I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11216

A Panel Study of Immigrants' Overeducation and Earnings in Australia

Le Wen
Sholeh A. Maani

DECEMBER 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11216

A Panel Study of Immigrants' Overeducation and Earnings in Australia

Le Wen
University of Auckland

Sholeh A. Maani
University of Auckland and IZA

DECEMBER 2017

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

A Panel Study of Immigrants' Overeducation and Earnings in Australia*

The recent literature on overeducation has provided divergent results on whether or not overeducation bears an earnings penalty. In addition, few studies have considered overeducation among immigrants. This paper uses panel data analyses to investigate the match between education and occupation and resulting earnings effects for immigrants from English Speaking, and Non-English Speaking, Backgrounds relative to the native-born population in Australia. Based on nine years of longitudinal data, the panel approach addresses individual heterogeneity effects (motivation, ability, and compensating differentials) that are crucial in overeducation analysis. First, we find that immigrants have significantly higher incidence rates of overeducation than the native-born. This probability increases, rather than diminishes, once we control for unobserved correlated effects. Second, based on panel fixed effects analyses there is no penalty for overeducation for ESB immigrants. However, NESB immigrants receive a lower return to required and overeducation compared to the other groups after controlling for individual heterogeneity.

JEL Classification: J24, J15, J31

Keywords: over-education, educational mismatch, immigrants, non-English-speaking, panel data, wage effects

Corresponding author:
Sholeh A. Maani
Graduate School of Management
The University of Auckland
12 Grafton Road,
Auckland, 1010
New Zealand

Email: s.maani@auckland.ac.nz

* An extended version of this paper is forthcoming in *International Migration*, Wiley.

1. INTRODUCTION

It is not uncommon to hear of immigrants being employed in occupations that are below the level of their educational attainment, such as those from professional occupations driving taxis or working in a kitchen. Are these isolated occurrences of overeducation which have attracted exaggerated interest, or are immigrants more likely to experience education and occupation mismatches compared to native-born populations with similar qualifications? If so, to what extent does overeducation explain immigrant earnings, and in particular returns to education among English-speaking background (ESB) and Non-English-speaking background (NESB) immigrants relative to native-born workers? In addition, do these labour market outcomes change with years since migration for immigrants?

These questions have received recent attention among researchers and policymakers due to the greater role that skilled migrants now play in developed countries. As an increasing number of immigrant-receiving countries (e.g. Canada, New Zealand, and Australia) have adopted skilled migration policies, and other countries are considering such policies, these questions continue to demand current in-depth and recent research across international settings to better understand their parameters for policy fine tuning.[1] In addition, the link between such mismatches and economic productivity is further recognized (e.g. Quinn & Rubb, 2006; Tsang & Levin, 1985).

Overeducation refers to the incidence of, or the extent to which, actual education exceeds the educational requirement to perform a job. There is evidence that immigrants suffer a higher incidence of overeducation (Aringa & Pagani, 2010; Fernández & Ortega, 2008; Kler, 2007), and that immigrants experience an earnings loss from education-occupation mismatches (Chiswick & Miller, 2008, 2010; Green et al., 2007; Lindley, 2009; Wald & Fang, 2008; Yeo & Maani, 2015). Aringa and Pagani (2010) further found that, for immigrants, work experience acquired in the host country (Italy) did not help to improve their occupation-education match.

However, very few studies in the overeducation literature have focused on overeducation with panel data for immigrants. Notably, the findings of the pioneering studies for overeducation among immigrants are generally based on cross-sectional or pooled OLS analyses (see e.g. Leuven and Oosterbeek (2011), and Sloane (2014) for a discussion of methods). The few emerging studies that have applied panel data techniques to the study of overeducation, have, in general, found controversial results on whether or not education and occupational mismatches with earnings penalties effects are verified once individual heterogeneity is accounted for (e.g. Tsai (2010), that finds no effect, versus Dolton and Vignoles (2000), Leuven

and Oosterbeek (2011) and Mavromaras et al. (2012), that indicate a reduced but significant effect, and Carroll and Tani (2013), who find wage penalties for overeducated older workers).

These studies have generally focused on overall populations combining native-born populations with immigrants or Non-English-Speaking background populations. In addition, they cover very different data sets, time periods, population groups and techniques. Given these facts, greater research in this area can shed light on the factors contributing to differing results.

Our study fits in this literature by examining the incidence and the determinants of overeducation among immigrants and the native-born in Australia, and the impact of overeducation on earnings after accounting for individual heterogeneity. We utilize nine years of a rich longitudinal data set (Household, Income, and Labour Dynamics in Australia Survey (HILDA)). We apply the correlated random effects (CRE) logit model and the fixed effects earnings model to address endogeneity and unobserved individual heterogeneity. These approaches alleviate the major estimation concerns that are pertinent for determining individual job mismatches. To the best of our knowledge, this is the first examination of the determinants of overeducation and its impact on earnings among immigrants using these longitudinal techniques based on panel data and for Australia. We further examine effects for both English-speaking background (ESB) and Non-English-speaking background (NESB) immigrants compared to the Australian born population.

The remainder of this paper is organized as follows: The next section provides a brief overview of immigrants' overeducation literature and identifies the main factors that affect immigrant mismatch and labour market outcomes in the host country. Section 3 develops the econometric framework. Section 4 outlines the data and variables. The results are presented in Section 5 and conclusions are presented in Section 6.

2. REVIEW OF THE LITERATURE

Incorporating the match between the educational requirements of a job and the actual education of a worker in earnings models extends the human capital model by incorporating earnings variations across individuals with similar qualifications. The standard ORU (overeducation, required education, and undereducation) framework of earnings as originally proposed by Duncan and Hoffman (1981) and Hartog (2000), is widely used in overeducation empirical research. While this framework is usually applied to the labour market in general, it lends itself well to further understanding of immigrants' differential returns to years of schooling.

