

DISCUSSION PAPER SERIES

IZA DP No. 14088

**Over-Education and Immigrant Earnings:
A Penalized Quantile Panel Regression
Analysis**

Sholeh A. Maani
Le Wen

JANUARY 2021

DISCUSSION PAPER SERIES

IZA DP No. 14088

Over-Education and Immigrant Earnings: A Penalized Quantile Panel Regression Analysis

Sholeh A. Maani

University of Auckland and IZA

Le Wen

University of Auckland

JANUARY 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Over-Education and Immigrant Earnings: A Penalized Quantile Panel Regression Analysis*

Despite evidence that immigrants experience a higher incidence of over-education, relatively few studies have considered the labour market outcomes of over-education for immigrants. Using longitudinal data and penalized quantile panel regression, we inspect the earnings effects of job mismatches for immigrants in Australia. This first application of the method to this question addresses both individual and distributional heterogeneity. Results confirm divergent effects across the earnings distribution. Immigrants from non-English speaking countries experience lower earnings returns and a substantial earnings penalty of up to 25 percent from educational mismatched employment. However, in contrast to conventional findings of a penalty based on regression at the mean, at the highest earnings quantile, both Australian-born and immigrants from non-English speaking countries with a host country qualification receive earnings premiums of 2.7 and 11 percent, respectively, from years of over-education. We discuss the policy implications of this new result.

JEL Classification: C23, I21, J24, J31

Keywords: penalized panel quantile regression, over-education, immigrants, overseas qualifications, language proficiency

Corresponding author:

Sholeh A. Maani
Department of Economics
The University of Auckland
12 Grafton Road
Auckland
New Zealand

E-mail: s.maani@auckland.ac.nz

* This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The HILDA data sets are made available to academics through the Australian Bureau of Statistics. The findings and views reported in this paper are those of the authors and should not be attributed to FaHCSIA or the Melbourne Institute.

1. Introduction

Immigrants comprise a significant proportion of the population in developed countries, and immigration is increasing internationally (Organisation for Economic Co-operation and Development (OECD), 2013). For example, in 2013, the percentage of the population that was foreign-born was 27.7 in Australia, 22.4 in New Zealand, 20 in Canada, 13.1 in the US, and 12.3 in the UK (OECD, 2013). As such, the labour market performance of immigrants has significant effects for both themselves and the economies of their host countries.

Another noticeable trend is that immigrants are increasingly more likely to have higher educational qualifications. In particular, the number of immigrants with university degrees in OECD countries increased by 70 percent within the first decade of the new millennium (United Nations, 2013). However, empirical evidence has confirmed that immigrants are more likely to be employed in mismatched jobs compared to the native-born population, and in positions where their educational qualifications exceed the level of education required (Aringa & Pagani, 2010; Fernández & Ortega, 2008; Kler, 2007). The narrative of doctors employed as taxi drivers is one example of such job mismatches.

It has also been empirically verified that educational mismatches are associated with earnings penalties (Aringa & Pagani, 2010; Chiswick & Miller, 2008, 2010; Cohn, 1992; Frenette, 2004; Green et al., 2007; Groot, 1996; Lindley, 2009; Rubb, 2003; Wald & Fang, 2008; Yeo & Maani, 2015), productivity decreases (e.g., Quinn and Rubb, 2006; Rumberger, 1987; Tsang & Levin, 1985), and less occupational mobility (Wen & Maani, 2019).

The extent to which over-education impacts the earnings of immigrants from non-English speaking backgrounds (NESB) and English speaking backgrounds (ESB), and the role of the transferability of human capital in this context, are important research questions for policy development. This is particularly true as more countries are now considering following Australia, Canada, and New Zealand in adopting skill-based migration policies. These questions also raise interesting methodological issues relating to the potential endogeneity and unobserved heterogeneity that are crucial to skill mismatch analysis. In addition, potential asymmetries of impacts across the earnings distribution raise further methodological issues.

Only a few studies have examined the impact of educational mismatches on the earnings penalties of immigrants within panel data settings. Indeed, the findings of the pioneering studies on job mismatches among immigrants have generally been based on pooled OLS (ordinary least squares) or cross-section analyses (see Leuven and Oosterbeek (2011) and Sloane (2014) for a discussion of these methods). Exceptions are: (a) Tsai (2010), who found no effect when worker heterogeneity is accounted for; (b) Dolton and Vignoles (2000) and Mavromaras et al. (2012), who found a reduced but significant effect; and (c) Carroll and Tani (2013) who found wage penalties for over-educated older workers. These studies considered general population groups and not immigrants. Yeo and Maani (2015), who specifically consider immigrant samples, found significant earnings penalties for job mismatched workers. The earnings penalty refers to each year of over-education receiving a positive, but at a lower rate than each year of required education compared to a well-matched employee.

These differences in findings, point to the need for international evidence based on panel data analyses across a wide range of studies. It also highlights the need to examine the potential drivers and the extent of earnings penalties associated with the over-education of immigrants. In particular, to what extent are earnings penalties asymmetric across the earnings distribution

spectrum is a relevant question that has not been examined within a panel data setting and is deserving of careful analysis. For example, whether the earnings penalty to job mismatches is different and more pronounced at the higher or lower ends of the earnings distribution is a question that we examine in this paper.

In addition, the effect of the transferability of human capital for immigrants has been verified in relation to immigrant earnings (Sarkar & Collier, 2019; Basilio et al., 2017; Friedberg, 2000), but not in connection with educational mismatches for immigrants. In particular, understanding the effect of the transferability of overseas academic degrees on the earnings penalties of immigrants merits attention in the analyses of job mismatches for immigrants, from both an advancing methodological perspective and for policy considerations. Moreover, the literature shows that host-country language proficiency is associated with a reduction in the probability of being over-qualified among immigrants (Budria & Martinez de Ibarreta, 2020). However, limited evidence exists on the impact of language proficiency on earnings associated with over-education.

In this paper, we examine the link between over-education and earnings. The following questions are of interest: How does over-education impact on earnings across the earnings distribution? Does having the highest qualification in the host country help ESB/NESB immigrants alleviate any earnings penalty to over-education? Is the impact of over-education on earnings affected by unobserved heterogeneity (e.g., personal ability and motivation, or employer perceptions of the qualifications of immigrants)? If these unobserved individual effects are correlated with explanatory variables, cross-section analysis would result in omitted unobservable biases.

Longitudinal studies capture data from the same individual over time. Thus, unobservable individual effects can be potentially eliminated by using a panel fixed effects model. Although we could use panel data techniques to answer the above questions, we also endeavor to investigate those effects that could vary based on the asymmetric features of the earnings distribution. The regression coefficients across quantile levels of earnings can provide useful information when strategically designing earnings equity policies for vulnerable groups. For this purpose, a penalized panel quantile regression model, proposed by Koenker (2004), is employed in this study. Our analysis incorporates a rich longitudinal data set (Household, Income, and Labour Dynamics in Australia (HILDA) Survey) for nine years.¹

The standard ORU (over-education, required education, and under-education) framework of earnings, as originally proposed by Duncan and Hoffman (1981) and Hartog (2000), is widely used in over-education empirical research. However, relatively few studies have been conducted on the impact of education-occupation mismatches on immigrant earnings. Applying the ORU framework, Green et al. (2007) and Kler (2007) utilized the survey of recent immigrants to Australia in the 1990s. They found that immigrants are more vulnerable to over-education in their first few years in the host country than the Australian-born population. They also found that NESB immigrants have lower returns to required and surplus education than Australian natives. The analysis employed OLS estimation. Chiswick and Miller (2008, 2010) reported that NESB immigrants have a lower rate of return to schooling,

¹ The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The HILDA data sets are made available to academics through the Australian Bureau of Statistics.

accompanied by over-education and under-education, in the US and Australia.

Tsai (2010) used a panel approach with US data and found that the earnings penalty associated with over-education diminishes significantly or disappears once individual heterogeneity is accounted for. She verified this finding by race and years of education. The analysis for immigrants was not a part of the study.

