many potential sources of bias as possible based on the principle of Propensity Score Matching (PSM), the wage penalty for overskilled workers quite similar to the conventional method, the Ordinary Least Square (OLS). For instance, the former authors discover severely overskilled workers face a pay penalty of 10 percent regardless of the OLS or PSM method. Similarly, the latter authors found that the penalty for overskilled workers is about 22 percent using the OLS and 25 percent when the PSM was concerned. Both the results strongly imply that the methods employed do not fully control the unobserved effect. This may be the fact that they are only deals with the effect of unobserved heterogeneity upon the mean wage, where the coefficients are assumed constant over the earnings. In this sense, such technique may not be really helpful to understand the wages gap between well-matched and overskilled people if unobserved heterogeneity varies across the workers' wages distribution. It may also be the case that overskilled workers reflect their position at the different point of the conditional wages distribution. Consequently, this could mislead about the effect of unobserved heterogeneity on individuals' wages (Budría & Moro-Egido 2008; Martins & Pereira 2004; McGuinness & Bennett 2007).

Therefore, in this paper, our analysis takes a different approach by employing quantile regression (QR) technique.[3] Unlike the mean regression (the OLS), the QR allows the study of the effect of overskilling on wages at different points of the distribution, consequently, describing changes in wages are not limited to the location but also in the shape of the distribution (Koenker 2004; Koenker & Bassett Jr 1978). Consequently, the QR approach explains for a non-trivial interaction between the explanatory variables and unobserved factors related to wages. Indirectly, the estimation of the effect of overskilling on conditional distributions allows the recognisation individual heterogeneity in the effect of overskilling on wages (Hartog, Pereira & Vieira 2001). This might suggest that conditional on observable characteristics, individuals located at higher quantiles of the wage distribution are to some extent those who have more ability than individuals located at the lower part of the wage distribution. Following Budría and Moro-Egido (2008), if the conditional distribution of wages emerges from the underlying distribution of unobserved skills, then differences in the overskilling wage effect between individuals at high-pay and low-pay jobs can be interpreted as differences between workers with high and low unobservable skills. Put simplify, individuals' unobserved heterogeneity in ability could be indexed by their position in different quantiles in the conditional wage distribution. By default, the technique we use here to some extent free from bias concerning unobserved heterogeneity.

Up to date, there is no study tries to examine the effect of overskilling on wages at a different quantile. Indeed, there is very few studies employ the QR in analysing the role of unobserved ability on individual's wages in the context of over-education apart from Hartog et al., (2001), McGuinness and Bennett (2007) and Budría and Moro-Egido (2008). Based upon quantile, these authors found some evidence of the wages penalty is greater amongst low-ability workers.

Up to our knowledge, there has been no overskilling study in Malaysia and in fact no such study available in developing countries. We believe that this is the first research attempts to explore the overskilling incidence and its effects in the Malaysia labour market.Therefore, by considering the conditional wage distribution of overskilled workers into account, we contribute, therefore to the overskilling literature, not only for Malaysia but also for developing country context. In doing so, the paper is structured as follows. Section 2 provides some literature review regarding unobserved heterogeneity and its relation to labour market mismatch. Section 3 discusses the data and the measurement of overskilling along the descriptive statistics while section 4 details the empirical estimation method. Section 5 explores the effect of overskilling on individuals' conditional wages distribution whilst the final section provides some concluding remarks.

MISMATCH AND INDIVIDUALS' UNOBSERVED HETEROGENEITY

Up to date, the effects of unobserved heterogeneity among mismatch workers on wages have been given less attention in the mismatch literature. Only few studies explore the relationship between overeducation or overskilling and individuals' heterogeneous.

In an earlier study, Hartog et al. (2001) utilised data from Quadros de Pessoal (Personnel Records) in Portuguese labour market for the years 1982, 1986 and 1992 to estimate the return to over-education across the workers' earnings distribution.4 Quantile regression analysis reveals that returns to years of education above the job requirement were not constant across the conditional wage distribution. Instead, the returns were increasing with quantile, i.e.- they are higher for those at higher quantiles and lower amongst those at the lower decile in the conditional wage distribution. For instance, the return to a year of education above the job requirement for men amounts to 46 percent of that to a year of required education at 10th decile and to 64 percent at 90th in 1982.

McGuinness and Bennett (2007) examined the role of unobserved individual heterogeneity on the graduates' earning distribution in the UK. By exploiting data from the second follow-up survey for a cohort study of all Northern Ireland domiciled students entering higher education in 1991/1992, they found evidence of over-education incidence were prevalent amongst low-skilled workers as proxied by their position within the graduate wage distribution. Using quantile regression techniques, the wage penalty of overeducation was heavily concentrated within low-ability segments, especially for men. For those low-ability workers (at the bottom part of the wage distribution), the earnings penalty stood at 35 percent and substantially reduces to just –11 percent at the mid-ability ranges. At the top decile, i.e.- 90th decile, there was no penalty reported, instead, overeducated men experience a wage premium about 16 percent (though, the result did not significantly different from zero). However, the authors found that the impacts of overeducation were found to be much more pervasive and constant throughout the entirety of the female ability (wage) distribution. The results provide only partial support for the hypothesis linking overeducation with lower levels of ability.

The role of unobserved heterogeneity on the wage effect of overskilling is extensively examined by Mavromaras et al. (2009) and McGuinness and Sloane (2010). Using data from the Household, Income and Labour Dynamics in Australia (HILDA), the former authors employed the PSM method along with the conventional method of OLS to predict the earnings loss among overskilled workers.[5] They found that the penalty for overskilling, particularly severely overskilled using the PSM method does not vary to the OLS technique, approximately 10 percent. To be sure that the model estimated accounted for unobserved heterogeneity; they run a subsequent sensitivity analysis test using Rosenbaum bounds.[6] The test suggests that any remaining individual unobserved heterogeneity after the application of the PSM methodology is not a problem in the estimates presented in their analysis.

McGuinness and Sloane (2011) utilised data from the Flexible Professional in the Knowledge Society (REFLEX) to examine the effect of labour market mismatch, i.e. - over-education and overskilling on earnings in the UK. With respect to overskilling, they found that the corresponding penalty is about 22 percent using the OLS and 25 percent when the PSM was concerned. Using the Rosenbaum test, they reveal that the OLS method robust to unobserved individual heterogeneity bias.