As such, the potential effect of education and occupation mismatches in a major

immigrant-receiving country such as Australia can shed light on this topic.

Relatively few studies have been conducted on the impact of education and occupation mismatches on immigrant earnings. The following review is not intended as a comprehensive review of the literature, but it points to major studies related to the current paper.

Applying the ORU framework, Green et al. (2007) and Kler (2007) utilized the survey of recent immigrants to Australia (the Longitudinal Survey of Immigrants Australia (LSIA)) for two cohorts of immigrants who had arrived in the 1990s. The surveys followed recent immigrants during the first 5 to 41 months of their residence in Australia. The researchers examined the incidence of overeducation and an augmented human capital earnings model (Frenette, 2004) of returns to required schooling and surplus schooling in the Australian labour market. They found that recent immigrants, even those with skill-assessed visas, are more vulnerable to overeducation than Australian natives in their first few years in the host country. NESB immigrants are more likely to be overeducated, with the incidence of overeducation being between 32 per cent and 49 per cent. Their evidence also signaled increasing rates of overeducation for the immigrant groups as the initial months of residence increased. NESB immigrants also had lower returns to required and surplus education than did Australian natives. Tighter welfare and support policies for immigrants may increase employment at the expense of under-utilizing their skills.[2] The analysis employed ordinary least squares (OLS) estimation. Given the nature of the data set on recent immigrants only, the examination of the link between educational mismatch and immigrants' economic integration with years since migration was outside of the scope of that study.

Chiswick and Miller (2010) reported that NESB immigrants have a lower rate of return to schooling accompanied by overeducation and undereducation. In their research they utilized the 1 per cent sample of the Australian 2001 Census of Population and Housing. They found the same payoff to required years of schooling for Australian-born, ESB and NESB immigrant groups. They reported penalties for years of overeducation for all groups. These results were consistent with their findings for the US (Chiswick and Miller, 2008), based on the 2000 Census, that the earnings return to each year of overeducation was lower than returns to each year of required education, but that the return to each year of overeducation was twice as much for the US-born (10.6%) compared to immigrants (5.2%).

Kler (2007) further examined the effects of overeducation among tertiary-educated recent immigrants, utilizing the two cohorts of the LSAI and random effects estimation methods for immigrant earnings. The study found that the payoff for overeducation was smaller than the payoff to required education for all groups. However, a result that differed from earlier studies

was that there was no significant effect of overeducation on earnings among Asian immigrants in the initial settlement months.

Tsai (2010) used a panel approach with US data and found that the earnings penalty to overeducation 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 in the scope of the study.

We extend the analysis of overeducation to Australian data on both immigrants and native-born Australians (longitudinal data (HILDA) across nine years of data). This allows us to extend our analyses by panel data methods for both incidence of overeducation and its wage effects across the two groups. We employ a correlated random effects (CRE) logit model with Mundlak correction (1978) to examine the incidence of overeducation. Importantly, the endogeneity due to the correlation between explanatory variables and error terms is addressed by Mundlak correction. We also employ panel fixed effects (FE) models to examine the effects of overeducation on earnings from years since migration. In the initial sample year, there is a wide range of years since migration among immigrants from recent arrivals to forty years since migration. These aspects of our study for the determinants of overeducation and earnings effects, along with the panel feature of the analysis for Australia, extend the international literature on the subject. We further consider both English-speaking and Non-English speaking background immigrants in relation to the same base group of the Australian-born for standardized comparisons.

3. DATA

The data used in this paper are taken from wave 1 to wave 9 (year 2001-year 2009) responding person file of the Household, Income, and Labour Dynamics in Australia (HILDA) Survey. The HILDA Survey, Australia's first nationally representative household panel survey, began in 2001 and interviews are conducted annually.

We use an unbalanced panel of the first nine waves of the HILDA survey. With pooled 2001-2009 data, the full sample size used in this study is composed of 15,833 observations of 2,502 individuals.

The inaugural HILDA survey for 2001 provides a nationally representative survey. In the initial year of 2001, the immigrant samples in HILDA demonstrate a wide range of birth countries and years-since-migration (0<, to >40 years). Our analysis follows this population group for nine consecutive years. A major advantage of the HILDA data set is that it allows the use of

appropriate panel data techniques based on longitudinal information, and a rich coverage of relevant variables. These factors allow controls for potential endogeneity and unobserved individual heterogeneity that are particularly important in the analyses of job mismatches. A second advantage of the data set employed is that it allows a significant duration for the analysis. During this time period the individuals in the sample move in and out of jobs and receive wages that change. This aspect is important given our interest in the effect of years-since-migration on job mismatches for immigrants. Thus, the strengths of the data lie in the careful analyses made possible for the population observed in 2001, and a lengthy longitudinal coverage of the individuals in the labour market after that year. A limitation of the data set, however, is that it mainly covers immigrants who were present in Australia by year 2001, with small additions to the sample over time. As a result, 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.

The sample for the current study includes all full-time male workers aged 23 to 64 in the initial survey year. [3] We consider full-time workers for a more comparable group of employees and earnings scales. Workers in part-time jobs may have chosen to do so for reasons of family or other personal commitments or preferences (e.g., flexibility of hours of work, shorter distances to work). Therefore, part-time workers may be more likely to accept mismatched jobs in terms of education and occupation match. These mismatches are likely to reduce workers' productivity and result in wage penalties. In addition, part-time jobs are shown to have a different pay structure, which adjusts for other job-related fringe benefits.