The heterogeneity of returns to over-education was considered by McGuinness and Bennett (2007), who adopted a cross-section quantile regression to study the impact of over-education on the earnings of a cohort of graduates. The study found asymmetries across the earnings distribution. The incidence and penalty for educational mismatch were the greatest at the lowest end of the earnings distribution, which was considered to reflect lower ability in the general population of graduates. Adjusting for endogeneity due to individual heterogeneity through panel data analysis was outside of the scope of the study. Bender and Roche (2018) support this finding, and conclude that the earnings penalty from educational mismatch shrinks at the upper earnings decile. In contrast, Budría (2011), and Hernández and Serrano (2012) showed the opposite evidence, indicating that the wage inequality is caused by educational mismatch, and not by unobservable characteristics such as innate ability or motivation. Furthermore, Johnes (2019) found a mixture of results of earnings penalty over the wage distribution across countries.

Billger and Lamarche (2015) applied a panel data quantile regression analysis to examine native-immigrant earnings differentials and verified heterogeneity across immigrant earnings distributions. However, the migrant earnings differential from the education-occupation mismatch perspective was outside the scope of that analysis.

In this paper, we employ nine years of longitudinal data from Australia and extend this literature.

Our study contributes to this literature as follows: First, it extends the analysis of over-education by using a penalized quantile regression approach. Penalized regression analysis allows us to examine the distributional earnings effects associated with education-occupation mismatches of immigrants by taking both distributional heterogeneity and individual/job heterogeneity into account. The distributional heterogeneity of penalties pertaining to job mismatches for immigrants is a question that is deserving of careful evidence. In addition and notably, estimates we obtain with the penalized regression method also allow us to address the endogeneity caused by the correlation between explanatory variables and error terms in the cross-sectional quantile modelling. To the best of our knowledge, this is the first application of the penalized quantile regression approach to the impact of over-education on earnings. Second, to further examine the drivers of divergent results by countries of origin of immigrants, we extend this analysis to incorporate questions relating to the transferability of overseas degrees. Developing our analysis for two sub-samples of ESB and NESB immigrants and country of the highest degree, we find divergent distributional effects for these two groups with relevant policy implications.

The remainder of this paper is organized as follows: The following section provides a brief overview of the analytical framework. Section 3 outlines the data and variables examined. The results are then presented in Section 4, followed by conclusions in Section 5.

2. Analytical framework

2.1 ORU model

We use the standard ORU (over-education, required education, and under-education) earnings modelling set up, originally proposed by Duncan and Hoffman (1981) as a benchmark model to estimate the impact of over-education on earnings.

As an extension of human capital models, the ORU model decomposes actual years of education (S_a) into required years of education (S_r), years of over-education (S_o), and years of under-education (S_u). Thus $S_a = S_r + S_o - S_u$, where $S_o = S_a - S_r$ for the over-educated (i.e., if $S_a > S_r$), and 0 otherwise. Similarly, $S_u = S_r - S_a$ for the under-educated (i.e., if $S_r > S_a$), and 0 otherwise. Using the mode method, years of required education is based on the within-sample mode years of education in the specific occupation of the job (two-digit occupation classifications). Years of over-education measures years of education held that exceed the mode level of required education in the occupation.

The model was originally applied to general populations, but it lends itself well to analyses for immigrants. In its original form, designed for mean estimations, the model takes the form of equation 1. Then the log of earnings in the ORU model can be written as:

$$\ln y_{it} = \beta_r S_{r,it} + \beta_o S_{o,it} + \beta_u S_{u,it} + \delta x_{it} + \alpha_i + \varepsilon_{it} \quad (1)$$
$$i = 1, \dots, n; t = 1, \dots, T$$

Where $\ln y_{it}$ is the natural logarithm of earnings of individual i at year t ; x_{it} is personal characteristics and other control variables; $S_{r,it}$, $S_{o,it}$, $S_{u,it}$ are, respectively, the years of required education, over-education, and under-education of individual i at year t ; α_i denotes the unobservable individual-specific effect; and $\varepsilon_{i,t}$ is an error term assumed to be independent and identically distributed i.i.d $(0, \sigma_\varepsilon^2)$. β_r , β_o , β_u are the rates of returns to required education, over-education, and under-education, respectively.

For individuals who have well-matched education levels for their jobs, the rate of return to their education is reflected by β_r . For individuals in mismatched jobs, the rate of return to their education is $(\beta_r + \beta_o)$ or $(\beta_r + \beta_u)$ across the two components of their matched and mismatched years of education. Previous studies for the general population find that $\beta_r > 0$, $\beta_o > 0$, and $\beta_u < \beta_r$, where the difference between β_o and β_r is the penalty for each year of over-education and $\beta_u < 0$. The model also accounts for the cohorts of immigration and includes states and time fixed effects.

2.2 Estimation of the model

Two estimation issues are relevant in this context. The first is the endogeneity of education-job mismatch variables. The second is the possible heterogeneity of results across the earnings distribution. We elaborate on these points below.

The econometric model we have developed, based on the Koenker (2004) penalized quantile regression approach, has the advantage of addressing both issues.

2.2.1 Endogeneity

Endogeneity can occur due to unobserved factors, resulting in a correlation between an explanatory variable and error terms. Unobserved factors associated with educational mismatches may reflect typical supply-side personal characteristics, such as motivation, effort, and diligence. In the case of immigrants, and especially for immigrants from NESB countries of birth, additional unobservable factors could result in endogeneity on the demand side, due to information asymmetries on the part of employers about the quality of degrees and potential productivity. Immigrants, especially NESB immigrants, may use higher years of education to signal greater productivity and, as a result, hold jobs at which they are over-educated. Fixed effects estimation, applied within the quantile setting, captures these factors that are constant over time in the panel setting in our analysis.

2.2.2 Asymmetry of coefficients across the earnings distribution

The assumption of homogeneity of earnings returns for each year of over-education/required education across the earnings distribution can be restrictive, in particular for NESB immigrants where greater information asymmetries also exist.

For example, at the lower portions of the earnings distribution, NESB immigrants who are over-educated (e.g., university graduates in lower occupations such as trades) may be utilizing their other knowledge and qualities to earn relatively more for each year of over-education/required education (Hartog & Oosterbeek, 1988). The high hourly rate in trades can also contribute to the relatively high return to over-education at the lower and mid-earnings quantile levels. However, at the higher ends of the earnings distribution, asymmetries in the transferability of previous educational learning, or in information about the quality of immigrants' educational backgrounds, may result in greater penalties for years of over-education.

2.3 Econometric Model

Koenker (2004) established a penalized quantile regression estimator with fixed effects by adding the ℓ_1 penalty to a quantile regression function proposed by Koenker and Bassett (1978) in order to solve heterogeneity associated with the covariates under the less stringent distributional assumption. Koenker (2004) pointed out that when the number of individuals is larger than time period T, the penalized panel quantile regression estimator is advantageous in controlling the variability stemming from a large number of estimated α parameters. The ℓ_1 penalty is selected because ℓ_1 shrinkage offers statistical and computational advantages over conventional Gaussian ℓ_2 penalties, as claimed by Tibshirani (1996) and Donoho et al. (1998).

A contribution of our analysis is that we extend the ORU model by embedding the ORU model into the penalized quantile regression model to solve,

$$\min_{(\alpha, \beta_r, \beta_o, \beta_u, \delta)} \sum_{k=1}^q \sum_{i=1}^n \sum_{t=1}^T \eta_k \sigma_{\tau_k} \left(y_{it} - \alpha_i - S_{r,it} \beta_r(\tau_k) - S_{o,it} \beta_o(\tau_k) - S_{u,it} \beta_u(\tau_k) - x_{it}^T \delta(\tau_k) \right) + \lambda \sum_{i=1}^n |\alpha_i| \quad (2)$$

$i = 1, \dots, n; t = 1, \dots, T$

where τ is the index of quantile levels, ranging from 0 to 1; α_i denotes the unobservable fixed effect of individual i ; $S_{r,it}$ is the years of required education for individual i at year t ; $S_{o,it}$ is the years of over-education; and $S_{u,it}$ is years of under-education. β_r (β_o , β_u) measures the rate of returns to required education (return to over-education, penalty to under-education). x_{it} represents the other control variables of individual i at time t . $Q_{y_{it}}(\tau|\cdot)$ is the τ -th conditional quantile function. $\beta_r(\cdot)$, $\beta_o(\cdot)$, $\beta_u(\cdot)$, and $\delta(\cdot)$ are the vectors to be estimated. $\sigma_\tau(u) = u(\tau - I(u < 0))$ is the piecewise linear quantile loss function introduced by Koenker and Bassett (1978) and η_k is the relative weight of quantile levels $\{\tau_1, \dots, \tau_q\}$ on the estimation of α_i . We use equally weighted quantiles $\eta_k = 1/q$. A weighted bootstrap method is used to compute standard errors. The tuning parameter λ expresses the magnitude of the penalty term and is set at $\lambda = 1.5$ in the analysis. An alternative value of λ is also used in our robustness check. The penalized quantile fixed effects estimator (Billger & Lamarche, 2015) allows us to estimate the impacts of over-education on earnings while addressing individual-specific factors.