Although the PSM estimator is capable of reducing substantially biases generated by unobserved confounding factors, implicitly they cannot eliminate the impact of unobserved factors. We still can observe the overskilled workers experience the wages penalty, indeed, the penalty does reduce to a significant degree. The failure to explain the unobserved effect might be by the fact that such technique implicitly assumes that the impact of the overskilling penalty along the workers' conditional distribution is constant. What happen if the wages are not constant across the conditional quantile distribution? In other words, overskilled workers might reflect their position at the different point of the conditional wages distribution. According to Koenker (2004), coefficient of explanatory variable at certain decile of the wages distribution, normally reflect unobserved heterogeneity (p. 076). Hence, they may influence the conditional distribution of the response in many other ways. Using the quantile regression, therefore, we are willing to estimate the whole distribution of the conditional quantiles of the overskilling penalty, and to be able to study the influence of the unobserved ability along the distribution.

The above studies imply that high-ability workers able to reduce the penalty pay to the well-matched workers relative to low ability workers. Unfortunately, there is much less we know about the overskilling impact at different point of distribution since the use of quantile regression up to our knowledge not readily available. As we noted earlier, overskilling measurements explicitly reflect worker's ability to perform job. If this were the case, we would expect that the return to overskilling penalty would be greater at the lower part of the wages distribution and smaller at the end decile.

DATA DESCRIPTION AND THE MEASUREMENTS OF OVERSKILLING

The data used in this paper is taken from the second Malaysia Productivity Investment Climate Survey (PICS-2) which was carried out in 2007. This survey was a collaborative effort of the World Bank and the Malaysian Government via the Economic Planning Unit and the Department of Statistics. The focus of the survey was on the role of industrial relations in understanding the investment climate faced by enterprises and how this affects business performance. The respondents of PICS-2 were randomly selected from two main sectors, the manufacturing and business support services sectors across five (region) local labour markets. Since business services is not representative of services in general, this paper only utilises samples from the manufacturing sector due to the fact that it is representative of the manufacturing sector as a whole (World Bank 2009).[7]

TABLE 1. Descriptive statistics (mean and standard deviation) of the selected variables used

Variable	All (n = 10,302)		Male (n = 5,610)		Female (n = 4,692)	
	Mean	SD	Mean	SD	Mean	SD
Age	34.89	9.83	35.86	9.99	33.91	9.56
Years of schooling completed (yearsch)	10.35	3.52	10.21	3.63	10.92	3.34
Education level						
No/informal qualification	0.03	0.18	0.04	0.21	0.02	0.14
Primary education	0.12	0.33	0.13	0.33	0.12	0.33
Lower secondary	0.25	0.43	0.28	0.45	0.21	0.41
Upper secondary	0.38	0.49	0.36	0.49	0.41	0.49
Diploma	0.13	0.34	0.11	0.31	0.15	0.36
University	0.09	0.29	0.08	0.29	0.09	0.29
Work experience in month (*Exp*)	165.45	120.05	181.26	123.15	149.38	114.61
Attended training at the workplace, *Training* (0 = No, 1 = Yes)	0.42	0.49	0.43	0.50	0.40	0.49
Male (0 = No, 1 = Yes)	0.55	0.45				
Married respondent (0 = No, 1 = Yes)	0.65	0.48	0.68	0.47	0.62	0.49
Ethnicity						
Malay	0.55	0.50	0.58	0.49	0.52	0.50
Chinese	0.35	0.48	0.33	0.47	0.39	0.49
Indian	0.10	0.29	0.09	0.29	0.10	0.30
Region						
Central	0.35	0.48	0.35	0.48	0.34	0.47
North	0.23	0.42	0.24	0.42	0.23	0.42
South	0.33	0.47	0.31	0.46	0.34	0.47
East coast	0.03	0.16	0.03	0.18	0.02	0.13
Malaysia East	0.07	0.25	0.07	0.25	0.07	0.25
Occupation						
Managerial	0.15	0.36	0.13	0.33	0.17	0.38
Professional	0.08	0.28	0.09	0.28	0.08	0.27
Skilled job	0.37	0.48	0.45	0.50	0.28	0.45
Clerical/Non-production	0.23	0.42	0.22	0.41	0.24	0.43
Unskilled job	0.17	0.38	0.12	0.32	0.23	0.42
Hours of work per week	45.82	12.23	46.81	12.56	44.81	11.81
Lack of skills in doing current job (0 = No, 1 = Yes)						
English communication skill	0.56	0.50	0.56	0.50	0.56	0.50
Professional communication skill	0.35	0.48	0.37	0.48	0.34	0.47
Team work skill	0.15	0.35	0.14	0.35	0.15	0.36
Leadership skill	0.17	0.37	0.16	0.36	0.18	0.38
Time management skill	0.14	0.35	0.14	0.35	0.14	0.35
Numerical skill	0.11	0.31	0.11	0.31	0.11	0.31
Problem solving skill	0.15	0.35	0.14	0.34	0.16	0.37
Information technology skill	0.44	0.50	0.46	0.50	0.41	0.49
Technical skill	0.38	0.49	0.36	0.48	0.41	0.49
Industry						

TABLE 1. (*Continue*)

Variable	All (n = 10,302)		Male (n = 5,610)		Female (n = 4,692)	
	Mean	SD	Mean	SD	Mean	SD
Food processing	0.22	0.41	0.23	0.42	0.21	0.41
Textiles	0.04	0.19	0.04	0.19	0.04	0.19
Garments	0.07	0.26	0.02	0.15	0.12	0.33
Chemical	0.08	0.27	0.09	0.28	0.07	0.25
Rubber & plastics	0.25	0.44	0.25	0.43	0.26	0.44
Machinery & equipment	0.09	0.28	0.12	0.32	0.05	0.23
Electric & electronic	0.04	0.18	0.03	0.18	0.04	0.19
Auto parts	0.11	0.31	0.11	0.31	0.11	0.31
Wood & furniture	0.11	0.31	0.11	0.32	0.10	0.31
Hourly wages - (RM)	11.71	6.05	13.21	3.82	10.19	8.04

Samples used in this study however, are restricted to respondents who were in full-time employment, aged between 15 and 64 and who reported no missing in wages. By such restriction, this leaves about 10,302 respondents, of which 54.5 percent are males and 45.5 percent are females across nine major industries. Table 1 shows the descriptive statistics for some of the key variables (see Appendix section for variable definition). Generally, respondents were reported to have about 11 years of schooling on average which is equivalent to Malaysia Certificatie Examination (MCE) qualification. However, women are slightly better educated with 24 percent of them possessing higher degree qualifications (both diploma and university qualification) relative to 19 percent among men. Meanwhile, men have more work experience than women (*181 months vs. 149 months*, respectively). Furthermore, about 25 percent of the women occupied higher job levels (management and professional) with a corresponding figure of 22 percent for men. In contrast, men seem overrepresented in the skilled job with 45 percent relative to 28 percent for women. Though, women are twice as likely as men to be working in non-production level jobs, i.e. clerical jobs. Meanwhile, men are found to have more working hours than women every week, roughly 2 hours more than for women. Perhaps, the most striking figure revolves around wages where women earn about 83 percent of male wages.