To control for potential sample selection, in addition to the analyses summarized in the paper, the potential impact of selection into both employment and full-time employment was examined using a Heckman selection adjustment. We also examined the impact of selection across years of data. [4] The results showed that control for selection for either type of selection did not change the results reported in this paper. [5]

Of this full-time sample, 79 per cent are Australian natives (born in Australia) and 21 per cent are immigrants (born overseas). Most ESB immigrants (specified by country of birth) come from developed countries such as the United Kingdom (50 %), New Zealand (23 %), South Africa (3 %), and the United States (3 %). Unlike the ESB immigrants, NESB immigrants are diverse, coming from over 60 different countries, including Vietnam (13 %), China (including mainland China, Hong Kong and Taiwan, 10 %), India (6 %), Philippines (5 %), and The Netherlands (4 %). NESB immigrants generally experience greater difficulty in adapting to their new lives even if they work in skilled categories. Furthermore, NESB immigrants may work in occupations that

require lower levels of educational attainment in instances in which their overseas credentials are not recognized by Australian employers.

The mean characteristics of the samples of the Australian-born population, ESB, and NESB immigrants are shown in Table 1.

[Table 1 here]

There is significant difference between these three groups across a number of personal characteristics. Notably, 44 per cent of NESB immigrants and 30 per cent of ESB immigrants have qualifications at or above a Bachelor degree, compared to 22 per cent of Australian natives. The average hourly wage for NESB immigrants is lower than that for ESB immigrants. There are also differences in the age profile and years since migration of these two groups of immigrants. This evidence encourages the test of the hypothesis that in comparison with ESB immigrants and Australian natives, NESB immigrants are more likely to undertake jobs in which they are overeducated; while controlling for relevant human capital factors and individual heterogeneity.

4. VARIABLES

The overeducation measure in our analysis is based on the Mode method (realized match method (Kiker et al., 1997), and it is derived at the two-digit occupational category level for greater accuracy.[6] We adopt the Mode method estimates of the level of required education by computing the amount of education that most commonly occurs within an occupational category (Rubb, 2003). Based on this cross-wave Mode method, there is a high incidence rate of overeducation in Australia. Evidence can be found from Table 2 that migrants are more likely to be overeducated than Australian natives.

[Table 2 here]

Notably, NESB immigrants experience more overeducation than their ESB counterparts. Table 2 shows that among full-time male workers aged 23 to 64, NESB immigrants have the highest rate of overeducation: 42 per cent of full-time NESB migrant workers are employed in positions that are below their educational attainment levels, compared to 31 per cent for ESB and 25 per cent for the Australian-born population. It reveals that mismatch is very serious among NESB immigrants.

The earning variables used in this study are log hourly wage from main job in constant dollars. Years since migration (YSM) measures years of residence in Australia for immigrants. We are interested in examining whether or not job mismatches are conditions that immigrants initially face, but that improve over time.

Our data set provides information on both actual years of education and qualifications completed. Years of actual education are derived by four variables from HILDA. To evaluate the effects of educational qualifications, we categorize degrees into five categories: Postgraduate, Bachelor, Diploma, Certificate (including the high-school completion certificate), and no qualification (high-school incompletion).

Age at migration is considered as a variable that affects immigrants' economic integration due to effects on language proficiency and education in the host country (see e.g. Pendakur and Pendakur, 2016). We define four cohorts based on their age at migration: 0-12, 13-22, 23-34, and 35-60. This specification is consistent with previous research that shows less than perfect transferability of university degrees, but that elementary school education is portable across national boundaries (Friedberg, 2000).

In addition, the proportion of poor English among NESB immigrants increases with age at migration, which as expected suggests that language proficiency is affected by age on arrival. As English is the main language in Australia, NESB immigrants with difficulties in speaking and writing English are more likely to decrease their expectations while job searching, and to accept jobs which require education below their level of attainment. Therefore, proficiency in spoken English may have a significant effect on the rate of overeducation and on immigrants' assimilation. We collapse four classifications into two: those who speak English well and those who speak English poorly. This variable identifies only four percent of Non-English Speaking immigrants that are in full-time employment as having poor spoken English. It may be that a larger percentage of the sample that experiences lower English fluency in reading and writing are not picked up by this variable. Given that, we also rely on the age at migration variable available in our data to examine the incidence of overeducation, which is expected to convey language proficiency as well.

The unemployment rate represents the annual percentage of the labour force that is unemployed and actively looking for work. It is also a common indicator of a country's economic conditions in a given year. In this study, it is used as a control for labour market conditions. We have collected the annual unemployment rate (year 2001 to year 2009) from the Australian Bureau of Statistics (ABS). High unemployment rates may force some workers to accept mismatched employment positions due to the limited availability of positions. This variable is an

annual rate. We also include fixed effects controlling for states in all models.

5. ECONOMETRIC FRAMEWORK

A longitudinal analysis is applied in this study to address the potential problem of "omitted unobservable bias" from cross-sectional analysis. This is important to identify both the incidence and potential earning penalty to overeducation.

5.1 Model 1: Determinants of overeducation

Model 1, examines the determinants of the incidence of overeducation in the panel model setting, and the hypothesis that the incidence of overeducation for immigrants may decrease with their duration of stay (YSM) in Australia. This less-examined hypothesis has important implications for understanding the labour market assimilation of immigrants in earnings models. We employ both a random effects logit model (without correction for endogeneity) and a correlated random effects logit model with Mundlak correction to estimate the determinants of overeducation to adjust for potential endogeneity of explanatory variables (preferred model). The random effects logit model was applied as a benchmark. The results from these two models were significantly different, and auxiliary tests confirmed the choice of the correlated random effects logit model as the relevant estimation method to account for endogeneity, as in Model 1 below.