3. Data and variables

3.1 Data

In this research, we use data taken from the first nine waves of the responding person file of the Household, Income, and Labour Dynamics in Australia (HILDA) Survey. The HILDA Survey began in 2001 and is conducted annually. It is Australia's first nationally representative household panel survey. We have selected the HILDA Survey because of its many advantages.

The major advantage is that this longitudinal survey overcomes the disadvantages of cross-section surveys by following the same individuals over time, thus allowing researchers to analyze labour market conditions for the same individual over extended periods. The use of appropriate panel data techniques that account for potential endogeneity and unobserved individual heterogeneity is particularly important to immigrants and job-education-mismatch studies. The second advantage of the data is that it contains rich information on education credentials with regard to their type and the country where the highest qualification was acquired, in addition to information regarding other variables of interest. The duration of the data also permits us to analyze variation in the performance of immigrants in and out of mismatched jobs.

The sample group comprised an unbalanced panel of full-time male workers aged 23–64. Within the sample group, about 11 percent of the immigrants moved into jobs for which they were over-educated at least once during the period of the study.²

² The HILDA data set used has a remarkably high response rate throughout the survey period (HILDA Survey Annual Report 2010 (<http://melbourneinstitute.com/downloads/hilda/areport2010.pdf>)). Re-interview rates are reasonably high, rising from 87 percent in wave 2 to over 96.3 percent in wave 9. The period selected for this research provides continuous data and group inclusion in the data, and it is guided by some major changes in HILDA Survey data in 2011.

The sample selection question (i.e., potential non-random selection into employment/full-time) was also examined satisfactorily before carrying out the panel quantile regression analysis.³

A limitation of the data set is that it mainly covers immigrants who were present in Australia in 2001, with small additions to the sample over the period. Therefore, the analysis does not cover most new immigrants since 2001. However, we believe that given the objectives of the study, the advantages of the data set outweigh this particular limitation.

Our sample size is composed of 15,833 observations. Of this total, 12,606 of observations are for Australian-born, 2,025 for ESB immigrants, and 1,202 for NESB immigrants. Most ESB immigrants come from developed countries such as the UK, New Zealand, South Africa, and the US. The NESB immigrants, however, represent diverse countries of birth across 60 countries, such as Vietnam, China, India, and Philippines.

Analysis of the data confirms significant variation across time and sample groups during the period of the analysis, for the efficacy of our panel data estimation. For example, one-quarter of the Australian-born sample switched from over-educated to matched or under-educated occupation status at least once during the panel data period. For ESB immigrants, the transit rate from being over-educated to a matched or under-educated occupation is 24.27 percent, and it was 16 percent of NESB immigrants. Earnings also changed significantly during the period of the panel data.

The motivation for getting a well-matched job drives NESB immigrants to study in Australia and obtain local degrees, even though they may already have overseas qualifications. As a result, Australian qualifications may be more valued among NESB than ESB immigrants. Also, the highest qualification obtained in Australia or overseas is assumed to affect economic integration. Therefore, we consider Australian and overseas qualifications separately based on the country where the highest qualification was obtained.

Table 1 presents the summary statistics for English speaking and non-English speaking immigrants in terms of their country of birth and for the Australian-born population. Table 1 shows there is significant variation in personal characteristics between the three groups. For example, the average number of years of education shows that immigrants are highly educated. Notably, 4 percent more immigrants have obtained post-school qualifications (70 percent) than that for the Australian-born population (66 percent).

Immigrants' educational 'quality' (referring to human capital skills) is expected to affect their success in obtaining employment with a better match between their actual education and the educational requirements of the job. Particularly when immigration policy changes favor skilled migration, the earnings of immigrants is expected to increase compared to the earlier cohorts. Since 1980, Australian immigration policy has emphasized skilled migration, leading to a larger proportion of NESB immigrants arriving in the decades since 1980. Therefore, NESB immigrants have, on average, fewer years since migration, and fewer years of work experience. All migrants are also more likely to have higher levels of educational attainment than earlier cohorts (Department of Parliamentary Services (DPS), 2012; Department of Immigration and Border Protection (DIBP), 2014, 2016). Therefore, in our analyses, we

³ The test results show that accounting for either type of selection (employment or full-time) did not alter the results reported in this study.

control years since migration, years of experience, and the cohort of arrival.

Notably, the average hourly wage for ESB immigrants (\$32.49) is also higher than that for Australian-born (\$29.81) and for NESB immigrants (\$29.50). The earnings differential may be explained by the transferability of overseas qualifications. ESB immigrants come from economically advanced countries, and their qualifications are valued in Australia. In contrast, NESB immigrants may experience more difficulty in adapting to their new lives, even if they work in skilled job categories. Furthermore, NESB immigrants may work in occupations that require lower levels of educational attainment when their overseas credentials are not recognized by Australian employers. This evidence encourages us to explore what leads to this earnings differential. We test the hypothesis that NESB immigrants are likely to have a greater earnings penalty in educational mismatched jobs in comparison with ESB immigrants and Australian natives, and the extent to which an Australian qualification may help NESB immigrants improve their situation. In particular, we examine whether there are asymmetries in earnings effects across the earnings distribution.

3.2 Variables

The dependent variable in our models is the natural logarithm of the hourly wage rate from the main job in constant 2009 dollars. The models include the conventional human capital variables, including measurement of years of required education and over- or under-education.

The over-education measure is adapted based on the conventional mode method (Kiker, Santos, & Mendes de Oliveira, 1997).⁴ In this method, each individual's education level is compared to the mode level of education for the occupation held (at the two-digit level of occupation categories) for each year of data. For example, an individual is over-educated if their level of education is higher than the mode.

Years of actual education are derived from four variables included in HILDA. To evaluate the effects of qualification, we categorize qualifications into five categories: postgraduate (17 years of education), bachelor (16 years), diploma (15 years), high-school completion certificate (13 years, based on 12 years of schooling and one year of kindergarten), and no qualification for less than high-school completion (based on reported completed school education level).⁵

Years since migration (YSM) measures years of residence in Australia for immigrants, and the YMS coefficients estimate immigrants' assimilation effects.

As English is the primary language spoken in Australia, NESB immigrants are more likely to decrease their expectations while job searching and are more likely to accept jobs that require education below their level of attainment. Therefore, proficiency in spoken English is expected to have a positive effect on the rate of over-education and immigrants' assimilation. This

⁴ Kiker, Santos, and Mendes de Oliveira (1997) and Verhaest and Omey (2006) show that the mode method is preferred to Verdugo and Verdugo's (1989) mean criterion. They found that the mean criterion changed gradually and that it could produce classification errors before correcting itself. In contrast, the mode changes more freely, reflecting each period's educational requirements for most workers at any given time.

⁵ We use a standardized system for accounting for years of education for high-school completion. High-school completion age across the countries in the data is the age of 18, universally requiring 12 years of schooling plus a kindergarten/preparation year at age 6.

variable identifies 74 percent of non-English speaking immigrants with well-spoken English.⁶

The arrival cohorts' quality may have a subtle effect on immigrants' economic integration (e.g., greater skills, quality of education and resources (Borjas, 1985)). Successive arrival cohorts of higher-skilled immigrants could be expected to have greater productivity in the host country, which may accelerate their earnings compared to their counterparts. The heterogeneity of average productivity among different arrival cohorts would affect the coefficient of the number of years since migration (YSM). Thus, the estimated impact of years since migration may reflect either immigrants' assimilation or changes in the quality of the cohorts. Failing to consider arrival cohorts quality would bias the estimated effect of immigrants' assimilation (Borjas, 1985). Thus, in order to investigate whether there are unobserved differences in productivity across immigrants from different arrival periods, we define three cohorts based on their years of arrival in 1947–1979, 1980–1989, and 1990–2001. We also include time and states fixed effects and work experience variables in the model.

⁶ The relevant variable in HILDA asks 'How well do you speak English?' among the population who speaks another language at home. Answer: 1-very well, 2-well, 3-not well, 4-not at all. 'Proficient in English' corresponds to category '1'.