TABLE 2A. Percentage of the raw responses of "*your current job offers you sufficient scope to use your knowledge and skills*"

	All (n = 10,302)	Male (n = 5,610)	Female (n = 4,692)
Do not agree at all	8.1	8.8	7.3
Somewhat agree	22.9	23.5	22.1
Agree	54.7	53.0	56.7
Agree completely	14.3	14.7	13.9
Total	100	100	100

TABLE 2B. Percentage distribution of the incidence of overskilling

Overskilling	All (n = 10,302)	Male (n = 5,610)	Female (n = 4,692)
Well-matched	69.0	67.7	70.6
Moderately overskilled	22.9	23.5	22.1
Severely overskilled	8.1	8.8	7.3
Total	100.0	100.0	100.0

To measure overskilling, here, we utilise a worker's own assessment approach where a respondent was asked about "*Your current job offers you sufficient scope to use your knowledge and skills".* Four responses were available, from 1 (do not agree at all), 2 (somewhat agree), 3 (agree), and to 4 (agree completely).[8] The percentage of the corresponding responses are 8.1 percent, 22.9 percent, 54.7 percent and 14.3 percent respectively (see Table 2a). There is little difference across gender. Then, we regrouped the four responses into three to create a new variable, termed overskilling (Table 2b). Here those with response 1 are classified as severely overskilled, those with response 2 on the scale are called as moderately overskilled, and well-matched for those with responses 3 and 4 on the scale.[9] The extent of overskilling, i.e. both severely and moderately overskilled in Malaysia is relatively low as compared to 44 percent in Mavromaras et al. (2010).

We have to acknowledge that the advantage of the workers' own assessment is that the method incorporates all information about a respondent's specific job as the worker is the only one who knows the best position to understand the skill requirements of an occupation. But the problem is respondents may lack sufficient benchmarks against which to assess their job requirements, as may be evident for young workers who have little work experience. Furthermore, whether or not workers are evaluating the actual skill required to get or to do the job may be unclear. Indeed, workers may inflate or overstate the requirements of the jobs as a form of self-worth, which may lead to an under- or overestimation of overskilling.

Nevertheless, we also plot separately the percentage of moderately (Figure 1a) and severely overskilled workers (Figure 1b) across the wage distribution for male and female. These plots give us a good preliminary indication of the relationship between overskilling and unobserved ability. There is certainly evidence to support the notion that overskilling is heavily inversely correlated with ability, much apparent in the case of severely mismatched. On average, between 8 and 12 percent of males and females in the bottom three quantiles of the wage distribution are classified as being severely overskilled (Figure 1b). However, the percentage continuously decline as we moved from the left to the right of the distribution. Within the mid-ability range, i.e. - 4th to 6th decile, the average incidence stands roughly at 4 percent and falling off to 3 percent on average in the top three quantiles.

For moderately overskilled workers, Figure 1a also suggests that the fraction of moderately is higher at the bottom part of the wage distribution irrespective of gender. The percentage decreases gradually once we moved from left to the right of the distribution. Shortly, the percentages of moderately overskilled are 10 percentages points higher at the bottom part than at the upper part of the respondents' wage distribution.

Above all, the inter-relation between wages, unobserved ability and overskilling can be explored in both figures as the figures provide a more flexible approach to characterising the effect of overskilling on different percentiles of the conditional wage distribution. The wages for moderately and severely overskilled are higher for the lower quantiles of the conditional wage distribution. Therefore, both figures offer some evidence of what so-called 'substitutability' between overskilling and unobserved ability. Workers who are moderately and severely overskilled located at the lower end of the conditional wage distribution

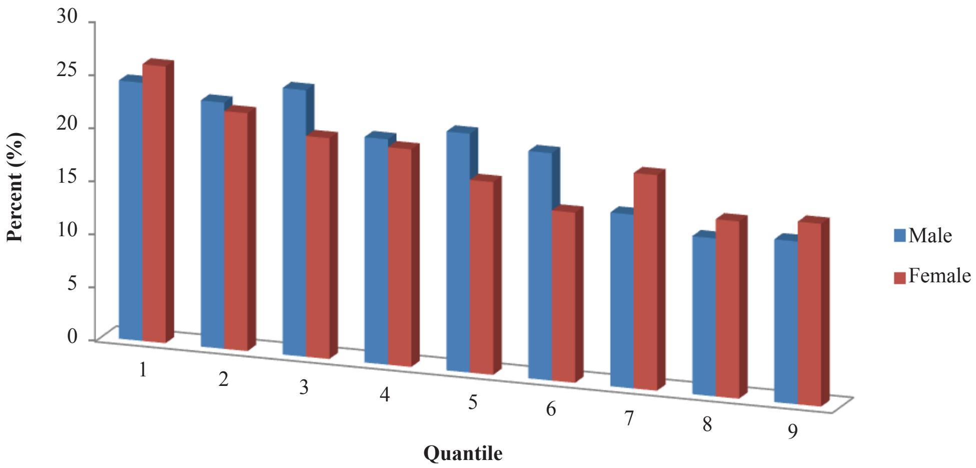

FIGURE 1a. Percentage distribution of moderately overskilled workers across wage quantile by gender

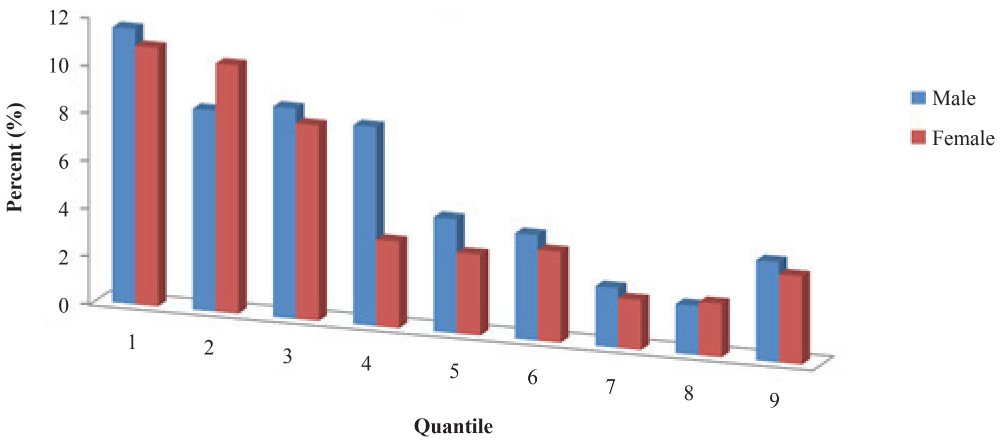

FIGURE 1b. Percentage distribution of severely overskilled workers across wage quantile by gender

TABLE 3. Means and standard deviation of hourly wages by incidence of overskilling (RM)

	All		Male		Female	
	Means	SD	Means	SD	Means	SD
Skills mismatch						
Well-matched	12.33	9.00	13.96	6.14	10.65	3.61
Moderately overskilled	10.70	7.39	11.80	7.16	9.61	7.57
Severely overskilled	8.32	1.39	9.45	1.99	7.28	1.73

will, *ceteris paribus*, have unobserved characteristics, leading to lower wages.

