I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15943

Earnings Penalty of Educational Mismatch: A Comparison of Alternative Methods of Assessing Over-Education

Le Wen
Sholeh A. Maani

FEBRUARY 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15943

Earnings Penalty of Educational Mismatch: A Comparison of Alternative Methods of Assessing Over-Education

Le Wen
University of Auckland

Sholeh A. Maani
University of Auckland and IZA

FEBRUARY 2023

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

Earnings Penalty of Educational Mismatch: A Comparison of Alternative Methods of Assessing Over-Education*

In this paper we systematically evaluate the impact of using the alternative methods conventionally used in the international literature on the measured incidence of educational mismatch and its earnings effects. We use a rich Australian longitudinal data set for a controlled group of full-time employed workers. Using panel data estimation, we address individual heterogeneity and measurement error, which are important in educational mismatch analysis. We show that alternative methods of measurement result in a range of estimates, with the Mode measure providing the most stable results across instrumental variables (IV) selections in panel fixed effects instrumental variables (FEIV) estimations. Based on the Mode measure, the incidence rate of over-education is 32.3%. The earnings penalty for each year of over-education is 2.5%, which is larger than 0.6% in fixed effect estimation and also larger than 1.9% in OLS estimations.

JEL Classification: J24, J31

Keywords: over-education, education mismatch, earnings, measurement error, instrumental variable

Corresponding author:
Sholeh A. Maani
Department of Economics
University of Auckland
12 Grafton Road
Auckland 1010
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 findings and views reported in this paper, however, are those of the authors and should not be attributed to either FaHCSIA or the Melbourn.

1. Introduction

When a worker's formal education is not fully utilised in his or her occupation, this under-utilisation is referred to as over-education. Education mismatches in employment have important potential impacts on economic productivity and policy design. The stylised fact is that, by controlling for required education for a job, over-educated workers earn more than their adequately educated counterparts (holding other characteristics constant), but earn less than workers with similar levels of schooling who work in occupations that demand the level of schooling they have acquired (Hartog 2000; Rumberger 1987; Sicherman 1991). In addition to this negative effect on an individual's earnings, there is evidence that under-utilisation of education reduces company revenue and potentially impedes economic growth, reflecting inefficiencies in the allocation of human resources (Tsang et al. 1991). As such, the provision of skills that enhance job opportunities for the population and lead to enhanced economic production are among important goals and budgets of governments.[1] But 38.7% of workers were observed to be qualification mismatched, and 20.2% were over-qualified in Australia (OECD 2016).

In explaining the existence of a lower pay rate for years of over-education, human capital theory (Becker 1964) does not focus on job characteristics. Instead, it is based on assumptions where workers' earnings are determined by their educational attainment. This theory, therefore, can explain the existence of over-education based on the supply side, and the assumption that productivity is an increasing function of the human capital level of the worker.

An extension of human capital theory in this context is occupational mobility theory (or career mobility theory), which postulates that over-education is a temporary, short-term mismatch phenomenon (Rosen 1972). This theory predicts that an over-educated worker will accept low-level jobs to gain work experience and training to move to higher levels on the occupational ladder. In contrast, Thurow's (1975) job competition (job assignment) theory interprets the existence of over-education based on demand-side theory. Marginal productivity depends on job characteristics, and the individual's earnings depend on his or her job's characteristics, rather than his or her own personal characteristics. Including both supply and demand-side parameters, assignment theory explains the phenomenon of over-education according to the position and the worker's characteristics (Duncan and Hoffman 1981).

According to human capital theory and occupational mobility theory, over-education is expected to be only a temporary period for workers who will get promoted within the firm or change to a matched job in a different firm. Job competition theory and assignment theory explain that over-education is a persistent phenomenon and brings inefficient resource allocation, which should be researched seriously. Based on the same data set as this paper, Wen and Maani (2019) have shown that, contrary to career theory, education-occupation mismatch results in a significant disadvantage in occupational growth for individuals. But the study did not examine the impact of over-

[1] For example, government tertiary education expenditure in Australia was AU$33.03 billion, accounting for 2% of GDP in 2015-2016 (ABS 2016).

education on wages directly. Thus, this paper complements Wen and Maani's (2019) research in its scope.

In determining whether a worker is over-educated or not, evidence-based analyses use alternative methods to define the education required for a job. In the literature on educational mismatches, different studies use one or another of the alternative conventional measures. Each of the widely-used conventional measures (i.e., Job Analysis (JA), Realised Match (RM) and Self-Reported (SR)) have some advantages and some drawbacks. Studies utilising these alternative measures have often found a wide range of, and at times opposing, results based on which alternative method of educational mismatch measurement is used. For example, Verhaest and Omey (2010) noted the sensitivity of results to the measure used. Additionally, Robst (1994) examined more than 200 individuals who were over-educated according to one measurement method, but were under-educated according to another method. The incidence of over-education further ranges from 11% using RM (Verdugo and Verdugo 1989) to over 50% using SR and JA measures (Tsang et al. 1991). In other studies, researchers have found that somewhat compatible effects may be obtained by adopting alternative measures of over-education (Cohn and Khan 1995; Rumberger 1987; Sicherman 1991).