Model 1 is presented in Equation (1) below. The functional form of the logit model is written as:

$$ln\left(\frac{Pr(over_{it})}{1-Pr(over_{it})}\right) = \delta_0 + \delta_1 \mathrm{Z_{it}} + \delta_2 \mathrm{M_i} + \delta_3 ED_{it} + \delta_4 QUA_{it} + \delta_5 (QUA_{it} * M_i) + \delta_6 YSM_{it} + \delta_7 YSM_{it}^2 + \delta_8 EXP_{it} + \delta_9 EXP_{it}^2 + \sum_{j=1}^{m} [\bar{X}_\iota \ \delta_j] + \eta_i + \varepsilon_{it} \quad (1)$$

$$\mu_i = \sum_{j=1}^{m} [\bar{X}_\iota \, \delta_j] + \eta_i \text{ , where } \eta_i \sim \mathrm{N}(0, \sigma_\eta^2); \quad \varepsilon_i \sim \mathrm{N}(0, \sigma_\varepsilon^2); \ \varepsilon_i, \text{iid}$$

$$i = 1, \dots, N; \quad t = 1, \dots, T; \quad j = 1, .., m$$

The observed variable $over_{it}$ takes the value of 1 if worker *i* is overeducated and is

defined as 0 otherwise. Z_{it} denotes a set of personal or job characteristics of individual *i* at time period t; ED_{it} denotes actual years of education, and variables QUA_{it} represent a series of binary variables for educational qualifications. M_i represents immigrant status. The coefficient of M_i, δ_2, measures the initial overeducation gap of immigrants upon arrival relative to the comparable Australian-born population. YSM_{it} denotes the number of years of residence since migrating to Australia. The coefficient of YSM_{it}, δ_6, measures the way in which the overeducation gap varies as immigrants spend time in the host country. The overeducation rates of immigrants are expected to signify their levels of economic assimilation. Therefore, the coefficient of YSM_{it} is predicted to be negative. To further examine the effects of experience in the host country on the probability of being over-educated, in the empirical section (Tables 3) we also report results in which we replace YSM_{it} and YSM_{it}^2 with age-on-arrival dummy variables in alternative specifications.

Finally, $\sum_{j=1}^{m}[\bar{X}_i\ \delta_j]$ represents the Mundlak adjustments (where *m* is the number of explanatory variables). These added components in the model, based on average values in the model across individuals and years of panel data, address endogeneity (such that the coefficients approximate unbiased fixed effects estimates (e.g. as shown by Wooldridge, 2010)).

5.2 Model 2: Impacts of overeducation on earnings

Model 2 focuses on the link between overeducation and earnings. The following questions are of interest in the empirical analysis: How does overeducation impact, directly or indirectly, on earnings via years since migration and migration status?

The standard ORU (overeducation, required education, and undereducation) earnings model, as originally proposed by Duncan and Hoffman (1981), decomposes actual years of education (S_a) into required years of education (S_r), years of overeducation (S_o), and years of undereducation (S_u). Thus $S_a = S_r + S_o - S_u$, where $S_o = S_a - S_r$ for the overeducated (i.e., if $S_a > S_r$), and 0 otherwise. Similarly, $S_u = S_r - S_a$ for the undereducated (i.e., if $S_r > S_a$), and 0 otherwise.

Then the log of earnings in the ORU model can be generally written as:

$$lny = \alpha_1 + \beta_r S_r + \beta_o S_o + \beta_u S_u + \delta_1 X_1 + \varepsilon \qquad (2)$$

lny is the natural logarithm of earnings; X_1 is a vector of control variables that generally includes personal characteristics and job characteristics; S_r, S_o, S_u are, respectively, the years of required education, overeducation, and undereducation; α_1 is the intercept term; and ε is an

error term.

Equation (2) estimates β_r, β_o, β_u continuously, as the rates of returns to required education, overeducation, and undereducation, respectively. Prior literature on overeducation has consistently found that $\beta_r > \beta_o$ and $\beta_o > 0$, such that the return of overeducation is lower than the return to required education, and the return to overeducation is positive (Cohn, 1992; Groot, 1996; Rumberger, 1987). In contrast, these studies also found that $\beta_u < \beta_r$ and $\beta_u < 0$, which means the return to undereducation is lower than the return to required education, and that it is a negative return (Hartog, 2000).

The extended ORU earnings model (Model 2 below) in a panel data setting, is built by adding interaction terms to examine the impacts of educational mismatch, years since migration and migrant status on the returns to overeducation, after controlling for the individual effects. By doing so, we can examine the earnings gap between immigrants and the Australian-born population via educational mismatch. These results examine an added and less-studied explanation for the existing earnings disadvantage for immigrants in the Australian labour market.

$$
\begin{aligned}
lny_{i,t} = {} & \beta_r S^r_{i,t} + \beta_o\left(S^a_{i,t} - S^r_{i,t}\right) + \beta_u\left(S^r_{i,t} - S^a_{i,t}\right) + \theta_1 \mathrm{Z_{it}} + \theta_2 \mathrm{M_i} + \theta_3 YSM_{it} + \theta_4 YSM^2_{it} \\
& + \left(S^r_{i,t} \times \mathrm{M}\right) + \beta_{oM}\left[\left(S^a_{i,t} - S^r_{i,t}\right) \times \mathrm{M}\right] + \beta_{uM}\left[\left(S^r_{i,t} - S^a_{i,t}\right) \times \mathrm{M}\right] \\
& + \left[\theta_{oYSM}(Over_{it} \times \mathrm{YSM}_{it}) + \theta_{oYSM2}(Over_{it} \times YSM^2_{it})\right] + \left[\theta_{uYSM}(Under_{it} \times \mathrm{YSM}_{it})\right. \\
& \left. + \theta_{uYSM2}(Under_{it} \times YSM^2_{it})\right] + \mu_i \\
& + \varepsilon_{\mathrm{it}} \qquad (3)
\end{aligned}
$$