Table 3 shows preliminary findings of the relationship between overskilling and wages. It is clear that the incidence of overskilling result in substantially lower wages, especially for the severely overskilled workers. Data shows that moderately overskilled workers earn about RM1.04 to RM2.16 less than their well-matched counterpart regardless of gender. The corresponding figure of severely overskilled workers was roughly RM3.37 to RM4.51. This implies that wages for those severely overskilled is about 32 percent less than that of the well-matched group. Therefore, we hope that to what extent the return to overskilling based on raw data will be tested empirically in the regression analysis.

EMPIRICAL METHOD

To estimate the wage impact of overskilling, we utilised an augmented human capital model following Mavromaras et. al (2010). The model can be written as follow:

$$\ln w_{ij} = \alpha_0 + \alpha_1 X_{ij} + \gamma_1 MOS_{ij} + \gamma_2 SOS_{ij} + \delta Z_j + \varepsilon_{ij} \qquad (1)$$

$i = 1, \ldots$ individual N $\qquad j = 1, \ldots$ firm J

where ln w is a natural logarithm of wages (hourly), X is a vector of explanatory variables (as in Table 1 such as work experience, educational level, gender, marital status, region, occupation and types of industry), *MOS* and *SOS* respectively denote dummies for moderate overskilling and severe overskilling, for individual i at firm j with the well-matched being the reference group. Z_j is a set of characteristics describing the workplace and firm at which individual i is employed (such as firm size, foreign ownership, union and age of establishment). α, γ and δ represent the estimated returns to characteristics. All parameters are estimated using OLS. Given the literature and the descriptive statistics in Table 3, we hypothesise that $\gamma_1 < 0$ and $\gamma_2 < 0$.

OLS allows the effect of overskilling to be estimated on the mean of the conditional individuals' wages distribution. As we mentioned earlier, it implicitly assumes that the impact of the explanatory variables along that conditional distribution is constant. This fact is referred to as a pure location shift. In other words, the control variables are unable to cause a scale effect or any other consequence on the distributional shape. However, as covariates may influence the conditional distribution of the response in many other ways, we are willing to estimate the wage effect over the whole distribution using a quantile regression. The model can be written as the following functional form (McGuinness & Bennett, 2007):

$$\ln w_i = X_i \beta_\theta + \mu_{i\theta} \quad with \quad Quant_\theta(\ln w_i|X_i) = X_i \beta_\theta \qquad (2)$$

where X_i is a vector of exogenous variables for individuals as explained in equation (1). $Quant_\theta$ (ln $w_i|X_i$) denotes the θ_{th} conditional quantile of log hourly wage (ln$w$) when X is given. The distribution function of the error term for the θ_{th} quantile is left unspecified but with the assumption that $Quant_\theta(\mu_\theta|X_i) = 0$. The θ_{th} regression quantile, $0 < \theta < 1$, is defined as the solution of problem:

$$\min \beta\varepsilon R^k \left(\sum_{i:\ln wi \geq Xi\beta} \theta|\ln w_i - Xi\beta_\theta| + \sum_{i:\ln wi < Xi\beta} (1-\theta)|\ln w_i - Xi\beta_\theta| \right) \qquad (3)$$

The above equation could be simplified as:

$$\min \beta\varepsilon R^k \sum_i \rho_\theta(\ln w_i - X_i\beta_\theta) \qquad (4)$$

where $r_{\theta(e)}$ is the check function defined as $r_{\theta(e)} = \theta e$ if $e \geq 0$ or $r_{\theta(e)} = (\theta - 1)\, e$ if $e < 0$. It should be noted that the QR models are estimated with the same controls as the OLS specification. The first quantile is obtained by setting θ = 0.1 and followed by 0.25 0.5 0.75 and 0.90. As θ increases, the entire distribution of wages for individual $i(w_i)$ is traced, conditional on X.

This econometric approach is, therefore, a working compromise to offer evidence and exploring the interplay between overskilling and unobserved characteristics as well. We might presume that controlling for unobserved heterogeneity among workers would explain some of the difference in wages between well-matched and overskilled workers. This is in line with our observation as shown in Figure 1a and 1b.

Nevertheless, the interpretation of the quantile regression coefficients is theoretically quite similar to the OLS. In the latter, the regression coefficient measures

TABLE 4. Wage effects of overskilling

Log hourly wages	Model 1		Model 2		Model 3		Model 4		Model 5	
Incidence of overskilling (ref-Well-matched)										
Moderately-overskilled	-0.035	**	-0.044	***	-0.044	***	-0.022	*	-0.023	*
	0.017		0.017		0.017		0.012		0.012	
Severely-overskilled	-0.137	***	-0.132	***	-0.129	***	-0.116	***	-0.099	***
	0.027		0.026		0.026		0.017		0.017	
Education (ref-No/Primary)										
Lower secondary	0.177	***	0.114	***	0.112	***	0.107	***	0.106	***
	0.023		0.022		0.022		0.016		0.016	
Upper secondary	0.355	***	0.268	***	0.260	***	0.183	***	0.165	***
	0.022		0.022		0.022		0.016		0.016	
College diploma	0.873	***	0.657	***	0.638	***	0.409	***	0.370	***
	0.029		0.029		0.030		0.022		0.022	
Univ degree	1.244	***	0.991	***	0.963	***	0.656	***	0.606	***
	0.032		0.033		0.034		0.026		0.026	
Work experience in month (*Exp*)	0.004	***	0.003	***	0.003	***	0.001	***	0.001	***
	0.000		0.000		0.000		0.000		0.000	
Work experience square in month (Exp^2)	-0.000	***	-0.000	***	-0.000	***	-0.000	***	-0.000	***
	0.000		0.000		0.000		0.000		0.000	
Training	0.143	***	0.137	***	0.137	***	0.109	***	0.060	***
	0.015		0.015		0.015		0.011		0.012	
Female (0=No, 1=Yes)	-0.192	***	-0.187	***	-0.184	***	-0.247	***	-0.214	***
	0.014		0.022		0.022		0.015		0.015	
Demographic background	No		Yes		Yes		Yes		Yes	
Skills to perform current job	No		No		Yes		Yes		Yes	
Job characteristics	No		No		No		Yes		Yes	
Workplace characteristics	No		No		No		No		Yes	
_cons	1.092	***	1.292	***	1.123	***	3.175	***	3.046	***
	0.024		0.034		0.067		0.054		0.064	
N	10,243		10,046		10,046		9,952		9,787	
R-square	0.264		0.353		0.355		0.677		0.691	
Adjusted R-sq	0.263		0.351		0.353		0.676		0.689	
Log-likelihood ratio test (χ^2)			1667.46	***	43.13	***	7097.53	***	603.09	***

Robust standard error in italics
* , **, and *** denote 0.1,0.05, and 0.01, respectively

the influence of the explanatory variables based on the conditional mean of the explained variable. Instead, for the QR, the coefficient of β_θ indicates the influence of the predictor variables based on the conditional θ quantile of the predicted variable. Estimation is by minimising the sum of weighted absolute deviations and can be performed using linear programming methods. An estimated variance-covariance matrix for the chosen system of quantile regressions is obtained using a bootstrap re-sampling method using Stata 13.