But remarkably, there is no systematised agreement in the literature as to how the choice of the measurement method is expected to affect the estimated incidence and the earnings impacts of mismatch. Thus, from a research or policy point of view, a better understanding of the reasons for how alternative measures lead to different results is imperative. Notably, if a mismatch measurement method systematically identifies lower rates (or years of) over-education than the other measures, it will classify some 'mismatched' over-educated individuals, identified as 'mismatched' based on other measures, as 'matched' workers. As a result, in the econometric estimation of earnings regression models, it is further expected to systematically result in larger coefficients of earnings penalty for over-education than other measures.

We contribute to this literature by providing systematised analyses based on conventional measures used in the literature. Notably, we extend the literature by incorporating econometric methods that adjust for individual heterogeneity and potential measurement error as part of our systematised analyses for more precise estimations. We focus on the conventional measures of Job Analysis (JA) and Realised Match (RM) and show the importance of considering which of the specific measures is used when interpreting results in the literature. These measures are discussed in Section 2.

The JA method is a systematic evaluation by professional job analysts who specify the level and type of education required from resources such as the Standard Occupational Classification System (SOCS) in the United Kingdom, the Australian and New Zealand Standard Classification of Occupation (ANZSCO) in Australia, and the Dictionary of Occupational Titles (DOT) Handbook in the United States.

The RM method includes Mean and Modal Education (Mode) measures of required education for an occupation. The Mean method refers to the empirical measure of over-education in which a worker's education is higher than one standard deviation above the average for his or her occupation code (e.g., Verdugo and Verdugo 1989). The Mode method, proposed by Kiker et al. (1997), estimates the level of required education by comparing the actual years of education of the individual to the mode in the occupation (Rubb 2003). Mode is the statistical measure of the amount of education that most commonly occurs within an occupational category. In this method, employees who have more years of education than the mode in their occupation are classified as over-educated. This method provides both incidence measures and years (or fraction of years) of over- or under-education. The Mode method has been extensively applied to job mismatch studies in panel data settings, such as examining the relationship between mismatch and labour market outcomes (Mavromaras et al. 2013), assessing immigrants' over-education and earnings (Wen and Maani 2018), and determining the impact of job mismatches on occupational advancement and wage growth (Wen and Maani 2019).

The Self-Reported (SR) method, adopted in some papers, evaluates over-education by asking a survey respondent about the educational level required for his/her occupation.

Although there are previous studies on educational mismatch with comparative results, a majority of analyses apply cross-sectional data (e.g., Robst 1994; Verhaest and Omey 2010; Leuven and Oosterbeek 2011). As such, they assume unsystematic unobservables and a random assignment of workers to jobs. Other studies have used panel data sets to examine the labour market outcomes of educational mismatches while controlling for individual heterogeneity (Bauer 2002; Dolton and Silles 2008; Tsai 2010; Mavromaras et al. 2013; Sellami et al. 2017; Wen and Maani 2018a,b). However, it was outside the scope of these studies to systematically compare the impact of alternative measures of over-education by adopting multiple alternative instrumental variables when addressing measurement error.

Our analysis compares and highlights, for future research, the systematic ways in which these measures are expected to result in classifying comparatively higher or lower incidences of educational mismatch (Section 4). We provide a systematic comparison of estimates associated with each of these alternative measures. We apply those measures to examine the incidences of over-education and under-education. We then examine the expected effects of these incidence categorisations, as determined by each measure, on regression estimates of earnings returns to years of matched and over- or under-education. Notably, employing a rich Australian longitudinal data set over a nine-year period, we use the panel features of the data set to control for individual heterogeneity, based on fixed effects models. Based on these results, we find earnings penalties to years of over-education (and under-education) based on econometric models of earnings that account for potential educational job mismatches (Sections 5 and 6).[2] We further provide comparative results with and without adjustments for individual unobserved heterogeneity and measurement error

[2] The earnings penalty in this literature refers to the difference between the earnings returns to each year of over-education (and under-education), in comparison to the returns for each year of required education. Therefore, the penalty estimate also provides a measure of the difference between the earnings returns to each year of over-education (and under-education), compared to workers with similar years of education, but who are engaged in occupations that match their education level.

to alleviate potential estimation issues and, for comparative purposes, on stability of conventional measures before and after these adjustments.

Our contribution to the literature is threefold. First, the study statistically examines alternative measures to define over-education; why some measures have certain advantages over the other measures based on their statistical properties; and the expected directions of variation in results (Sections 2-4). Our comparative empirical results for these four measures, based on the same rich data set, provide supporting evidence on the direction of variations in the estimated results based on which measure is used. This feature of the analysis highlights the importance to better scrutinise and consider the role of the measures applied in identifying the penalty/reward of over-education in the studies conducted and considered by researchers and policymakers.

Second, the study extends the literature in the analysis of the earnings penalty to over-education by applying panel fixed effects estimation using multiple alternative instrumental variables (FEIV) to address both individual heterogeneity and measurement error (Iriondo and Pérez-Amaral 2016; Verhaest and Omey 2012), along with our comparative analyses (Section 5). Addressing both of these issues is important to any analysis of educational mismatch, for a precise assessment of the earnings penalty of educational mismatch, and aimed at reflecting the extent of human resource allocation efficiency.