Table1 Summary statistics by country of birth group

VARIABLES	Australian-born	ESB Immigrants ¹	NESB Immigrants ²
	Mean (sd ³)	Mean (sd)	Mean (sd)
Human Capital			
Years of experience (total)-EXP	21.36 (10.30)	24.63 (10.38)	23.25 (10.44)
Years of actual education (total)-ED	13.76 (2.40)	14.08 (2.55)	14.57 (2.52)
<i>Based on country in which achieved</i>			
Australian qualification	0.66 (0.46)	0.41 (0.49)	0.43 (0.49)
Overseas qualification	/	0.29 (0.46)	0.29 (0.46)
Personal Characteristics			
Age	41.12 (9.96)	44.70 (9.91)	43.81 (9.79)
Years since migration (YSM)	/	26.32 (13.08)	20.65 (12.65)
Year of arrival	/	1978 (13.03)	1984 (12.74)
Arrived 1947-1979	/	0.50 (0.50)	0.30 (0.46)
Arrived 1980-1989	/	0.29 (0.46)	0.30 (0.46)
Arrived 1990-2001	/	0.21 (0.41)	0.39 (0.49)
Proficient in English	/	/	0.74 (0.44)
Hourly wage of main job (AUD)	29.81 (15.75)	32.49 (18.23)	29.50 (15.24)
Individuals	1,987	317	198
Observations	12,606	2,025	1,202

Sample: HILDA Release 9 (Pooled Waves 1-9), full-time employed males.

Notes:

Standard deviations are reported in parentheses. Standard deviations for binary variables represent changes for individuals across nine years of longitudinal data.

1 ESB Immigrants (English speaking background immigrants)

2 NESB Immigrants (non-English speaking background immigrants)

3 sd (standard deviation)

Table 2 reports the extent of the educational mismatch among the Australian-born natives and ESB and NESB immigrants. This shows that immigrants are more likely to be over-educated than Australian natives. Specifically, at 42 percent, the over-education rate of NESB immigrants significantly exceeds the rates of 25 percent for Australian natives and 31 percent for ESB immigrants. The evidence presented in Table 2 reveals that mismatch is especially severe among NESB immigrants.

Table 2 Incidence and extent of over-education by country of birth group

	Australian Natives	ESB Immigrants	NESB Immigrants
VARIABLES	Mean (sd)	Mean (sd)	Mean (sd)
<u>Educational mismatch</u>			
Years of over-education	2.31 (1.74)	2.26 (1.61)	2.55 (1.76)
Years of under-education	2.99 (1.53)	3.05 (1.60)	2.83 (1.53)
Years of required education	14.39 (1.37)	14.54 (1.38)	14.84 (1.36)
Over-educated	0.25 (0.43)	0.31 (0.46)	0.42 (0.49)
Under-educated	0.36 (0.48)	0.33 (0.47)	0.28 (0.45)
Matched	0.39 (0.49)	0.36 (0.48)	0.30 (0.46)
Individuals	1,987	317	198
Observations	12,606	2,025	1,202

Sample: HILDA Release 9 (Pooled Waves 1-9), full-time employed males.

In the next section, we present the results of our analyses regarding the impact of over-education on immigrant earnings, based on our penalized quantile regression models. This approach allows us to account for asymmetries in the impacts of explanatory variables across the earnings distribution segments. The method also allows adjustments for potential heterogeneity and endogeneity.

4. Results

Following the extended ORU earnings model in equation (2), we first examine the impact on earnings by country of birth language group in Section 4.1. In Section 4.2 we focus on the impact of country of highest qualification effects for both ESB and NESB immigrants. Finally, to check the validity of our results, we carry out a robustness analysis by using alternative lambda values in Section 4.3.

4.1 Impact of over-education on earnings by country of birth group

According to prior literature, and controlling for the required years of education, over-educated workers earn less than their co-workers who are employed in the same occupation, but are employed in education-occupation matched jobs (Yeo & Maani, 2015; Chiswick & Miller, 2008, 2010; Green et al., 2007; Lindley, 2009; Wald & Fang, 2008). The extent of over-education is greater for NESB immigrants than for Australian natives (see Table 2). Therefore, we expect that the earnings return to over-education varies across Australian-born, ESB, and NESB immigrants. We further expect that over-education may have different impacts on earnings at different portions of the earnings distribution due to potentially divergent labour

market conditions.

The results of the penalized quantile regressions are reported in Figure 1 and Table 3 for the three samples of Australian-born, ESB immigrants, and NESB immigrants. We also provide conventional OLS regression results at the mean, as a benchmark for comparison. As an estimation method, OLS analysis has the drawback of omitting unobserved fixed effects, resulting in biased estimators due to the unobserved heterogeneity of individuals or jobs. The penalized quantile regression results of interest have fixed effects features, in addition to allowing differences in coefficients across portions of the earnings distribution. The earnings returns reported across quantiles are conditional on the specific group (Australian-born, ESB, and NESB immigrants).

Tests of the null hypothesis that the quantile coefficients are constant as a function of the quantile index statistically reject the null hypothesis for the over-education variable of interest across most categories. We report on these results later in this section.

First, as expected, we find that the impact of over-education and required education is smaller in the penalized quantile regression than in conventional regression, as shown in Figure 1 and Table 3. This result supports expectations that unobserved factors, such as ability, motivation, effort, or preference for other aspects of jobs, may partly explain earnings returns to education for these groups. Moreover, the return to over-education is generally smaller than the return to required education at each quantile level for these three groups, a result that is consistent with the literature (Chiswick & Miller, 2008, 2010; Tsai, 2010; Green et al., 2007; Kler, 2007; Voon & Miller, 2005; Hartog, 2000; Groot, 1996; Cohn, 1992; Sicherman, 1991). We also find other results as expected based on human capital explanations, such as the positive impact of years of experience on earnings at all quantiles.

However, the results add to earlier findings by further illustrating larger variations in returns to over-education and required education across different quantile levels for immigrants compared to the Australian-born. For NESB immigrants, in particular, a significantly lower return to both over-education and required education is observed at each quantile level, compared to ESB immigrants and Australian natives, indicating the earnings disadvantage in returns to education.

An interesting finding revealed by the quantile results is that for the Australian-born group in the 90th percentile, a penalty for over-education is not verified. In fact, the results show a small gain for each year of over-education relative to returns to required education.

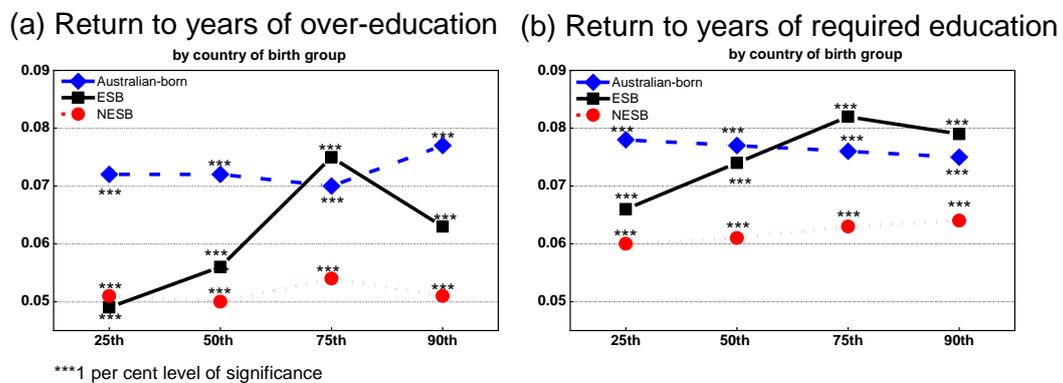
Another finding is regarding the differential impact of years since migration, which is generally considered as an indicator of the economic integration of immigrants in the host country over time. We find that, for NESB immigrants, the effect becomes smaller in the higher quantiles (0.038 at the 25th and 0.025 at the 90th quantiles). A similar decrease is not observed for ESB immigrants. These results indicate that added barriers might slow earnings for NESB immigrants, despite the passage of time since migration. Examples of such obstacles are the transferability of overseas institutional knowledge and educational qualifications, or less-than-full language proficiency beyond what the self-reported variable controls for, or other demand-side factors related to perceptions of productivity.