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

$S^a_{i,t}$ denotes the years of actual education, and $S^r_{i,t}$ is the years of required education for individual i at year t. Thus, $(S^a_{i,t} - S^r_{i,t})$ is the years of overeducation when $S^a_{i,t} > S^r_{i,t}$, and 0 otherwise. Likewise, $\left(S^r_{i,t} - S^a_{i,t}\right)$ is years of undereducation when $S^r_{i,t} > S^a_{i,t}$ and 0 otherwise. Coefficients $\beta_o, \beta_u, \text{and } \beta_r$ estimate the magnitude of earnings effect of a one-unit change in years of overeducation, years of undereducation, and required years of education, respectively, among the Australian-born population. The coefficients of the interaction terms, $\beta_{\mathrm{oM}}, \beta_{\mathrm{uM}} \text{ and } \beta_{\mathrm{rM}}$ evaluate the difference of earnings effects between Australian natives and immigrants who have the same type of educational mismatch.

$\mathrm{Z_{it}}$ denotes a set of personal characteristics, such as years of experience. $lny_{i,t}$ is the natural log of hourly wage from main job in constant (2009) dollars for the *i*th individual in period t. Variable $Over_{it}$ is a time-variant binary variable, which takes the value of 1 if individual *i* is

overeducated at time period t; $Under_{it}$ takes the value of 1 if individual *i* is undereducated at time period t. Educationally matched is the reference category.

The coefficient of M_i, θ_2, denotes the initial earnings gap between immigrants and Australian natives.

YSM_{it} denotes the number of years of residence since migrating to the host country for individual *i* at time *t*. The coefficient of YSM_{it}, θ_3, denotes assimilation effects. Based on previous studies, θ_3 is expected to have a positive sign. But if immigrants work in jobs requiring qualifications that are below their educational attainment, this may lengthen their assimilation process with the consequence that they catch up with the earnings of the Australian-born population slowly over time, or not at all. Thus, the coefficient of interaction terms, $Over_{it} \times YSM_{it}$, $\theta_{oYSM,}$ is negative if overeducation slows immigrants' earnings assimilation in Australia. Unobserved factors, such as ability, motivation or work effort influence earnings, and also are correlated with observed education and skills. If these unobserved individual effects, $u_{i,}$ are correlated with explanatory variables, cross-section analysis would result in omitted unobservable biases. Longitudinal data captures the same individual over time. Thus, unobservable individual effects are eliminated by using a panel fixed effects model, such that estimation results from fixed effects models are consistent. However, this model cannot evaluate the time-invariant explanatory variables because they are removed by within-group transformation. In contrast, a random effects generalized least squares (GLS) model assumes that u_i is uncorrelated with explanatory variables in which GLS uses the optimal combination of within-group and between-group variations. If individual effects do not matter, then the GLS estimator is equal to the ordinary least squares (OLS) estimator. A Hausman test is used to identify whether the random effects GLS estimator is biased.

The error term is denoted by $\mu_i + \varepsilon_{i,t}$. The unobservable individual-specific effect μ_i is assumed not to change over time and the random disturbance, ε_{it}, is assumed to be independent and identically distributed, i.i.d (0, σ_ε^2).

6. ESTIMATIONS AND RESULTS

In this section we summarize the results on both the incidence of job mismatches and earnings effects. Throughout this analysis we employ two samples separately for comparison purposes. The first sample contains ESB immigrants and Australian natives, and the second sample consists of NESB immigrants and Australian natives. Thus, we can determine specific effects for ESB and NESB immigrants by comparing them with the Australian-born population as a common base.

The dependent variable for the outcome equation is the odds ratio of being overeducated.

Our earnings model, in turn, examines overeducation effects on earnings via years since migration. This model examines potential earnings penalties associated with overeducation and it demonstrates the effects of overeducation on immigrants' assimilation.

6.1 Model 1 Results: Determinants of overeducation

As discussed previously, duration of residency (YSM) and age on arrival (e.g. migration as a child or later) may influence the rate of overeducation. While there is a strong overlap of impacts as measured by these two variables, they reflect somewhat different effects on language proficiency, country of higher education, institutional knowledge of the host country, and social networks. To examine the effects of both variables, based on Equation (1), we consider two alternative specifications, based on 'years since migration' and 'age at migration'. The examination of these two aspects of an immigrant's history in the host country provides a more comprehensive analysis.

Marginal effects are reported, and derived as the coefficient multiplied by the density function (the probability of a positive outcome), evaluated at sample mean values of explanatory variables.

Based on Model 1 in Equation (1), the results of the estimations for Australian natives and ESB immigrants, and for Australian natives and NESB immigrants, are reported in Table 3. Overall, the results show that university graduates and immigrants regardless of country of origin group, are more likely to be overeducated. However, there are major differences between ESB and NESB immigrants in relation to age at migration.

The propensity of overeducation for immigrants is 88 to 94 per cent higher than for the Australian-born population. This result indicates that the incidence of overeducation among NESB immigrants is worthy of closer examination and review with panel data techniques. This evidence of effects for NESB immigrants is significantly greater than results in studies for recent immigrants (e.g. Green et al., 2007; and Kler, 2007).

Among NESB immigrants, overeducation measures are also higher for immigrants who have arrived as adults, compared to child immigrants. Migrating as a child helps NESB migrants to reduce the probability of being overeducated in employment. Immigrants who migrated at less than 12 years of age have a 10 per cent lower probability rate of overeducation compared to others who migrated at age 34 to 60. Immigrants who arrived in the host country between the ages of 13 and 22 also have an advantage, but the advantage (marginal effect of − 6.5 %) is

smaller than for child arrivals. This effect is robust and it is more pronounced. As noted earlier, we expect that this variable reflects the impact of language fluency, education in the host country, and institutional knowledge of the country. These effects do not apply to ESB immigrants. The finding that arrival as a child significantly reduces the likelihood of overeducation among Non-English Speaking Background immigrants, but is not relevant to English-Speaking Background immigrants, is revealing.