Third, in an extension of our analysis, we incorporate the self-reported 'skilling' variable that directly captures utilisation of 'ability and skills' into the model (Section 6). Combining over-education and skill-underutilisation, with different instrumental variables incorporated in our analysis, offers further new evidence by showing that over-educated and over-skilled workers are the group which policymakers should focus on.

The plan for the remainder of the paper is as follows. Section 2 examines the characteristics of alternative measures commonly used across studies, and provides a brief overview of the literature on the link between over-education and earnings. A discussion of data and variables is provided in Section 3. Section 4 evaluates four measures to assess the incidence of educational mismatch. Sections 5 and 6 describe the methodology and the empirical results, respectively. Section 7 concludes the paper.

2. An overview of the empirical literature

Alternative measures have been applied in over-education studies. In the education job mismatch literature, there is notable variation across the estimates of both the incidence and the earnings impacts of educational mismatch, based on the mismatch measures applied. Therefore, our task in this paper is to compare and standardise the impact of the specific measure chosen on the relative size of the estimated impacts. In our analysis, we provide auxiliary results based on an extension of the model with combined measures of over-education and over-skilling. The use of skilling variables can also address the bias from potential unobserved heterogeneity. Additionally, using instrumental variables in the panel data model addresses measurement error.

In this section, we first discuss the merits and drawbacks of a broader set of measures and then briefly review the literature on impacts of over-education on earnings. The estimation results in this paper aim to shed light on the impact of the different conventional measures of job mismatch in relation to this international literature, including when these important corrections are also implemented in our modelling.

2.1 Alternative measures

Job Analysis (JA) Measure

The JA measure represents the core concept of over-education, defined as the under-utilisation of skills (Hartog 2000). Evaluations generated by professional job analysts include explicit definitions and comprehensive information about the qualifications required to undertake a job. The extent of substitution of various types of education can be shown by analysing the particular type of activities to be performed when doing that job.

However, despite this advantage, a challenge in using this measure relates to job aggregation. JA does not account for the educational variations in tasks within occupations; the job analyst considers that all of the roles that come under the same occupation title necessitate the same educational requirement. This aggregation inevitably leads to both an aggregation error, by ignoring the heterogeneity within an occupation (Halaby 1994), and a classification error (Verhaest and Omey 2006a). These errors in random measurement can lead to heterogeneity error.

A second challenge is that keeping JA educational requirements up to date, by establishing new occupations and regularly updating the qualification level required to perform in an existing occupation, involves frequent reviews across occupations. Due to high expenditures, existing requirements may at times be prone to becoming out of date and lacking depth. These failings, if present, are likely to introduce bias into the criteria established for any particular qualification. Given this potential drawback, both the reliability and the validity of JA have been questioned by some researchers. For example, Verdugo and Verdugo (1992) found that, based on the DOT Handbook, a single job analyst visits a job site and discusses job requirements with the employer. Van der Meer (2006) evaluated the validity of two JA measures on reflecting up-to-date requirements. Although JA is a general method that is widely used in the literature, due to its aggregation and judgement errors, including time-lag occurring in the observation, it is important to consider these underlying questions in interpreting the results.

Realised Matches (RM) Measures

The RM measure consists of Mean and Modal Education (Mode) measures.

The benefit of the Realised Matches methods is that both the mean and the mode can be derived directly from the existing data, so they are always available and up to date.

The Mean measure has its drawbacks. Notably, the Mean measure only assesses frictional mismatches, while it fails to consider the structural sources of over- and under-education (Verhaest and Omey 2006b). In addition, the Mean measure is less responsive to technological change and changes in workplace organisations than other measures, and so it may lead to erroneous results (Kiker et al. 1997). Also, "one standard deviation away from the mean" implies symmetry between over-education and under-education. To reduce errors from the Mean measure, the mean must be recalculated to take account of changes in the data over time.

The Mode method has a major advantage over the Mean method because it allows an asymmetry between over-education and under-education. In addition, it is more sensitive than the Mean measure to changes in skill requirements due to technological change. That is because the Mean measure is based on a range that can mask sudden changes in skill requirements due to technological change. Using a very simple example, Kiker et al. (1997) proved that the Mode measure is preferable to Verdugo and Verdugo's (1989) Mean criterion. They found that the Mean measure changes gradually and may produce classification errors; whereas, the Mode measure allows abrupt changes to be captured.

Self-Report (SR) Measure

The SR measure is applied in some early studies. In general, there are two types of questions in this measure. The first seeks information on the level of education sought by the employer for a respondent's job, such as: "What kind of education does a person need in order to perform your job?" (Alba-Ramírez 1993; Dolton and Silles 2008). The second type of question is based on the level of education required to get the job, such as "How much formal education do you require to get a job like yours?" (Duncan and Hoffman 1981; Rumberger 1987; Sicherman 1991; Dolton and Vignoles 2000; Dolton and Silles 2008).