There are five specifications proposed. In the first specification, we are only control for the basic individuals' human capital edowment and gender, and in the second specification, we add the demographic background (such as ethnic group, marital status, the present of children under 12 years old and region) as explanatory variables. For specification three, we further included individuals' skills to perform current job and followed up by job characteristics and workplace attributes in fourth and fifth specifications, respectively.

EMPIRICAL FINDINGS

Table 4 presents the result from the OLS method. There are five specifications examined, and all find that the coefficient on overskilling is negative and significant with smaller (larger) effects for moderate (severe) overskilling. Focus firstly on specification 1; when controlling the individuals' human capital and gender, the moderately overskilled earn about 4 percent less than their well-matched counterparts.10 This penalty is about 13 percent for the severely overskilled. These penalties remain unchanged even after controlling the demographic background (model 2) and individuals' skills (model 3). The inclusion of job characteristics and workplace attributes reduces the wage penalty of the overskilled to about 2 and 10 percent respectively for the moderately and severely overskilled. The addition of job characteristics (model 4) does have an impact since the wage penalty falls by around four percentage points for moderately overskilled and by around ten percentage points amongst the severely overskilled.

TABLE 5. Wage effects of overskilling by gender

Log hourly wage	Male		Female	
Incidence of overskilling (ref-well-matched)				
Moderately-overskilled	-0.024		-0.023	*
	0.017		0.013	
Severely-overskilled	-0.103	***	-0.083	***
	0.024		0.026	
Education (ref-No/primary)				
Lower secondary	0.087	***	0.113	***
	0.020		0.025	
Upper secondary	0.124	***	0.210	***
	0.021		0.024	
College diploma	0.332	***	0.409	***
	0.032		0.031	
Univ degree	0.575	***	0.645	***
	0.037		0.037	
Work experience in month (*Exp*)	0.002	***	0.001	***
	0.000		0.000	
Work experience square in month (Exp^2)	0.000	***	0.000	***
	0.000		0.000	
Training	0.050	***	0.085	***
	0.017		0.017	
Constant	3.198	***	2.729	***
	0.089		0.090	
N	5,273		4,514	
R-square	0.685		0.702	
Adjusted R-sq	0.681		0.697	

Robust standard error in italics
* , **, and *** denote 0.1,0.05, and 0.01, respectively

We now discuss briefly the effects of other variables. Wages is positively associated with education, work experience, and training. For education, the higher the education, the greater the wages received. The coefficients on work experience across all results are positive, while the estimated coefficients on work experience squared show negative signs indicate that wage increases with work experience but at a diminishing rate. The coefficients on training are positive and significant meaning that training is positively associated with wages. Women earn significantly lower than their men counterparts across the five specifications. For example, specification 5 suggests that women earn about 22 percent less than that of men. The greater wages loss for women, also found in elsewhere in Malaysia (see Amin 2004; Amin & DaVanzo 2004; Aminah Ahmad 2009; Ismail 2011; Ismail & Noor 2005) .11

Table 5 presents the wage impacts of overskilling separately for men and women (full sample, i.e.- Specification 5). If compared to previous analysis, the coefficient on moderate overskilling is insignificant for the male sample suggesting no male pay penalty for being moderately overskilled. By contrast, there is evidence that moderately overskilled women are paid about 2 percent less than their well-matched counterparts. Turning to severely overskilled, the coefficients are negative and significant for both men and women. The wage penalty seems to be higher for men than women (10 percent versus 8 percent). The results are in line with Mavromaras et al. (2010) who found that the penalty for being severely overskilled was higher for men than for women, especially with respect to Australian workers (9 percent against 6 percent).

From Table 4 and 5, it is clearly shown that an allowance for individuals' skills does not appear to have any impact on the overskilling penalty since the penalty remains largely unchanged. This could be interpreted as overskilling penalty might be related to individuals' unobserved heterogeneity such as unobserved ability. To ascertain this, Table 6 presents the wage impact of overskilling across the conditional wages distribution (Specification 5). It should be acknowledged that we only discuss the effects of overskilling on wages as our prime interest. Therefore, the results of the effects of other variables are not discussed here but available upon request.

Similar to the findings for OLS, the quantile regression estimates also reveal a wage penalty that is smaller for moderate than severe overskilling. However, compared to the OLS, the pooled sample suggests that the wage penalty for the moderately overskilled is only confined to the lower and mid-range of the conditional wages distributions, and there is no penalty evident at the top two deciles. The Wald test (Table 7) indicates

TABLE 6. Quantile regression and wage impacts of overskilling

log wage hourly	q10		q25		q50		q75		q90	
Pooled										
Moderately overskilled	-0.067	***	-0.055	***	-0.039	**	-0.010		0.015	
	0.020		0.013		0.017		0.020		0.025	
Severely overskilled	-0.085	***	-0.113	***	-0.125	***	-0.090	***	-0.097	**
	0.026		0.021		0.023		0.033		0.040	
N	9,787		9,787		9,787		9,787		9,787	
Male										
Moderately overskilled	-0.087	***	-0.039	**	-0.023		0.004		0.041	
	0.024		0.019		0.015		0.030		0.033	
Severely overskilled	-0.070	*	-0.094	***	-0.097	***	-0.045		-0.041	
	0.039		0.034		0.037		0.044		0.060	
N	5,273		5,273		5,273		5,273		5,273	
Female										
Moderately overskilled	-0.044		-0.053	**	-0.057	***	-0.029		0.034	
	0.033		0.023		0.018		0.020		0.033	
Severely overskilled	-0.092	**	-0.097	*	-0.130	***	-0.151	***	-0.170	**
	0.036		0.057		0.026		0.046		0.067	
N	4,514		4,514		4,514		4,514		4,514	

Standard errors in italics
** , **, and *** denote 0.1,0.05, and 0.01, respectively*

TABLE 7. The equality test of quantile regresions (F-Wald statistics)