The results also show that for NESB immigrants, the incidence of overeducation decreases with years of experience, but the effect is very modest. In addition, for this group, in our panel model specifications that account for unobservables with the Mundlak correction the incidence of overeducation does not seem to change significantly with years since migration.

[Table 3 here]

6.2 Model 2 Results: The Impact of overeducation on Earnings

Following Model 2 in Equation (3), estimation results are given in Table 4. There are four columns for the two specific subsamples. Columns 1 and 2 report results from pooled OLS, fixed effects estimation for ESB immigrants and Australian natives. Columns 3 and 4 present the results for NESB immigrants and Australian natives. The Hausman test results for the models in Table 4 reject the null hypothesis that individual specific error is uncorrelated with the explanatory variables of the wage equation. Therefore, fixed effects estimates were selected over the random effects to address individual heterogeneity. We report pooled OLS estimation as a benchmark to examine unobserved heterogeneity effects.

The results in rows 1-3 show ORU model effects of educational mismatches for the immigrant populations, relative to the native born. A comparison of the fixed effects (FE) results that control for individual heterogeneity, compared to the pooled OLS results in Table 4 gives the following findings:

Firstly, a number of the coefficients change dramatically, indicating the importance of controlling for endogeneity of unobserved individual heterogeneity effects. The effects for ESB immigrants in FE estimations are indistinguishable from those for the Australian-born population (as shown by the insignificant effects on interaction terms between years of overeducation and immigrant status). This result (together with the indication of a positive bias) is consistent with the frequently expressed thought that unobserved factors such as motivation (or compensating job characteristics) might explain the lower earnings observed among the native-born and ESB

immigrant group of workers in positions for which they are overeducated.

In contrast, according to the panel fixed effects estimation in Column (4), NESB immigrants experience a 9 per cent lower return for the additional years of overeducation than comparable Australian natives. Similar effects are also found in the returns to years of required education. This is shown in Column (4) – for each year of required education, NESB immigrants have a 9 per cent lower return than Australian natives have. This indicates that NESB immigrants also have relatively lower earnings returns not only from education-occupation mismatch, but also when they possess adequate years of education. In addition, the impact of overeducation on earnings does not change for this group of NESB immigrants with years since migration. We expect that this result may reflect the effects of factors such as less than complete transferability of overseas degrees and language proficiency. The significant and larger effects from age at migration in our panel data analyses (Table 3) point to these two factors.

A second finding is that a comparison of FE and pooled OLS results indicates that for NESB immigrants the penalty for overeducation compared to the native-born is significantly larger in the FE models for each year of overeducation compared to the native-born (column 4). These results (and the indication of a negative bias) are consistent with the hypothesis that other types of unobserved factors contribute to the lower earnings observed among the group of NESB immigrants in mismatched jobs.

The fixed effects models further show that after accounting for individual effects, years since migration (YSM) have a significant effect on earnings of immigrants. That is, an ESB immigrant improves his earnings by 2.4 per cent for each year of staying in Australia. For NESB immigrants the effect of each year since migration on earnings is 1.4 per cent. The results suggest improvements for both groups in general, and a much stronger effect on earnings assimilation for ESB immigrants than for NESB immigrants.

In addition, we find a significant impact from poor spoken English in the Pooled OLS model, but not in the fixed effects results (Table 4), most likely since the data does not show major changes in language proficiency over the nine years of the data, for the small (4 percent) group who reported poor spoken English.

[Table 4 here]

7. CONCLUSION

In this study we have examined the impact of education-occupation mismatches for immigrants,

using nine years of longitudinal data and panel data analyses covering a desired wide range of age groups and years since migration. We have examined effects for ESB and NESB immigrants, as well as for Australian-born full-time employed males.

Firstly, we find that 42 per cent of Non-English Speaking Background (NESB) immigrants and 31 per cent of English Speaking Background (ESB) immigrants work in jobs which require a lower educational standard than the one they possess. The determinants of overeducation are examined by means of a correlated random effects logit model with Mundlak correction. After endogeneity is accounted for, immigrants demonstrate significantly higher (approximately 90 percent higher) probability of overeducation than does the Australian-born population. As time passes, the education-occupation mismatch status for immigrants does not change with increased years since migration. Among NESB immigrants, younger entrants (who migrated at less than 22 years of age) have a significantly lower probability of overeducation than older entrants, and immigrants who have migrated at less than age 12 have a significantly lower probability that is similar to the Australian-born population. For ESB immigrants, age at migration is also not a significant determining factor.

Secondly, we find divergent results for ESB and NESB immigrants in relation to the absolute and relative earnings returns to years of education. We find that returns to education are indistinguishable for the native born and ESB immigrants. However, NESB immigrants experience earnings that are close to 9 per cent lower for each year of overeducation and required education, compared to the native born, after controlling for individual heterogeneity. Since overeducation is more prevalent among immigrants, an implication of these results is that in the study of overeducation it is important to consider Non-English Speaking Background immigrants separately, since the effects for this group exhibit significant differences from the native-born and ESB groups.

The above evidence further highlights that the lower returns to education among NESB immigrants relative to the native born is due not only to skill underutilization, but also to a return to education disadvantage that cannot be accounted for by the extensive human capital variables included in our models. The significant and larger effects from age at migration in our panel data analyses point to factors such as aspects of language proficiency or perceptions of it that persist over time, and incomplete transferability of educational degrees for NESB immigrants. These findings have implications for immigration assimilation policies beyond attracting skilled immigrants. Policy tools that require and facilitate language proficiency in all relevant areas of writing, reading, and spoken language, among recent immigrants, such as those recently implemented in Australia, New Zealand, and Canada, target these mechanisms.