Moreover, divergent country-of-qualification earnings effects are found between ESB and NESB immigrants. These differences are stark, with positive effects for ESB and negative

effects for NESB immigrants. For example, for NESB immigrants, an overseas qualification impacts earnings by -0.088 at the 50th quantile level. These results point to the impacts of the transferability of overseas degrees for NESB immigrants with different impacts across jobs. In addition, differences by quantile show further insights in this regard. We examine these differences further in Section 4.3, with additional sub-sample break downs by the country-of-qualification group.

Finally, the results confirm the significant impact of proficiency in English on earnings for NESB immigrants across all earnings levels. Specifically, there are greater impacts at the higher portions of the earnings distribution (0.084 at the 25th quantile to 0.107 at the 90th quantile), indicating the importance of language proficiency for this group.

Figure 1 Earnings returns by quantile for country-of-birth groups



Note: The estimates shown are derived from both penalized regression at the 25th, 50th, 75th, and 90th quantiles based on the results shown in Table 3.

Source: Authors' computations, HILDA Release 9 (Pooled Waves 1-9), full-time employed males.

Table 3 Earnings returns: Panel quantile regression model results:
Coefficients (standard errors) ϕ

Variables	Panel quantile regressions				Conventional regression γ
	25th	50th	75 th	90th	
Australian-born					
Years of over-education	0.072*** (0.004)	0.072*** (0.003)	0.070*** (0.003)	0.077*** (0.005)	0.086*** (0.004)
Years of under-education	-0.052*** (0.003)	-0.055*** (0.002)	-0.055*** (0.003)	-0.050*** (0.003)	-0.038*** (0.003)
Years of required education	0.078*** (0.003)	0.077*** (0.002)	0.076*** (0.002)	0.075*** (0.003)	0.107*** (0.003)
EXP	0.027*** (0.002)	0.024*** (0.002)	0.020*** (0.002)	0.019*** (0.002)	0.023*** (0.002)
EXP ² /100	-0.050*** (0.004)	-0.042*** (0.003)	-0.036*** (0.004)	-0.034*** (0.004)	-0.045*** (0.003)
ESB immigrants					
Years of over-education	0.049*** (0.010)	0.056*** (0.007)	0.075*** (0.011)	0.063*** (0.012)	0.081*** (0.010)
Years of under-education	-0.060*** (0.008)	-0.058*** (0.007)	-0.062*** (0.009)	-0.058*** (0.011)	-0.033*** (0.008)
Years of required education	0.066*** (0.007)	0.074*** (0.006)	0.082*** (0.007)	0.079*** (0.009)	0.106*** (0.007)
EXP	0.023*** (0.005)	0.020*** (0.004)	0.017*** (0.005)	0.024*** (0.006)	0.010** (0.005)
EXP ² /100	-0.037*** (0.010)	-0.037*** (0.008)	-0.032*** (0.009)	-0.048*** (0.012)	-0.012 (0.009)
Years since migration (YSM)	0.012* (0.007)	0.011** (0.006)	0.014** (0.006)	0.009 (0.013)	-0.018*** (0.006)
YSM ² /100	-0.019** (0.009)	-0.016** (0.008)	-0.018** (0.008)	-0.013 (0.011)	0.009 (0.008)
Overseas qualification	0.055 (0.038)	0.082*** (0.031)	0.117*** (0.037)	0.109** (0.047)	-0.004 (0.031)
NESB immigrants					
Years of over-education	0.051*** (0.010)	0.050*** (0.009)	0.054*** (0.010)	0.051*** (0.014)	0.084*** (0.011)
Years of under-education	-0.045*** (0.012)	-0.043*** (0.010)	-0.039*** (0.011)	-0.031*** (0.011)	-0.022** (0.011)
Years of required education	0.060*** (0.010)	0.061*** (0.008)	0.063*** (0.009)	0.064*** (0.010)	0.100*** (0.008)
EXP	0.013** (0.006)	0.009* (0.005)	0.001 (0.005)	0.004 (0.006)	0.017*** (0.005)
EXP ² /100	-0.031** (0.014)	-0.018 (0.011)	-0.004 (0.012)	-0.009 (0.013)	-0.041*** (0.011)
Years since migration (YSM)	0.038*** (0.009)	0.034*** (0.007)	0.031*** (0.007)	0.025** (0.011)	0.015** (0.007)
YSM ² /100	-0.047*** (0.014)	-0.043*** (0.012)	-0.039*** (0.012)	-0.038** (0.017)	-0.021** (0.010)
Overseas qualification	-0.076 (0.048)	-0.088* (0.046)	-0.059 (0.045)	-0.050 (0.059)	-0.070** (0.034)
Proficient in English	0.084*** (0.030)	0.091*** (0.031)	0.085*** (0.031)	0.107** (0.048)	0.135*** (0.032)

Notes: ϕ lambda=1.5

γ OLS, conventional regression at the mean.

***1 per cent level of significance; **5 per cent level of significance, *10 per cent level of significance

The model also includes cohorts fixed effects, states fixed effects, and time periods.

Sources: HILDA-Release 9 (Wave 1-Wave 9), full-time employed males.

4.1.1 Test of heterogeneity of quantile coefficients

As recommended by Koenker and Bassett (1982), Wald tests are used to check for equality of slopes across quantiles for groups by country of birth and immigrant sub-groups by country of highest qualification obtained. Table 4 reports the test results for those variables of interest. The Wald test results reject the null hypothesis that coefficients on years of over-education are equal across most quantile levels. For example, the Wald statistic in the first row and column 3 in Table 4 shows that the coefficient of years of over-education at the 90th quantile level is significantly different from that at the 50th quantile. Similar test results are also found for coefficients on years of required education, emphasizing the necessity of employing quantile regression analysis instead of OLS or FE models. The Wald tests confirm the statistically significant differences in estimated returns.

Table 4 Wald test results for equality of slopes across quantiles by country of birth group¹

Variables	(1) 25 th -50 th	(2) 75 th -50 th	(3) 90 th -50 th
Australian born			
Years of over-education	0.9935	1.9713	18.2576 ***
Years of under-education	12.6425 ***	0.3094	0.1790
Years of required education	0.0028	0.6933	16.7595 ***
ESB immigrants			
Years of over-education	0.6556	8.4571***	8.0984 ***
Years of under-education	8.7608 ***	0.7787	0.0165
Years of required education	5.2761**	5.0879 **	4.4237 **
NESB immigrants			
Years of over-education	0.1152	3.1557 *	1.8876
Years of under-education	3.0467 *	3.1806 *	4.6769 **
Years of required education	4.6060 **	4.1644 **	7.4745 ***

Notes:

Wald test results correspond to the equality of the coefficients in Table 3.

The null hypothesis is that the quantile coefficients are constant. A statistically significant result rejects the null.

***1 per cent level of significance, **5 per cent level of significance, *10 per cent level of significance

4.2 Country of highest qualification and over-education effects on earnings

As shown in the previous section, overseas qualifications have the opposite effects on earnings for ESB and NESB immigrants. In this section, we further analyze the impact of over-education on earnings for ESB and NESB sub-samples, based on the country where the highest qualification was obtained.

The results of penalized quantile regressions for ESB and NESB immigrant samples are summarized in Figure 2 and Tables 5 and 6.⁷ Figure 2 presents a summary of earnings effects for each year of over-education and required education for the sub-samples with the overseas or domestic country of highest qualifications. Results are reported for ESB and NESB immigrants.