These two types of questions reflect different standards. The education required to perform the job is in line with the concept of over-education (referring to the Green et al. (1999) approach) by considering the use of skills to do the job. The level of education required to do the job compared to the level to get the job reflects closely related, but somewhat different measures.

The required years of education to get a job can be affected by skill requirements due to technological change, but also the supply and demand in the labour market. For example, when labour supply grows relative to labour demand, employers are more likely to hire people with higher levels of education, in comparison to a situation in which labour supply is lower. As a result, the apparent education required to get a job goes up, while the education needed to do the job does not change. The level of education required to get a job also puts emphasis on the hiring standard, below which level a company is disinclined to employ job applicants (Verhaest and Omey 2006b; Dolton and Silles 2008). Therefore, measurement criteria of over-education to get a job may also reflect signalling effects for selection based on labour market conditions, in addition to the educational requirements to perform the job.[3]

[3] Signalling effects for selection indicate that workers with higher levels of education are likely to have higher productivity than workers with lower levels of education, based on human capital theory. Therefore, based on signalling hypotheses, workers with higher levels of education are more likely to be hired by employers.

Dolton and Silles (2008) used two self-reported measures that include the qualifications required to get the job and the qualifications needed to do the job. They examined the determinants of over-education and its earnings effects in the U.K. graduate labour market. They found a 25% mean incidence of over-education in the current job, based on both the educational qualification required to get the job and to do the job.

Finally, while the SR measure draws on local, up-to-date information (Hartog 2000), this could be based on the workers' perceptions, state of mind and limited range of experiences, rather than information based on the entire industry. For example, workers might overstate job requirements based on limited information (Kler 2005).

Over-skilling

In addition to over-education as a measure of job mismatch, empirical evidence has shown that over-skilling provides added information on job mismatch (Allen and van der Velden 2001; Green and McIntosh 2007). Some studies have turned to using a combination of over-skilling and over-education to examine the labour market outcomes of job mismatches (e.g., Mavromaras et al. 2013). It is argued that a combination that includes both skill matching and education matching may better control for the effects of unobserved ability, rather than education matching alone. This is because, conceptually, over-skilling may reflect the presence of skills that could easily be related within the employment context (Mavromaras et al. 2007).

Therefore, over-skilling may capture unobserved individual ability, thus providing some additional insight into the quality of match (Mavromaras et al. 2010). Specifically, for some data sets, the skilling variable may directly assess a worker's utilisation of 'ability and skills'. In contrast, over-education measures the deviation between the formal education obtained by the worker and the education required by the employer in order for the worker to perform a job. Consequently, the two measures may be used in combination.

2.2 Impacts of over-education on earnings

This section presents stylised facts based on the literature on the overall relationships between educational mismatches and earnings. The literature on over-education consistently concludes that the returns to over-education are lower, by about half to two-thirds, than the returns to required education, although it is still a positive return on earnings (Cohn 1992; Groot 1996; Rumberger 1987; Sicherman 1991). The returns to years of over-education are positive and less than the returns to required years of education; the returns to years of under-education are negative, but the absolute values are smaller than the returns to years of over-education (Hartog 2000). Therefore, the years of over-education increase earnings, but by less in comparison to the required years of education. In addition, being in a job for which a worker is under-educated benefits the worker compared to a worker with similar years of education, but in a matched occupation. Additional schooling beyond that required for the job is not always rewarded (Rumberger 1987).

In Australia, based on the Negotiating the Life Course Survey and a Self-Report measure of required education, Linsley (2005) found that 27% of individuals were over-educated for the job they were doing. Using data from the 1996 Census of

Population and Housing and a Realised Matches measure, Voon and Miller (2005) reported that about 14-16% of full-time workers aged 20-64 were over-educated. This same study found around 70% of workers were adequately educated. The returns to years of actual education were 9.2% for men and 8.0% for women, and for over-education they were 6.6% and 5.3% for men and women respectively.

The studies cited above used cross-sectional data, which assume homogenous individuals and the random assignment of workers to jobs. As such, adjusting for unobserved heterogeneity of individuals and jobs was not pursued in those studies.

Adjustment for heterogeneity and measurement error

To control for unobserved heterogeneity, panel estimation techniques were applied to longitudinal panel data sets by Bauer (2002) and Tsai (2010) in order to examine the wage effects of educational mismatch. Contrary to previous findings, Bauer (2002) used the German Socio-Economic Panel (GSOEP) data set over the period of 1984-1998 to show that the estimated wage differences between adequately and inadequately educated workers become smaller or disappear totally after controlling for unobserved heterogeneity. Tsai (2010) used a fixed effects model to control for the non-random assignment of workers to jobs, based on data from the United States. Tsai found that over-educated workers do not earn less than other workers with similar years of education, but who are engaged in matched occupations. In contrast, after considering both skill-underutilisation and over-education, Mavromaras et al. (2013) found different estimates of wage penalties due to mismatch between cross-section and panel estimations, suggesting the presence and influence of unobserved heterogeneity.