ACKNOWLEDGEMENTS

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 findings and views reported in this paper, however, are those of the authors and should not be attributed to the above individuals, FaHCSIA or the Melbourne Institute.

NOTES

1. For example, the number of skilled migrants arriving in Australia grew on average by more than 10 per cent per year over two decades, from 18,300 in 1993-94 to 128,550 in 2013-14. In addition, with a gradual policy shift, the current intake places for skilled migrants make up 68 per cent of the 2014-15 Migration Programme (DIBP 2016).
2. For example, stringent entry standards were applied to skill and English language ability test scores and eligibility to claim welfare and unemployment benefits was extended from 6 to 24 months (except for humanitarian visa holders).
3. According to HILDA's (Wave 1) Person Questionnaire, full-time work is specified as 35 hours per week or more.
4. The HILDA data set applied has a remarkably high response rate throughout the period. Re-interview rates are high (e.g. 96.3 % in wave 9).
5. These results are available on request.
6. Kiker et al. (1997) and Verhaest and Omey (2006) show that the Mode method is preferred to Verdugo and Verdugo's (1989) mean criterion. They found Verdugo and Verdugo's mean criterion to be changing gradually and that it could produce classification errors before correcting itself, but that the Mode changes more freely, reflecting each period's educational requirements of most workers at a given time.

REFERENCES

Aringa, C. D., & L. Pagani
2010 "Labour market assimilation and over education: The case of immigrant workers in Italy", No.178. Universita degli Studi di Milano-Bicocca. Dipartimento di Economia Politica.

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

Chamberlain, G.
2010 "Binary response models for panel data: Identification and information", *Econometrica,* 78(1): 159-168.

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

Chiswick, B.R., & P.W. Miller
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.

DIBP (Department of Immigration and Border Protection)
2016 "Historical migration statistics", http://www.border.gov.au/about/reports-publications/research-statistics/statistics/live-in-australia/historical-migration-statistics (accessed 16 December 2016).

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

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

Fernández, C., & C. Ortega
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., et al.
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.

Kiker, B. F., et al.
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.
Leuven, E., & H. Oosterbeek
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., et al.
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.
Mundlak, Y.
1978 "On the pooling of time series and cross section data", *Econometrica,* 46(1): 69-85.
Quinn, M. A. & S. Rubb
2006 "Mexico's labor market: The impact of education-occupation matching on wages and productivity in developing economies", *Economics of Education Review*, 25: 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.
Sloane, P.
2014 "Overeducation, skill mismatches, and labor market outcomes for college graduates", *IZA World of Labor* (2014:88).
Tsai, Y.
2010 "Returns to overeducation: A longitudinal analysis of the U.S. labor market", *Economics of Education Review*, 29: 606-617.
Tsang, M. C., & H. Levin
1985 "The economics of overeducation", *Economics of Education Review,* 4(2): 93-104.
Verdugo, R. R., & N.T. Verdugo
1989 "The impact of surplus schooling on earnings", *Journal of Human Resources,* 24(4): 629-643.
Verhaest, D., & E. Omey
2006 "Discriminating between alternative measures of overeducation", *Applied Economics,* 38(18): 2113-2120.
Wald, S., & T. Fang
2008 "Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey", *Canadian Public Policy,* 34(4): 457-479.
Pendakur, K. & R. Pendakur
2016 "Which Child Immigrants Face Earnings Disparity? Age-at-immigration, Ethnic Minority Status and Labour Market Attainment in Canada", *International Migration*, 54(5): 43-58.
Wooldridge, J.M.
2010 *Econometric analysis of cross-section and panel data*, Second edtion, MA: MIT Press, Cambridge.
Yeo, J. Z. & S.A. Maani
2015 "Educational mismatches and earnings in the New Zealand labour market", *New Zealand Economic Papers,* 51: 28-48.

TABLE 1

SUMMARY STATISTICS BY COUNTRY OF BIRTH GROUP

VARIABLES	Australian Natives mean	sd	ESB Immigrants[1] mean	sd	NESB Immigrants[2] 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)
Highest Qualification						
Postgraduate	0.09	(0.31)	0.16	(0.37)	0.18	(0.40)
Bachelor	0.13	(0.36)	0.14	(0.37)	0.26	(0.43)
Diploma	0.11	(0.30)	0.11	(0.31)	0.12	(0.33)
Certificate	0.33	(0.47)	0.29	(0.45)	0.16	(0.37)
Age	41.12	(9.96)	44.70	(9.91)	43.81	(9.79)
Job Characteristics						
Hourly wage of main job	29.81	(15.75)	32.49	(18.23)	29.50	(15.24)
Unemployment Rate	0.05	(0.01)	0.05	(0.01)	0.05	(0.01)
Years since Migration						
YSM	/	/	26.32	(13.08)	20.65	(12.65)
Age on Arrival	/	/	18.38	(12.08)	23.16	(11.56)
Age 0-12	/	/	0.40	(0.49)	0.22	(0.41)
Age 13-22	/	/	0.17	(0.37)	0.24	(0.42)
Age 23-34	/	/	0.34	(0.47)	0.38	(0.49)
Age 35-60	/	/	0.10	(0.29)	0.16	(0.37)
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 NESBImmigrants (Non-English Speaking Background Immigrants)