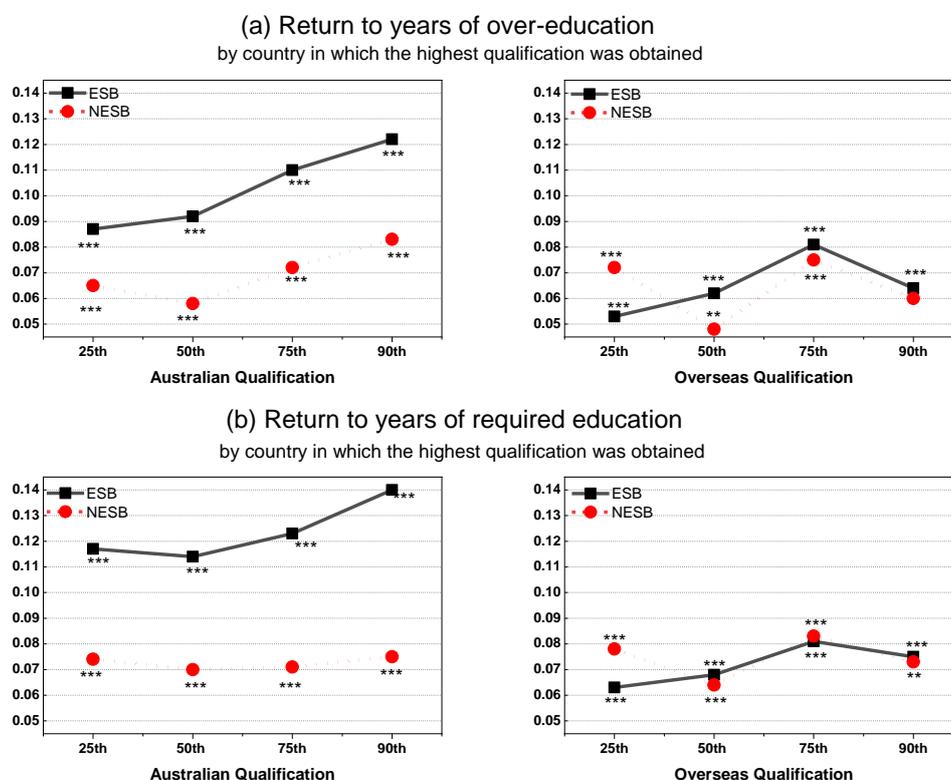
⁷ Tests results of heterogeneity of quantile coefficients in Table 5 are reported in Panel A of Appendix Table A1.

A result that stands out in Figure 2 is the significant premium from an Australian highest degree for both ESB and NESB immigrants. This effect is noteworthy in particular for ESB immigrants, but also for NESB immigrants at the highest portions of the earnings distribution.

4.2.1 Results for ESB immigrants

As Figure 2 and Table 5 show, ESB immigrants with an Australian qualification earn significantly more than those with an overseas qualification for each year of over-education or required education across all quantile levels. An additional year of over-education increases earnings by 0.087 at the 25th quantile to 0.122 at the 90th quantile level. Each extra year of required education raises earnings by 0.117 at the 25th quantile to 0.140 at the 90th quantile for ESB immigrants who obtained an Australian qualification.

Figure 2 Earnings returns per year by country-of-highest-degree (by quantile)



***1 per cent level of significance; **5 per cent level of significance.

Note: Estimates are obtained from penalized regression at the 25th, 50th, 75th, and 90th quantiles, as reported in Tables 5 and 6.

Source: Authors' computations, HILDA Release 9 (Pooled Waves 1-9), full-time employed males.

These results further confirm that over-educated ESB immigrants earn less than their colleagues who are in jobs for which they are adequately educated. In other words, each year of over-education has positive but smaller earnings returns compared to each year of adequate education.

Comparing the results in Figure 2 to those in Figure 1 shows that the returns to over-education and required education for ESB immigrants who hold an Australian qualification are higher

than that for Australian-born employees. These returns are higher than those at each corresponding quantile level. In contrast, ESB immigrants with an overseas qualification have a lower earnings return for each year of both over-education and required education at each corresponding quantile level (25th, 50th, and 90th) than Australian-born employees.

Table 5 Earnings returns for ESB immigrants:
Results by country of obtaining the highest degree:
Coefficients (standard errors)

Variables	Panel quantile regressions			
	25th	50th	75th	90 th
– Australian Qualification				
Years of over-education	0.087*** (0.016)	0.092*** (0.015)	0.110*** (0.015)	0.122*** (0.019)
Years of under-education	-0.118*** (0.028)	-0.121*** (0.030)	-0.110*** (0.033)	-0.129*** (0.060)
Years of required education	0.117*** (0.016)	0.114*** (0.014)	0.123*** (0.014)	0.140*** (0.018)
EXP	0.047*** (0.013)	0.036*** (0.010)	0.031*** (0.009)	0.030*** (0.010)
EXP ² /100	-0.086*** (0.029)	-0.065*** (0.021)	-0.055*** (0.019)	-0.044*** (0.021)
Years since migration (YSM)	0.000 (0.011)	-0.004 (0.012)	-0.007 (0.013)	-0.015 (0.015)
YSM ² /100	0.002 (0.016)	0.009 (0.016)	0.011 (0.018)	0.013 (0.020)
– Overseas Qualification				
Years of over-education	0.053*** (0.015)	0.062*** (0.009)	0.081*** (0.013)	0.064*** (0.014)
Years of under-education	-0.057*** (0.009)	-0.052*** (0.007)	-0.055*** (0.008)	-0.052*** (0.011)
Years of required education	0.063*** (0.009)	0.068*** (0.007)	0.081*** (0.008)	0.075*** (0.010)
EXP	0.027*** (0.008)	0.025*** (0.006)	0.020** (0.007)	0.031*** (0.008)
EXP ² /100	-0.040*** (0.015)	-0.040*** (0.011)	-0.033** (0.013)	-0.054*** (0.015)
Years since migration (YSM)	0.010 (0.007)	0.012** (0.006)	0.017** (0.007)	0.011 (0.010)
YSM ² /100	-0.009 (0.012)	-0.008 (0.010)	-0.012 (0.011)	-0.001 (0.015)

Notes: lambda=1.5

***1 per cent level of significance, **5 per cent level of significance, *10 per cent level of significance

The model also includes cohorts fixed effects, states fixed effects, and time periods.

Sample: HILDA-Release 9 (Wave 1-Wave 9), full-time employed males.

4.2.2 Results for NESB immigrants

The results in Table 3 showed that NESB immigrants are the most vulnerable group, with a lower return to years of over-education and required education compared to ESB immigrants and Australian-born employees.

Figure 2 and Table 6 show that an Australian qualification helps NESB immigrants increase their earnings more than an overseas qualification at the 50th and 90th quantiles. Also, more variation across income levels is found for NESB immigrants with a foreign qualification than for those with Australian qualifications. The distributional and heterogeneous effects of over-

education and required education, years since migration, work experience, and language proficiency across different quantiles are shown in Table 6.⁸

Notably, disaggregation by quantile reveals that at the highest earnings quantile, for the group of NESB immigrants with Australian degrees, over-education receives an earnings premium, compared to years of required education for the job. This finding adds to conventional findings that show mainly earnings penalties based on regressions at the mean. At the 90th earnings quantile, the return to each year of over-education is associated with a gain of 11 percent in earnings, compared to the return to a required year of education. (The coefficients for the over-education and required education effects at the 90th earnings quantile are 0.083 and 0.075, respectively, and the effect is statistically significant). This result points to the positive long-term effect of university studies among NESB immigrants in the host country, despite previous overseas qualification.

However, with an overseas qualification for NESB immigrants whose income is at the 50th quantile level, relatively lower returns to over-education (0.048) and required education (0.064) are found, compared to those at other quantile levels. This earnings penalty for each year of over-education is equivalent to a penalty of 25 percent from educational mismatched employment.

Combining our results in Sections 4.2.1 and 4.2.2, immigrants with less than proficient command of the language of their host country can benefit from an additional degree obtained in the host country, even though they may already have an overseas higher degree. As such, our results indicate that for NESB immigrants with an overseas degree, investing in a qualification in the host country is associated with earnings gains with jobs at the higher sections of the earnings distribution, pointing to productivity gains. These results, based on the penalized quantile panel estimation method that adjusts for unobservable personal characteristics, provide further confidence in these findings.

Finally, language proficiency significantly affects NESB immigrants' earnings by about 0.147 (15.8 percent)⁹ at the 50th quantile, to 0.164 (17.8 percent) at the 90th quantile.

⁸ Tests results of heterogeneity of quantile coefficients in Table 6 are reported in panel B of Appendix Table A1.

⁹ Using the conventional formula [$\text{antilog } \beta - 1$], to derive the percentage effect for binary variables in semi-log regressions.