Only a few longitudinal studies have examined the impact of over-education on earnings by using the fixed effects instrumental variables (FEIV) estimator to correct for both omitted variables bias and measurement error bias. These studies have mixed results. Overall, a more general application of instrumental variables estimators is advocated in over-education studies. The pioneering approach, proposed by Dolton and Silles (2008), uses instrumental variables to address measurement errors associated with the self-reported and retrospective survey data they used. The survey asked about the respondent’s first job and utilised data on the current job. They found that over-education lowers earnings for each year of over-education, as distinct from earnings for each year of required education for the job, by 35-40% and that the downward bias resulting from measurement error is equally balanced out by the upward bias caused by the omitted ability variable.

Using two subjective measures as instruments for a JA measure, Verhaest and Omey (2012) drew a slightly different conclusion based on the data for two cohorts of Flemish respondents born in 1976 and 1978. They found the upward bias resulting from unobserved heterogeneity was smaller than the downward bias resulting from measurement error. In contrast, using the Mean measure as an instrumental variable for the Mode measure, Iriondo and Pérez-Amaral (2016) found that measurement error results in attenuation bias, based on an EU longitudinal database for the period 2006-2009. They further found that the size of the measurement error bias is smaller than the omitted variables bias. Still differently, Sellami et al. (2017) found that the earnings effect differed when using different IV procedures in fixed effect models to address unobserved heterogeneity and measurement error for a sample of Belgian graduates. They recommended that more research on larger datasets is useful in this area.

In this paper, we further extend the literature by providing a systematic comparison of using alternative measures of educational mismatch and IV selections in this context, using a rich longitudinal Australian data set.

3. Data and variables

3.1 Data

The data are taken from the first nine survey years (2001-2009; waves 1-9) of the HILDA Survey, a longitudinal data set. This interview-based survey began in 2001 with a nationally representative sample. Following the respondents for nine years since 2001 allows the use of panel features of the data over a significant period of time. There are 54,595 observations in the full sample, with 11,664 employed individuals. Almost half (48.5%) of the sample are females and 78% are born in Australia. The sample we focus on is an unbalanced panel of full-time employed workers aged from 23 to 64 years old. Self-employed workers and students are excluded in the sample.

A major motivation for using this longitudinal survey is that it overcomes the weaknesses inherent in cross-sectional data. The HILDA Survey follows the same units over time, allowing researchers to analyse the dynamics of change at the individual (and household) level. Specifically, when exploring the effects of over-education on earnings, the application of panel data can identify the extent of the causal effect of unobserved heterogeneity on earnings. A second advantage of the data set is that it contains rich coverage of relevant variables, including detailed information on qualifications, job characteristics, and two-digit occupation information, all of which are important to this analysis. Also, the HILDA Survey re-interview rates are exceptionally high (e.g., 96.3% in wave 9).[4]

The dependent variable in this study is the hourly wage from main job. The key explanatory variables are educational mismatch and required years of education variables. Other variables on individual and job characteristics, as well as relevant controls through conventional fixed effects are employed. The definition of variables is available in Table A1 in the Appendix.

3.2 Measures of required years of education

As discussed in Section 2, the different measures capture different dimensions of over-education. We use Realised Matches (based on both the cross-wave mode (Mode) measure and Verdugo and Verdugo's (1989) Mean measure), and JA to measure the

[4] We further examined the potential impact of selection into full-time employment, and also employment, using Heckman (1979) selection adjustments. We found the coefficients in the earnings models stayed unchanged across the models with and without the selection adjustments. These results alleviated concerns regarding sample selection effects in our preferred fixed effects models. These results are available on request. Since the focus of the analysis is on a comparison of measurement methods, and not policy or time-sensitive periods, we decided to keep the analysis using the 2001-2009 data waves as a sufficiently long period for the current study.

level of required years of education.[5] The Mode and Mean methods measure the difference between the actual years of education in comparison to the required years of education by the employees in a job. As such, these measures are expected to reflect the educational qualifications required to perform a job. But market forces (e.g., labour supply and demand, signalling effects) for getting a job can also affect the required years of education over time. We expect that the market effect is the same regardless of the measurement method used – Mode, Mean, and also, to some extent, JA. The JA method may lag, incorrectly capturing these market-based effects.

Each of the following methods compares the actual years of education by the respondent, for each year of data with the specified measure of required education for that year. This analysis is carried out throughout this study based on the standardised two-digit occupational categories for all four methods.

Mode method

In the Mode method, a person is classified as over-educated if their years of education exceeds the Mode level of education in their occupation.

The Mode measure is defined as the modal education by occupation and year, and derived separately for each of the nine years of data, varying by survey waves (year). First, the actual years of education that most commonly occurs within an occupational category is calculated for each year. Then, the required years of education, for all nine survey waves, are derived based on the Mode for each year. A detailed explanation is: we sort the data by the two-digit Australian and New Zealand Standard Classification of Occupation (ANZSCO 2006) occupations category; then, for each occupation category, we generate the most frequently recorded years of education in 2001, in 2002, and so on, to 2009. Finally, we put these education data together and they are the required years of education from 2001 to 2009. Because the years of required education are generated by waves, we call this the cross-wave Mode measure.[6]

Mean method (Range-One and Range-Half)

Two range measures, based on Verdugo and Verdugo's (1989) RM method, are used in this study. Since, under this measure, required years of education is a range, we call this method a 'range measure'.