TABLE 2

THE EXTENT OF OVEREDUCATION BY COUNTRY OF BIRTH GROUP

	Australian Natives		ESB Immigrants		NESB Immigrants	
VARIABLES	mean	sd[4]	mean	sd	mean	sd
Educational mismatch						
Overeducated	0.25	(0.43)	0.31	(0.46)	0.42	(0.49)
Undereducated	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)
Years of overeducation[1]	2.31	(1.74)	2.26	(1.61)	2.55	(1.76)
Years of undereducation[2]	2.99	(1.53)	3.05	(1.60)	2.83	(1.53)
Years of required education[3]	14. 39	(1.37)	14.54	(1.38)	14.84	(1.36)
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:
1 Overeducated sample mean
2 Undereducated sample mean
3 Matched sample mean
4 sd (standard deviation)

TABLE 3

DETERMINANTS OF OVEREUCATION (MODEL1)
Marginal effects (standard errors)

	ESB Immigrants and Australian Natives			NESB Immigrants and Australian Natives		
	(1) Specification 1 YSM Effects	(2) Specification 2 Age on Arrival Effects	Pr ($over_{it}$ $\|u_i=0$)= 11.8% Mean of X	(3) Specification 1 YSM Effects	(4) Specification 2 Age on Arrival Effects	Pr($over_{it}$ $\|u_i=0$)= 11.9% Mean of X
Explanatory Variables						
Immigrant (M)	0.924***	0.937***	0.143	0.879***	0.937***	0.093
	(0.079)	(0.059)		(0.167)	(0.030)	
Human Capital						
Years of education	-0.029	-0.027	13.720	0.117***	0.116***	13.746
	(0.584)	(0.597)		(0.031)	(0.031)	
Postgraduate	0.223	0.229	0.115	0.258	0.261	0.114
	(0.362)	(0.365)		(0.395)	(0.396)	
Bachelor	-0.116*	-0.115*	0.146	-0.113	-0.113	0.154
	(0.064)	(0.064)		(0.071)	(0.071)	
Certificate	-0.215***	-0.214***	0.320	-0.211***	-0.211***	0.312
	(0.069)	(0.069)		(0.070)	(0.070)	
EXP	-0.149	-0.147	22.264	-0.006**	-0.007**	21.924
	(0.585)	(0.599)		(0.003)	(0.003)	
EXP^2	0.010	0.010	6.106	0.016**	0.015**	5.955
	(0.007)	(0.007)		(0.008)	(0.008)	
Years since migration (YSM)	-0.003	/	26.320	0.000	/	20.650
	(0.005)	/		(0.006)	/	
Age on Arrival						
Age 0-12	/	-0.032	0.399	/	-0.098***	0.220
	/	(0.043)		/	(0.018)	
Age 13-22	/	-0.052	0.166	/	-0.065*	0.235
	/	(0.038)		/	(0.037)	
Age 23-34	/	-0.001	0.340	/	-0.046	0.379
	/	(0.055)		/	(0.045)	
Individuals	2,313	2,313		2,185	2,185	
Observations	14,711	14,711		13,808	13,808	
Log likelihood	-4504	-4504		-4208	-4211	

Dependent Variable : $Over_{it}$ (overeducated in time period t)

Notes:

Correlated Random Effects Logit Estimations

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

Base-categories are the Australian-born population, no qualification, Age 35-60, Year 2009, and QLD. The model also includes an Interaction items between qualification categories and immigrants; disability; poor English; and States fixed effects; unemployment; time periods; and insignificant coefficient for Diploma. Full Results are available upon request.

Sources: HILDA-Release 9 (Wave 1-Wave 9).

TABLE 4

THE EFFECTS OF OVEREDUCATION ON EARNINGS (MODEL 2)

Coefficients (standard errors)

	ESB Immigrants and Australian Natives		NESB Immigrants and Australian Natives	
	(1) Pooled OLS	(2) Fixed Effects	(3) Pooled OLS	(4) Fixed Effects
Explanatory Variables				
Human capital				
Years of overeducation × M	-0.081***	-0.020	-0.030	-0.089**
	(0.019)	(0.035)	(0.025)	(0.040)
Years of undereducation × M	0.105***	-0.009	0.037*	0.070
	(0.016)	(0.037)	(0.021)	(0.045)
Years of required education × M	-0.075***	-0.004	-0.033	-0.087**
EXP	0.023***	0.040***	0.024***	0.039***
	(0.001)	(0.003)	(0.002)	(0.003)
$EXP^2/100$	-0.042***	-0.052***	-0.046***	-0.050***
	(0.003)	(0.005)	(0.003)	(0.005)
Years since migration (YSM)				
YSM	0.003	0.024***	0.023***	0.014*
	(0.004)	(0.007)	(0.005)	(0.008)
$YSM^2/100$	-0.008	-0.035***	-0.031***	-0.008
	(0.008)	(0.011)	(0.010)	(0.015)
Overeducated ×YSM	0.003	0.004	-0.010**	0.004
	(0.004)	(0.004)	(0.005)	(0.004)
Overeducated ×$YSM^2/100$	-0.011	-0.010	0.026**	-0.015
	(0.010)	(0.009)	(0.012)	(0.011)
Undereducated ×YSM	YES	YES	YES	YES
Undereducated ×$YSM^2/100$	YES	YES	YES	YES
Immigrant (M)	0.879***	/	0.070	/
	(0.208)	/	(0.296)	/
Poor English	/	/	-0.190***	0.052
	/	/	(0.066)	(0.078)
Individuals	2,313	2,313	2,185	2,185
Observations	14,711	14,711	13,808	13,808

Dependent Variable : The natural logarithm of hourly wage of main job

Notes:

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

Base categories are Australian-born population, no qualification, being matched × YSM, being matched YSM SQR/100, QLD. The models include years of overeducation, years of undereducation, years of required education, qualifications dummy variables, qualifications dummy variables x M, disability and and States fixed effects; unemployment; and time periods;. Full Results are available upon request.

Sources: HILDA-Release 9 (Wave 1-Wave 9).