Table 6 Earnings returns for NESB immigrants:
Results by country of obtaining the highest degree:
Coefficients (standard errors)

Variables	Panel quantile regressions			
	25th	50th	75th	90th
– Australian Qualification				
Years of over-education	0.065*** (0.015)	0.058*** (0.011)	0.072*** (0.012)	0.083*** (0.016)
Years of under-education	-0.052*** (0.013)	-0.053*** (0.011)	-0.044*** (0.011)	-0.025* (0.012)
Years of required education	0.074*** (0.014)	0.070*** (0.010)	0.071*** (0.011)	0.075*** (0.011)
EXP	0.018** (0.006)	0.011** (0.006)	0.003 (0.006)	0.004 (0.006)
EXP ² /100	-0.040** (0.015)	-0.023* (0.013)	-0.011 (0.013)	-0.015 (0.014)
Years since migration (YSM)	0.057*** (0.011)	0.048*** (0.009)	0.047*** (0.009)	0.038*** (0.013)
YSM ² /100	-0.074*** (0.016)	-0.065*** (0.013)	-0.062*** (0.014)	-0.053*** (0.020)
Proficient in English	0.047 (0.040)	0.041 (0.037)	0.054 (0.038)	0.009 (0.060)
– Overseas Qualification				
Years of over-education	0.072*** (0.020)	0.048** (0.022)	0.075*** (0.026)	0.060 (0.037)
Years of under-education	0.025 (0.091)	0.049 (0.065)	0.024 (0.093)	0.102 (0.142)
Years of required education	0.078*** (0.020)	0.064*** (0.021)	0.083*** (0.026)	0.073** (0.038)
EXP	0.021 (0.020)	0.015 (0.017)	0.008 (0.015)	0.026 (0.022)
EXP ² /100	-0.062 (0.048)	-0.036 (0.039)	-0.016 (0.038)	-0.056 (0.051)
Years since migration (YSM)	0.004 (0.016)	0.012 (0.014)	0.009 (0.014)	-0.002 (0.021)
YSM ² /100	0.055 (0.054)	0.023 (0.046)	0.010 (0.047)	0.033 (0.071)
Proficient in English	0.065 (0.058)	0.147*** (0.046)	0.163*** (0.057)	0.164** (0.080)

Notes: lambda=1.5

***1 per cent level of significance, **5 per cent level of significance, *10 per cent level of significance

The model also includes cohorts fixed effects, states fixed effects, and time periods.

Sample: HILDA-Release 9 (Wave 1-Wave 9), full-time employed males.

4.3 Robustness analysis

In order to check the validity of our results in relation to the tuning parameter λ used, a robustness analysis was carried out by using alternative values of the tuning parameter λ in the penalized regression models, as shown in equation (2). These results are reported in Table 7.

We find that both the sign and magnitude of the coefficients of the explanatory variables at each quantile level in Table 7, with an alternative λ of 1.8, are consistent with those obtained in Table 4 with a λ of 1.5. For example, some of the estimated coefficients for job-mismatched variables show only minimal change in magnitude, while others remain the same as those shown in Table 4. Similarly, the effects of language proficiency remain consistent with those shown in Table 4. These results show that this range of the penalty term addresses the unobserved heterogeneity appropriately. Based on the above analysis, the findings from the penalized panel quantile regression model can be considered to be robust.

5. Conclusion and policy implication

In this paper, we have examined the impact of over-education on immigrants' earnings by applying a penalized panel regression model. This approach addresses both individual/work heterogeneity and distributional heterogeneity. We have examined the effects of job mismatches for immigrants, as well as for Australian-born full-time employed males. We have also assessed the robustness of our findings by applying an alternative value of the tuning parameter λ .

The application of the penalized panel regression model has allowed us to differentiate between the effects of over-education on immigrant earnings at different parts of the earnings distribution. The results that adjust for endogeneity due to unobservable personal characteristics using the fixed effects features of the method also provide further confidence in these findings.

The analysis adds new results to conventional research findings based on regressions at the mean. Notably, with disaggregation by quantile, our results reveal that, for the highest earnings quantile, an earnings penalty from over-education for this group of NESB immigrants with Australian degrees is not substantiated. This group experiences a gain in earnings for over-education, compared to years of required education. This result is noteworthy for international job mismatch research by pointing to outcomes that differ from those our models would indicate based on analyses at the mean, when distributional heterogeneity is accounted for. These results can also explain why the uptake of university studies in Australia is prevalent among NESB immigrants, regardless of previous overseas qualifications. This has implications for relevant immigration policy options.

We further find noteworthy variations in returns to over-education and required education across different quantile earnings levels for immigrants in comparison to the Australian-born employees.

Divergent results across the earnings distribution are found for ESB and NESB immigrants in terms of returns to both over-education and required education. ESB immigrants perform well in the job market and earn more, comparable to Australian-born employees. NESB immigrants,

in comparison, have a significantly lower return to over-education and required education, relative to ESB immigrants and those who are Australian-born. This is despite controlling for educational qualifications, potential work experience, years since migration, cohort effects, and language proficiency.

Immigrants who have obtained an Australian highest qualification earn more than those with overseas qualifications. This is true for both ESB and NESB immigrants. Notably, our results highlight that for NESB immigrants, obtaining a qualification in Australia is an important factor that contributes to earnings gains at the highest quantiles, even though accompanied by over-education. Also, significant earnings effects, in relation to years since migration, can be noted for NESB immigrants who hold Australian qualifications and who are proficient in English.

NESB immigrants with an overseas qualification, however, show a substantial earnings penalty resulting from educationally mismatched employment at every level of the earnings distribution spectrum, compared to their co-workers who are in well-matched jobs. This penalty is at its highest, at 25 percent, at the 50th quantile.

For ESB immigrants who hold an Australian qualification and are employed in educationally matched or over-educated employment, earnings returns are notably greater than Australian-born employees at each corresponding quantile earnings level.

Language proficiency is an important factor that significantly affects NESB immigrants' earnings. English proficiency increases earnings by about 15.8 percent at the 50th quantile, to 17.8 percent) at the 90th quantile. This variable signifies the language fluency at the highest level in the survey. The positive influence of language proficiency increases successively at the higher quantiles of the earnings distribution, reflecting the importance of language fluency in related occupations.

Our findings have implications for immigration policies intended not only to attract skilled immigrants, but also to facilitate finding suitable educationally matched positions for those immigrants. The analytical framework we applied can assist successful policy design aimed at achieving the economic integration of immigrants by considering impacts on specific income groups, and by focusing on instruments that can provide the greatest potential. Our results further point to taking the interplay of country of qualification and language proficiency into consideration in skilled migration policy design.

Table 7 Robustness Test
Earnings returns: Panel quantile regression model results
Alternative lambda value=1.8
Coefficients (standard errors)

Variables	Panel quantile regressions				OLS regression
	25th	50th	75th	90th	
Australian					
Years of over-education	0.074*** (0.004)	0.073*** (0.003)	0.071*** (0.003)	0.078*** (0.004)	0.086*** (0.004)
Years of under-education	-0.052*** (0.003)	-0.054*** (0.002)	-0.055*** (0.003)	-0.050*** (0.003)	-0.038*** (0.003)
Years of required education	0.081*** (0.003)	0.080*** (0.002)	0.078*** (0.002)	0.076*** (0.003)	0.107*** (0.003)
ESB immigrants					
Years of over-education	0.051*** (0.010)	0.058*** (0.007)	0.076*** (0.010)	0.068*** (0.011)	0.081*** (0.010)
Years of under-education	-0.060*** (0.009)	-0.058*** (0.007)	-0.063*** (0.008)	-0.060*** (0.011)	-0.033*** (0.008)
Years of required education	0.070*** (0.008)	0.077*** (0.006)	0.085*** (0.007)	0.082*** (0.009)	0.106*** (0.007)
NESB immigrants					
Years of over-education	0.056*** (0.011)	0.053*** (0.009)	0.056*** (0.011)	0.057*** (0.015)	0.084*** (0.011)
Years of under-education	-0.043*** (0.013)	-0.044*** (0.011)	-0.041*** (0.012)	-0.030** (0.011)	-0.022** (0.011)
Years of required education	0.064*** (0.011)	0.062*** (0.009)	0.066*** (0.010)	0.069*** (0.011)	0.100*** (0.008)

Notes:

Robustness test –Different value of lambda

***1 per cent level of significance, **5 per cent level of significance, *10 per cent level of significance

The model also includes all variables as in Table 3.

Sample: HILDA-Release 9 (Wave 1-Wave 9), full-time employed males.

References

Aringa, C. D., and Pagani, L. (2010) Labour market assimilation and over education: The case of immigrant workers in Italy, *Universita degli Studi di Milano-Bicocca. Dipartimento di Economia Politica*.

Basilio, L., Bauer, T.K., and Kramer, A. (2017) Transferability of human capital and immigrant assimilation: An analysis for Germany. *Labour* 31:245-264.

Bender, K. A., and Roche, K. (2018) Educational mismatch and the earnings distribution. *Southern Economic Journal*. 85(2): 441-456.