The Range-One measure of over-education compares each individual's actual years of education to a mean plus one standard deviation within occupations. Over-education is defined as the actual years of education exceeding one standard deviation above the mean education for the occupation; under-education is defined as the actual years of education one standard deviation below the mean education for

[5] In this study, the HILDA Survey does not include questions such as "What level of education do you need to get your current job?" or "What level of education is required to perform your job?". Thus, we do not have a measurement on "over-education" based on the Self-Report measure. Thus, worker's self-report is not applicable.

[6] Both the Mode and the actual years of education measures are in full years, corresponding to the official degree completion year requirements. For example, if a respondent takes longer to complete a degree, such as 3.5 years to complete a 3-year degree, the recorded completed years of education is 3 years.

the occupation. Similarly, Range-Half measures over-education by replacing one standard deviation in the Range-One measure with one-half of a standard deviation. The mean plus one standard deviation and the mean plus a one-half of a standard deviation are adopted as complementary methods in this study. These measures are derived for each year of data separately, based on the mean value for that year.

Job analysis

A person is classified as over-educated if his or her years of actual education exceed the JA level of education in their occupation. The JA measure is derived based on this comparison, by two-digit occupation classification, utilising the Australian and New Zealand Standard Classifications of Occupations (ANZSCO 2006) and the Australian Qualifications Framework (AQF). The ANZSCO 2006 provides qualifications and skill levels required to work in specific occupations (to perform a job). The HILDA Survey provides a two-digit ANZSCO 2006 occupation category for each employed worker's main job for each wave.

The Australian Qualifications Framework (AQF) is the national policy for regulated qualifications in Australian education and training. Each qualification type is described in terms of its purpose and required knowledge, skills, and volume of learning. We compare each qualification type in the AQF with matched qualifications required in ANZSCO 2006 at the two-digit occupation code level, then convert the qualification to required years of education. Since the performance of the JA method is largely dependent on the relevance of the specific job classification, we adopt the two-digit, rather than one-digit, occupation code level available in the HILDA Survey for greater accuracy of the job description.

Both measures of the incidence and the years (or fractions of years) of over-education and under-education are derived based on the four methods and are used in the study. In Section 4, we report on the incidence of over-education as is commonly described in the literature. In the earnings regression models in Section 5, we employ the years of educational mismatch in keeping with the literature and to use all of the information available on the magnitude of educational mismatch. In Section 6, and for auxiliary models that combine education and skill mismatches, binary measures of both education and skill mismatches are used as required and consistent with the literature.

3.3 Statistical Analysis of comparative measures

Required years of education measures an interval under the Range-One and Range-Half measures rather than a point under the Mode or JA measures. Following alternative measures as discussed, we use pooled data and plot the histogram of actual years of education and its normal density curve in Figure 1.

Both the Mean measure (Range-One and Range-Half) and the Mode measure are derived based on the two-digit ANZSCO 2006 occupations category. The smaller range of the Range-Half measure is expected to perform more closely to the Mode measure than the Range-One measure. In contrast, the JA measure is derived from the occupational classification code and qualification framework. It is updated intermittently for different occupations. Therefore, it is not surprising for it to perform differently from both the Mode and the Mean measures.

Figure 1 shows that the mean years of actual education in our data is 13.8 (at point *c*). The required years of education is 14.01 (at point *d*) for the Mode measure.

Based on the Range-One definition, actual education years between points *a* and *g* represent an education-occupation matched job; workers who have less than point *a* years of education are under-educated and workers who have more than point *g* years of education are over-educated. Figure 1 also shows the narrower boundary years of required education (points b. and f.) for the Range-Half (mean plus and minus one-half of a standard deviation) measure. If workers' years of education fall into this range, those workers are matched. Because of the larger range of the required years of education for the Range-One measure, than for the Range-Half measure, a smaller proportion of mismatched workers are expected to be classified in the Range-One measure.

[Figure 1 here]

4. Incidence of educational mismatch

For simplicity and clarity, the four measures are called Mode, Range-One, Range-Half, and JA. The Spearman correlation test reported in Table 1 across the nine waves of data confirms that there are correlations between the various measures, but it also confirms that some individuals are classified differently under these alternative measures. The correlation ranges from 51.4% (relatively weak correlation) for Mode and Range-One to 77% for Range-Half and JA, which indicates that the same individual is over-educated for one measure, but may be adequately educated for the other measures.

[Table 1 here]

4.1 Educational mismatch by educational qualification

Figure 2 illustrates the average required years of education, aggregated across all occupations, held by those who have a given qualification, using the four alternative measures of specifying the required years of education for each occupation. Table 2 summarises the various incidences of educational mismatch. These analyses show results by educational qualifications, with comparative and variation of results, based on the four methods applied.