	Pooled		Male		Female	
	equality test	Inter-quantile test	equality test	Inter-quantile test	equality test	Inter-quantile test
Incidence of overskilling (Table 6 and 7)						
Moderately overskilled	3.82***	10.55***	3.15***	3.87**	5.15***	9.02***
Severely overskilled	0.82	0.07	1.18	0.07	2.31**	4.01***

* , **, and *** denote significant at 0.1,0.05, and 0.01, respectively.

that the wage impacts of moderate overskilling are heterogeneous in different parts of the conditional wages distribution.[12] For the severely overskilled, the wage loss appears across the conditional wages distribution and has an inverted u-shape; highest at the median decile, approximately 12 percent, and lowest at the bottom and top end decile, about 8 percent and 9 percent respectively. By gender, the wage impact of moderate overskilling seems to follow the general sample. For severe overskilling there is a gender difference. For men, the penalty is concentrated at the first 50 deciles, on average 9 percent and there is no evidence of the penalty at the top end decile. On the other hand, the wage penalty (severely) for women increases linearly with quantile, from 9 percent at the 10th decile to 17 percent at the 90th decile which means that the penalty at the top end is seven percentage points higher than that at the bottom decile.[13]

Above all, the quantile regression analysis provides some evidence that the wage penalty of overskilling is partly attributable to unobserved individual heterogeneity such as ability.14 This is particularly true for the male sample. The penalty for overskilled men is more crowded into the lower segment of the distribution and largely disappears or is insignificant as one moves towards the end of the distribution. The wage penalty for the overskilled female is not particular to the lower quantiles but also occurs across the female's conditional wages distribution. This to some extent suggests that a higher wage penalty for the overskilled women is not due to a consequence of omitted unobserved ability, but is due perhaps to them being genuinely mismatched.

CONCLUSION

This paper investigates the incidence and wage effects of overskilling on individuals' wages in the Malaysian labour market using quantile regression approach. Previous research indicated that wage effect of overskilling upon the mean wage, i.e.- OLS estimation is far from perfect to capture the unobserved ability if such elements vary over the workers' wages distribution. Instead, the QR enables us to explore the effect of overskilling for each point of the workers' wages distribution. If the reason why the overskilled workers earn less is that they are low-ability than that of well-matched workers, then once unobserved ability is controlled for, the penalty of overskilling should be reduced as we moved up along the wages distribution.

Nevertheless, the incidence of overskilling was reported around 31 percent - moderately and severely overskilled workers represented 23 and 8 percent irrespective of gender. However, the incidence of overskilling was very heavily concentrated amongst lower ability segments of the wage distribution, particularly severely overskilled workers (see Figure 1b). Only 3 percent of males and females in the top three quantiles classified as overskilled relative to 9 percent at the lower quantile. Looking at the wages outcomes, the OLS estimation suggests that being overskilled results in a 2 percent to 10 percent wages loss compared to well-matched workers. A higher loss of wages is found for the severely overskilled relative to the moderately overskilled workers. By gender, men faced a higher penalty than their women counterparts.

Using a quantile regression, the penalty for overskilled workers was heterogeneous and not constant over the conditional wages distribution. This was particularly true for the severely overskilled workers. Instead, the wages loss for moderately overskilled was only evident at the lower and mid-range of the conditional wage distributions. For the severely overskilled, the penalty was evident throughout the conditional wages distribution, and has an inverted u-shape; highest at the median decile, and lowest at the bottom and top end decile. There is, however, a gender matter where the penalty for men was only evident at the first 50 deciles and no evidence at the top end decile. Contrary, for the severely overskilled women, the penalty increased linearly with quantile where the penalty at the top end is seven percentage points higher than that at the bottom decile.

Therefore, this is not true to say that overskilling was absolutely exclusive to low-ability workers. This may suggest overskilling is a phenomenon that is imposing a real wage cost for a particular group, once they are overskilled, this would reduce their potential productivity regardless of their ability distribution. Nevertheless, the fact that the wage penalty for moderately overskilled workers was not evident at the top of the conditional individuals' wages distribution may suggest the importance of including explicit controls for unobserved heterogeneity, i.e.- individuals' unobserved ability where possible, in order to avoid overstating the wage impacts of the phenomenon.

By differentiating between quantiles, we discriminated between groups of workers with different (unobservable) skills. We found that the detrimental effects of overskilling among the high-skilled are, if not higher, as large as amongst the low skilled. This was interpreted as evidence that overskilling is an event that reduces the worker's potential productivity, regardless of their skills.

Similar to overeducation phenomenon, the impact of overskilling on individuals' wages might be an important policy issue for the government as well, given the fact the significant rise in enrolment in higher education in the last 10 years, both in public and private HEIs.15 While the number of graduates has been in an increasing trend, the technological changes have been taking place during that period, leading to the upgrading of skills requirement of jobs. One the one hand, both trends suggest that education systems in Malaysia and labour markets have been progressing in harmony. On the other hands, one should have admitted that there is substantial possibility for serious mismatches, with respect to economic outcomes such as wasting of resources and the costs inherent for not meeting the economic potential of an economy due to the lack of appropriate human capital.

Therefore, embedding soft skills elements in the curriculum design at the school level will probably be a better solution so that students may have some soft employability skills. The introduction of Integrated Cummulative Grade Point Average (ICGPA) at university level should be extended to all higher education institutions as the IGCPA recognizes academics and non-academic skills of each students, and allow students to gain necessary skills early prior to graduation which are particularly important in the modern labour market. Moroever, fostering industrial training programmes at higher education institutions by extending the attachement to all programs of study might be useful as the program may provide better platform for graduates to find jobs that match their skills background. Apart from gaining practical experience, students will realise the importance of possessing a certain level of soft skills and make time to improve the skills they lack prior to entering the job market. From the individuals' perspective, overskilled workers must have also initiative in the first place to improve or upgrade their human capital stock or skills. This can be done through attending off-the-job training or off-work courses or by participating in any training programme at the workplace. Alternatively, upon completing study at the tertiary level, individuals should take part in any skills development course offered by their institutions. This will help them improve their employability skills before entering the job market. Workers who equip themselves with extra training and skills have a better chance of getting higher positions (i.e., positions that are commensurate to their skills background) within their organisations.

ACKNOWLEDGEMENT

The authors acknowledge the World Bank Enterprise Survey (WBES) from whom the 2007 Productivity Investment Climate Survey (PICS) data was acquired. This organisation has no bear any responsibility for the authors' analysis and interpretations of the data.