Billger, S. M., and Lamarche, C. (2015). A panel data quantile regression analysis of the immigrant earnings distribution in the United Kingdom and United States. *Empirical Economics* 49: 705-750.

Borjas, G. J. (1985) Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics* 3: 463-489.

Budria, S., and Martinez de Ibarreta Zorita, C. (2020) Educational and skills mismatches among immigrants: The impact of host language proficiency. *IZA Discussion Paper No. 13030*. Available at SSRN: <https://ssrn.com/abstract=3558328>

Carroll, D., and Tani, M. (2013) Overeducation of recent higher education graduates: New Australian panel evidence. *Economics of Education Review* 32: 207-218.

Chiswick, B.R., and Miller, P.W. (2008) Why is the payoff to schooling smaller for immigrants? *Labour Economics* 15(6): 1317-1340.

Chiswick, B.R., and Miller, P.W. (2010) The effect of educational-occupational mismatch on immigrant earnings in Australia, with international comparisons. *International Migration Review* 44(4): 869-898.

Cohn, E. (1992) The impact of surplus schooling on earnings. *Journal of Human Resources* 27(4): 679-682.

Department of Immigration and Border Protection (DIBP), (2014) *Continuous survey of Australia's migrants: Cohort 1 report*. August 2014, Canberra.

Department of Immigration and Border Protection (DIBP) (2016) *Historical migration statistics*. September, Canberra.

Department of Parliamentary Services (DPS) (2012) *Skilled migration: temporary and permanent flows to Australia*, September, Canberra.

Dolton, P. and Vignoles, A. (2000) The incidence and the effects of overeducation in the U.K. graduate labor market. *Economics of Education Review* 19:179-198.

Donoho, D., Chen, S., and Saunderson, M. (1998) Atomic decomposition by basis pursuit, *SIAM J. Sci. Computing* 20, 33-61.

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

Fernández, C., and Ortega, C. (2008) Labor market assimilation of immigrants in Spain: Employment at the expense of bad job-matches? *Spanish Economic Review* 10(2): 83-107.

Frenette, M. (2004) The overqualified Canadian graduate: The role of the academic program in the incidence, persistence, and economic returns to overqualification. *Economics of Education Review* 23(1): 29-45.

Friedberg, R. M. (2000) You can't take it with you? Immigrant assimilation and the portability of human capital. *Journal of Labor Economics* 18(2): 221-251.

Green, C., Kler, P. and Leeves, G. (2007) Immigrant overeducation: Evidence from recent arrivals to Australia. *Economics of Education Review* 26(4): 420-432.

Groot, W. (1996) The incidence of, and returns to overeducation in the UK. *Applied Economics* 28(10): 1345-1350.

Hartog, J. (2000) Overeducation and earnings: Where are we, where should we go? *Economics of Education Review* 19(2): 131-147.

Hartog, J. and Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: Overschooling? *Economics of Education Review* 7(2): 185-194.

Hernandez, L. and Serrano, L. (2012). Overeducation and its effects on wages: A closer look at the Spanish regions. *Investigaciones Regionales—Journal of Regional Research* 24: 59-90.

Johnes, G. (2019). The incidence of and returns to 'overeducation': PIAAC Evidence on the G7. *Minerva*. 57(1): 85-107.

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

Kler, P. (2007) A panel data investigation into over-education among tertiary educated Australian immigrants. *Journal of Economic Studies* 34(3): 179-193.

Koenker, R., and Basset, G. (1978) Regression quantiles. *Econometrica* 46:33-50.

Koenker, R., and Bassett Jr, G. (1982) Robust tests for heteroscedasticity based on regression quantiles. *Econometrica: Journal of the Econometric Society*, 43-61.

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

Leuven, E., and Oosterbeek, H. (2011) Overeducation and mismatch in the labor market. *IZA Discussion Paper No. 5523*.

Lindley, J. (2009) The overeducation of UK immigrants and minority ethnic groups: Evidence from the Labour Force Survey. *Economics of Education Review*. 28(1): 80-89.

Mavromaras, K., Sloane, P. and Wei, Z. (2012) The role of education pathways in the relationship between job mismatch, wages and job satisfaction: A panel estimation approach. *Education Economics* 20(3): 303-321.

McGuinness, S. and Bennett, J. (2007) Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review* 26(5): 521-531.

Organisation for Economic Co-operation and Development (OECD) (2013) https://www.oecd-ilibrary.org/social-issues-migration-health/foreign-born-population/indicator/english_5a368e1b-en

Quinn, M. A. and Rubb, S. (2006) Mexico's labor market: The impact of education-occupation matching on wages and productivity in developing economies. *Economics of Education Review* 25(2): 147-156.

Rubb, S. (2003) Post-college schooling, overeducation, and hourly earnings in the United States. *Education Economics* 11(1): 53-72.

Rumberger, R. W. (1987) The impact of surplus schooling on productivity and earnings. *Journal of Human Resources* 22(1): 24-50.

Sarkar, D., and Collier, T.C. (2019). Does host-country education mitigate immigrant inefficiency? Evidence from earnings of Australian university graduates. *Empirical Economics* 56: 81-106.

Sicherman, N. (1991) Overeducation" in the labor market. *Journal of Labor Economics* 9(2), 101-122.

Sloane, P. (2014) Overeducation, skill mismatches, and labor market outcomes for college graduates. *IZA World of Labor* (2014:88).

Tibshirani R. (1996) Regression shrinkage and selection via the lasso, J.Roy. *Statist. Soc.B* 58 (1996) 267-288.

Tsai, Y. (2010) Returns to overeducation: A longitudinal analysis of the U.S. labor market. *Economics of Education Review* 29(4): 606-617.

Tsang, M. C., and Levin, H. (1985) The economics of overeducation. *Economics of Education Review* 4(2): 93-104.

United Nations (2013) Department of Economic and Social Affairs, Population Division, World Migration Figures. <http://www.oecd.org/migration/mig/World-Migration-in-Figures.pdf>

Verdugo, R. R., and Verdugo, N.T. (1989) The impact of surplus schooling on earnings. *Journal of Human Resources* 24(4): 629-643.

Verhaest, D., and Omeij, E. (2006) Discriminating between alternative measures of overeducation. *Applied Economics* 38(18): 2113-2120.

Voon, D., and Miller, P.W. (2005). Undereducation and overeducation in the Australian labour market. *Economic Record* 81, S22-S33.

Wald, S., and Fang, T. (2008) Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey. *Canadian Public Policy* 34(4): 457-479.

Wen, L., and Maani, S. A. (2019). Job Mismatches and Career Mobility. *Applied Economics* 51(10), 1010-1024. 10.1080/00036846.2018.1524569.

Yeo, J. Z., and Maani, S.A. (2015) Educational mismatches and earnings in the New Zealand labour market. *New Zealand Economic Papers* 51: 28-48.

Appendix

Table A1 Wald test results for the equality of slopes across quantiles:
ESB and NESB country of highest degree

	(1)	(2)	(3)
Panel A			
Variables	25 th -50 th	75 th -50 th	90 th -50 th
ESB			
– Australian Qualification¹			
Years of over-education	0.9814	5.2857 **	6.8083 ***
Years of under-education	0.0064	0.4982	0.2124
Years of required education	3.5700 *	4.0452 **	4.7897 **
ESB immigrants			
– Overseas Qualification¹			
Years of over-education	3.1046 *	5.1790 **	0.9661
Years of under-education	17.5668 ***	2.9901 *	0.0031
Years of required education	5.0292 **	7.9313 ***	6.1362 **
Panel B			
Variables	25 th -50 th	75 th -50 th	90 th -50 th
NESB immigrants			
– Australian Qualification²			
Years of over-education	0.0075	7.6368 ***	10.9094 ***
Years of under-education	0.1504	7.7442 ***	8.8170 ***
Years of required education	2.6761	0.8922	3.0366 *
NESB immigrants			
– Overseas Qualification²			
Years of over-education	7.5670 ***	7.2899 ***	0.8886
Years of under-education	1.4438	1.0891	5.5990 **
Years of required education	6.0001 **	7.0877 ***	5.5686 **

Notes:

1. Tests correspond to quantile coefficients in Table 5

2. Tests correspond to quantile coefficients in Table 6

Null hypothesis is that the quantile coefficients are constant. A statistically significant result rejects the null.

***1 per cent level of significance; **5 per cent level of significance, *10 per cent level of significance.