Figure 2 and Table 2 show that, as expected, workers with the highest qualifications are more likely to be over-educated and workers with no qualifications are least likely to be over-educated. But, for example, based on the Mode method, 11.5% of individuals with 'no qualification' were also over-educated (Table 2). This result reflects cases when the required years of education in a particular occupation is lower than the actual years of education for a respondent without secondary school completion. (e.g., employees with no qualifications, with 11 years of schooling, are over-educated in occupations with 10 years of required education).

The extent of over-education in general rises from 18.6% for Range-One to 40.3% for JA among bachelor graduates. The actual years of education and required education in Figure 2 explain the differences in results. With a qualification higher than certificate level, the required years of education is highest for Range-One than the other three measures. Thus, a high level of adequately-educated matches is found under Range-One (49-94.6%). Comparing Mode with JA, the required years of education is higher for JA than for Mode, except for the postgraduate group.

[Figure 2 here]

[Table 2 here]

For intermediate levels of qualification, the dependence of the incidence of over-education on the level of qualification is less clear and depends on the method used to define the educational requirements of the job. Using the JA or Range-One method leads to a monotonic pattern, while using the Mode and Range-Half methods gives an incidence of over-education for the advanced diploma group that is substantially greater than that for the bachelor group.

4.2 Educational mismatch by occupation over time

To further investigate the extent of over-education over time under different measures, we report on cross-sectional data with four-year intervals (2001-2005-2009; these are wave1, wave5, and wave9) to summarise the results on the incidence of educational mismatch by occupation over time. We report this analysis of the actual and required years of education based on the four measures investigated by occupation groups in Figure 3. Table 3 shows data related to Figure 3 on the actual and required years of education and over-education and under-education for the first and last year of the data (wave 1 and wave 9) by occupation groups.

[Figure 3 here]

[Table 3 here]

Figure 3 illustrates that for some occupation groups, such as Professionals, the four measures of required education are similar and within a narrow band. In contrast, for other groups, such as Mangers and Machinery Operators and Drivers, the variation based on the four measures is greater. Figure 3 and Table 3 further show that the required years of education based on the JA and Mean measures generally show minor changes over the period of the study.

In addition, the required years of education show more variations for the Mode measure than for the Range-One, Range-Half, and JA methods. This is particularly true in the occupations of Clerical and Administrative, Machinery Operators and Drivers, Community Service workers, Managers, and Labourers, reflecting either technology updates or excess supply that contributes to this change. Therefore, under the Mode measure, large adjustments in the incidence of over-education over time are expected in these five groups. For example, within Machinery Operators and Drivers, the over-education rate was 37% in 2001, but it decreased to 18% in 2005, and then further decreased to 9% in 2009.

As Figure 3 shows, the JA measure of required education changes less than the other measures over the time period, and, for most occupation groups, it differs from the other methods. Figure 3 and Table 3 further show that the average actual years of education are increasing slightly over time in most occupation groups.

4.3 Incidence of educational mismatch (standardised)

The extent of educational mismatch for each of the four methods, with standardised measures across methods, is reported in Table 4. Table 4 confirms that the incidence rate of over-education varies with different measures. For full-time workers, over-education ranges from 13% for the Range-One measure to 34.8% under JA. The percentage of adequately-educated workers is higher under the Range-One measure (71.4%) than under the other three measures. JA reports the lowest rate of adequately matched (23.6%) and the highest rate of under-education (41.6%).

[Table 4 here]

The standardised measures are derived by setting the incidence of educational mismatch as 1 under the Mode measure. Those standardised results confirm that, compared to the Mode method, Range-One categorises both over- and under-education to less than half the incidence in the Mode method.

In general, the JA method reports slightly higher rates of over-education, as our results in Table 4 confirm. This result can be explained because the JA method is prone to periodically ignoring market changes such as technological change. As noted earlier, since the JA method relies on official analyses that are not updated regularly enough, it is susceptible to lagging behind changes in required education in the market for over-education. As Table 4 shows, the JA method also estimates higher rates of under-education compared to the other three measures. JA classifications generally assign higher required education levels for the lower-skilled occupations, such as the assignment of a high-school degree for labourer jobs, as shown in Table 3.

The Mean measure, in particular Range-One, is likely to report a lower rate of over- or under-education due to assigning a wider range to ‘adequate education’.

In our results, Range-Half performs similarly to the Mode analysis for categorising over-education, but it reports under-education at rates that are 20% lower than the Mode method. This effect is consistent with the symmetry assumption of educational job requirements for under- and over-education discussed earlier.

Comparing results from earlier studies, reported in Panel B of Table 4, we found that the incidence of mismatches that include both over-education and under-education in our data under the Range-One measure is quite close to the findings of Voon and Miller (2005) and Tsai (2010). Using a different data set, Voon and Miller (2005) applied the Range-One measure and found that the incidence of over-education is about 16% of male (14% of female) full-time workers aged 20-64.

The Range-One measure in the study originally proposed by Verdugo and Verdugo (1989) examined the incidence of over-education based on the 1980 U.S. Census.

They found that 10.9% of European background males were over-educated and 79.2% were adequately educated.