NOTES

1. Details about this problem has been addressed in Dolton and Silles (2001); Bauer (2002); McGuinness(2006); Chevalier and Lindley (2009); Korpi and Tåhlin, (2009).
2. Overskilling is a counterpart of over-education and can be defined the extent to which workers are not able to utilise all their skills and knowledge gained in their current employment.
3. A detail discussion about this approach can be found in Koenker & Bassett Jr (1978) and Koenker, (2004).
4. Relative to McGuinness & Bennett (2007), Hartog et al. (2001) employed the Overeducation, Required and Undereducation (ORU) specification in estimating the effect of over-education on earnings. For detail about this approach, see Hartog, (2000).
5. Propensity score matching is a way to "correct" the estimation of treatment effects controlling for the existence of these confounding factors based on the idea that the bias is reduced when the comparison of outcomes is performed using treated and control subjects who are as similar as possible (Becker & Ichino, 2002).
6. This analysis allows assessing the extent to which an unobserved variable influenced the selection process in order to render the matching estimates unreliable.
7. There are nine major manufacturing industries covered in the PICS-2: food processing, textiles and garments, wood and furniture, chemical and chemical products, rubber and plastics, machinery and equipment, electrical machinery and electronics, equipment and components, and motor vehicles and parts.
8. The question is similar to that used in both the work of Allen & van der Velder (2001) and Green & McIntosh (2007) and Mavromaras et. al (2010) .
9. In line with overskilling literature (see Mavromaras et. al, 2010), we decided to cluster these scales into three categories as the statement captures the degree of skills utilised. The respondent might have answered this question with the notion that he or she possesses a skill set that is deemed to be irrelevant to the job held which leads to "underskilled" or skill deficits rather than "overskilled". But, in the PICS, there is another question available in capturing the degree of skills deficit, where the respondents were asked about "*You would perform better in your current job if you possess additional knowledge and skills*" followed by a four response - 1 (do not agree at all), 2 (somewhat agree), 3 (agree) and 4 (agree completely). We found that over 80% of respondents have a skills deficit (tho who agreed and completely agreed) in the jobs they currently occupy. We then run a simple Spearman rank order test to figure out how strong the correlation between them and found that the correlation was very low, at 0.097, and not statistically significant at the 10% level. The test might suggest that both statements capture different traits – one for overskilling and another one for skill deficits.
10. Since the wages regression specification is in semi-logarithmic form, the percentage point effect (PE) is obtained using the following formula:

    PE = $(e^{\beta} - 1) \times 100$, where β is the coefficient estimate.

    The percentage point effect will be used throughout the discussion in this article.
11. Other controlled variables are demographic backgrounds (ethnic group, marital status, the present of childeren under 12 years old, citizenship, distance to workplace and region), skills to perform current job (adaptability, communication, language profeciency, problem solving, computer literacy, teamwork and numerical skill), job characteristics (occupation, tenure, hours of work, semployment sector and employment status) and workplace attributes (firm size, ownership, union, age of establishement, exporter firms and capital intensive firm). For a full regression, it is available upon request.
12. The Wald test shows no evidence that the pay penalties are heterogeneous at different parts of the worker conditional wage distribution.
13. The inter-quantile test suggests that the pay penalty between the top end and bottom decile is statistically significantly different from zero.
14. We have to acknowledge that we reregressed all model specifications, this time we replaced level of education with years of schooling completed. We found that the results did not change too much in terms of sign and magnitude of the interested coefficient across sample.
15. In 2015, there were nearly 1.24 million students enrolled across public and private HEIs, increased from 921,548 in 2008 (2015 Higher Education Statistics, Ministry of Higher Education)

REFERENCES

Allen, J., & Velder, R. van der. 2001. Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers* 53(3): 434–452. https://doi.org/10.1093/oep/53.3.434

Amin, S. 2004. Ethnic differences and married women's employment in Malaysia: do government policies matter? *The Journal of Socio-Economics* 33(3): 291–306. https://doi.org/10.1016/j.socec.2003.12.023

Amin, S., & DaVanzo, J. 2004. The impact of wives' earnings on earnings inequality among married-couple households in Malaysia. *Journal of Asian Economics* 15(1) : 49-70. https://doi.org/10.1016/j.asieco.2003.12.002

Aminah Ahmad. 2009. Gender Differences and Trends in the Participation of Malaysians in Education : Implications on Employment Outcomes. *The Journal of International Management Studies* 4(2): 65-74.

Bauer, T. K. 2002. Educational mismatch and wages : A panel analysis *19*: 221-229.

Becker, S. O., & Ichino, A. 2002. Estimation of average treatment effects based on propensity scores. *The Stata Journal 2*(4): 358-377. article.

Budría, S., & Moro-Egido, A. I. 2008. Education, educational mismatch, and wage inequality: Evidence for Spain. *Economics of Education Review* 27(3): 332-341. https://doi.org/10.1016/j.econedurev.2006.10.005

Chevalier, A., & Lindley, J. 2009. Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172(2): 307-337. https://doi.org/10.1111/j.1467-985X.2008.00578.x

Dolton, P., & Silles, M. 2001. Over-Education in the Graduate Labour Market: Some Evidence from Alumni Data, (June).

Green, F., & McIntosh, S. 2007. Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics* 39(4): 427-439. https://doi.org/10.1080/00036840500427700

Hartog, J. 2000. Over-education and earnings: Where are we, where should we go? *Economics of Education Review* 19(2):131-147. https://doi.org/10.1016/S0272-7757(99)00050-3

Hartog, J., Pereira, P. T., & Vieira, J. a. C. 2001. Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators. *Applied Economics 33*(8):1021-1037. https://doi.org/10.1080/000368401300182798

Ismail, R. 2011. Gender wage differentials in the Malaysian services sector. *African Journal of Business Management* 5(19) : 7781–7789. https://doi.org/10.5897/AJBM10.047

Ismail, R., & Noor, Z. 2005. Gender wage differentials in the Malaysian manufacturing sector. *International Journal of Economics and Management* 2(2): 119-137. Retrieved from http://journals.iium.edu.my/enmjournal/index.php/enmj/article/view/110

Koenker, R. 2004. Quantile regression for longitudinal data. *Journal of Multivariate Analysis* 91(1): 74-89. article.

Koenker, R., & Bassett Jr, G. 1978. Regression quantiles. *Econometrica: Journal of the Econometric Society*, 33-50. article.

Korpi, T., & Tåhlin, M. 2009. Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974-2000. *Labour Economics* 16(2): 183-193. https://doi.org/10.1016/j.labeco.2008.08.004

Lim, H. 2011. The determinants of individual unemployment duration: the case of Malaysian graduates. *Journal of Global Management* 2(1): 184-203. Retrieved from http://www.globalresearch.com.my/journal/management_v02n01n02/JoGM_July2011_V2N2.pdf#page=66

Lim, H. 2013. Overeducation and happiness in the Malaysian graduate labour market. *International Journal of Business & Society* 14(1): 93-110. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:OVEREDUCATION+AND+HAPPINESS+IN+THE+in+the+malaysian+graduate+labour+market#0

Martins, P. S., & Pereira, P. T. 2004. Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics* 11(3): 355-371. https://doi.org/10.1016/j.labeco.2003.05.003

Mavromaras, K., Mcguinness, S., & Fok, Y. K. 2009. The incidence and wage effects of overskilling among employed VET graduates. NCVER Monograph Series 03/2009. National Centre for Vocational Education Research Ltd. PO Box 8288, Stational Arcade, Adelaide, SA 5000, Australia.

Mavromaras, K., Mcguinness, S., O'Leary, N., Sloane, P., & Fok, Y. K. 2010. the Problem of Overskilling in Australia and Britain. *The Manchester School* 78(3): 219-241. https://doi.org/10.1111/j.1467-9957.2009.02136.x

McGuinness, S. 2006. Overeducation in the labour market. *Journal of Economic Surveys* 20(3): 387-418.

McGuinness, S., & Bennett, J. 2007. Overeducation in the graduate labor market: A quantile regression approach. *Economics of Education Review* 26(1): 521-531.

McGuinness, S., & Bennett, J. 2007. Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review* 26(5): 521-531. https://doi.org/10.1016/j.econedurev.2005.12.003

McGuinness, S., & Sloane, P. J. 2011. Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review* 30(1): 130-145. https://doi.org/10.1016/j.econedurev.2010.07.006

World Bank. 2009. *World Bank. 2009. Malaysia productivity and investment climate assessment update. No. 49137-MY.* Washington, DC.

Zakariya, Z. 2014. The Effects of Over, Required and Under-Education on Earnings in Manufacturing Sector in Malaysia. *International Journal of Management Studies* 21(1): 83-109.

Zakariya, Z. 2014b. Wage effect of over-education and mismatch in Malaysia: a random effect approach. *Jurnal Ekonomi Malaysia* 48(2): 3-17. http://journalarticle.ukm.my/8508/

Zakariya, Z., & Battu, H. 2013. The effects of overeducation on multiple job satisfaction towards enhancing individual s ' well-being in Malaysia. *Business and Management Quarterly Review* 4: 38-51.

Zakariya, Z., & Mohd. Noor, M. A. 2014. Workplace characteristics and determinants of over-education and overskilling in the manufacturing sector in Malaysia. *Jurnal Pengurusan* 40: 125-136.

Zulkifly, O., Ishak, Y., & Abu Hassan, M. S. 2010. Inflasi pendidikan mengikut ketentuan dalam pasaran buruh Malaysia. *Jurnal Ekonomi Malaysia* 44(1): 61-71.

Zainizam Zakariya*
Department of Economics,
Universiti Pendidikan Sultan Idris
35900 Tanjung Malim
Perak
E-mail: zainizam@fpe.upsi.edu.my

Norasibah Abdul Jalil
Department of Economics,
Universiti Pendidikan Sultan Idris
35900 Tanjung Malim
Perak
E-mail: norasibah@fpe.upsi.edu.my

Khoo Yin Yin
Department of Economics,
Universiti Pendidikan Sultan Idris
35900 Tanjung Malim
Perak
E-mail: khoo@fpe.upsi.edu.my

*Corresponding author

APPENDIX

List of key variables	Explanation
Age	Age of respondents (continuous variable)
Years of schooling completed (*yearsch*)	How many years of Formal Education have you completed? (continuous variable)
Education level	What is the highest level of formal education you attended?
No/informal qualification	No or informal education (0 = No, 1 = Yes)
Primary education	Primary education was the highest formal education attained (0 = No, 1 = Yes)
Lower secondary	Lower secondary was the highest formal education attained (0 = No, 1 = Yes)
Upper secondary	Upper secondary was the highest formal education attained (0 = No, 1 = Yes)
Diploma	Diploma qualification was the highest formal education attained (0 = No, 1 = Yes)
University	University qualification was highest formal education attained (0 = No, 1 = Yes)
Exp	Potential work experience in month (Age – 6 years - years of schooling completed) (continuous variable)
Training	Have you received formal training since you joined this firm?(0 = No, 1 = Yes)
Male	Male respondent (0 = No, 1 = Yes)
Married	Respondents' marital status (0 = No, 1 = Yes)
Ethnicity	**What is your ethnicity?**
Bumiputera	Bumiputera (0 = No, 1 = Yes)
Chinese	Chinese (0 = No, 1 = Yes)
Indian	Indian (0 = No, 1 = Yes)
Region	
Central	Central region (0 = No, 1 = Yes)
North	North region (0 = No, 1 = Yes)
South	South (0 = No, 1 = Yes)
East coast	East Coast region (0 = No, 1 = Yes)
Malaysia East	Malaysia East region (0 = No, 1 = Yes)
Occupation	**What kind of work are you doing now?**
Managerial	Managerial (0 = No, 1 = Yes)
Professional	Professional (0 = No, 1 = Yes)
Skilled job	Skilled jobs (0 = No, 1 = Yes)
Clerical/Non-production	Clerical or non-production jobs (0 = No, 1 = Yes)
Unskilled job	Unskilled job (0 = No, 1 = Yes)
Hours of work (weekly)	On average, how many hours a week are you working at present? (continuous variable)
Skills to perform current job	**What are the three skills that you lack the most in doing your job?**
English communication	English communication is the skill that respondents lack the most (0 = No, 1 = Yes)
Professional communication	Professional communication is the skill that respondents lack the most (0 = No, 1 = Yes)
Team work	Team work is the skill that respondents lack the most (0 = No, 1 = Yes)
Leadership	Leadership is the skill that respondents lack the most (0 = No, 1 = Yes)
Time management	Time management is the skill that respondents lack the most (0 = No, 1 = Yes)
Numerical	Numerical is the skill that respondents lack the most (0 = No, 1 = Yes)
Problem solving	Problem solving is the skill that respondents lack the most (0 = No, 1 = Yes)
Information technology	Information technology is the skill that respondents lack the most (0 = No, 1 = Yes)
Technical	Technical is the skill that respondents lack the most (0 = No, 1 = Yes)
Industry	**Type of industry**
Food processing	Food processing (0 = No, 1 = Yes)
Textiles	Textiles (0 = No, 1 = Yes)
Garments	Garments (0 = No, 1 = Yes)

APPENDIX (*Continue*)

List of key variables	Explanation
Chemical	Chemical (0 = No, 1 = Yes)
Rubber & plastics	Rubber & plastics (0 = No, 1 = Yes)
Machinery & equipment	Machinery & equipment (0 = No, 1 = Yes)
Electric & electronic	Electric & electronic (0 = No, 1 = Yes)
Auto parts	Auto parts (0 = No, 1 = Yes)
Wood & furniture	Wood & furniture (0 = No, 1 = Yes)
Wages (hourly wages in RM)	What is your current monthly salary in 2007? (before Tax, including all allowances and bonuses)