Tsai (2010) used the PSID data from 1979 to 2005 to examine returns to over-education in the U.S. labour market. She found that 22% of workers are over-educated and 9% of workers are under-educated, based on the Range-One measure.

Combining the incidence rates of over-education and adequate education based on Range-One measures from those studies with different data sets, the analysis shows that the Range-One measure is likely to report the lowest rate of over-education among the four methods, because of the high proportion of adequately educated individuals obtained from this measure, as explained in Figure 1.

In summary, the results from both our study and our standardisation of earlier studies confirm that the classification of the incidence of over-education is lower for the Range-One measure than with the Mode method. This is because the required years of education is a range that includes the point that is defined as the adequate education needed for the Mode measure. In contrast, the rate of over-education is lower for the Mode measure than for the JA measure. The education required based on the JA measure is prone to being lower than what is derived based on the other measures. This is expected when educational requirements increase, coupled with a time lag in the updating process.

Overall, our analysis in this section also provides support for the expectation that among the four measures, the Mode measure presents greater variation in mismatches. In this respect, the Mode overcomes the drawbacks of the range measures by changing more freely. Therefore, the Mode measure is more responsive to reflecting technological change and skill requirements in different occupations, compared to the Mean measure. The Mode measure can also reflect abrupt changes with less delay than the JA and Range-One measures.

5. Econometric framework

In this section we apply the econometric framework on the earning impacts of educational mismatch to our panel data. We compare results across the standard and panel fixed effects models with instrumental variables (FEIV) estimations, and we adjust for heterogeneity and measurement error through IV estimation.

5.1 Model of Over-education, Actual education and Under-education

The general form of the model we apply is based on the recent literature (Iriondo and Pérez-Amaral 2016; Verhaest and Omey 2012)[7], which has modified Verdugo and

[7] This is one of the two prominent approaches. The other approach is the standard ORU (Over-Education, Required Education, and Under-Education) earnings model proposed by Duncan and Hoffman (1981). Rather than controlling for actual years of education, the required years of education are controlled for in the model. In this case, an over-educated worker earns more and an under-educated worker earns less than matched co-workers, holding other characteristics constant. Thus, when this alternative specification is made, the coefficient of over-education is positive, but the coefficient of under-education is negative.

Verdugo's (1989) model. Our model, as in equation (1), incorporates actual years of education and years of over-education and under-education.

The earnings model is

$$ln\, y_{i,t} = \alpha_i + \beta_a S_{i,t}^a + \beta_o S_{i,t}^o + \beta_u S_{i,t}^u + \gamma X_{i,t} + \varepsilon_{i,t} \qquad (1)$$

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

where $lny_{i,t}$ denotes, in year *t*, the natural logarithm of the hourly wage rate from main job of individual *i* , $S_{i,t}^a$ denotes the years of actual education of individual *i,* in year t. $S_{i,t}^r$ is years of required education for individual *i* in year *t.* Thus, $S_{i,t}^o = \left(S_{i,t}^a - S_{i,t}^r\right)$ is years of over-education when $S_{i,t}^a > S_{i,t}^r$; 0, otherwise. In contrast, $S_{i,t}^u = \left(S_{i,t}^r - S_{i,t}^a\right)$ is years of under-education when $S_{i,t}^r > S_{i,t}^a$; 0 otherwise. β_a is the rate of return to actual education.

After controlling for actual education, if a worker is over-educated, he/she earns less than the worker who has equal education and works in matched employment (which would be in a higher-ranked occupation). Similarly, an under-educated worker earns more than the worker who has the same years of actual education and works in a matched job (which would be in a lower-ranked occupation). Therefore, compared to workers who have similar years of actual education and work in education-occupation matched employment, the coefficient of over-education β_o<0 represents the penalty of over-education for each year of over-education. Similarly, the coefficient of under-education β_u>0 implies the bonus for under-education and $X_{i,t}$ represents personal and job characteristics of individual *i* in year *t.* The earnings penalty refers to the percentage reduction in earnings returns for each year of over-education (under-education) in comparison to a worker with similar actual education who is employed in a matched job. Personal characteristics include a married status, presence of children, qualification, state, and country of birth. Job characteristics include current occupation job tenure and industry fixed effects. α_i denotes the unobservable individual-specific effect and $\varepsilon_{i,t}$ denotes the remainder disturbance assumed to be independent and identically distributed i.i.d (0, σ_ε^2).

5.2 Unobserved heterogeneity and measurement error

There are two concerns involved in over-education studies that we address here. The first is individual unobserved heterogeneity, which would bias the results upwards (Sloane 2003). For example, a worker with lower personal ability or quality of education might substitute more education to perform the job equally well as the worker with high ability but less education, or less education of higher quality. Thus, the substitution between components of human capital would overestimate the impacts of over-education. Therefore, in those cases one cannot conclude that a worker with surplus schooling for the job held is over-educated if it is accompanied by lower ability.

We use panel fixed effects estimation to control for the individual's unobserved heterogeneity, as in equation (2), by estimating the deviation from the mean. The notation of variables is similar to equation (1).

$$ln\, y_{i,t} - \overline{ln\, y_i} = \beta_a(S_{i,t}^a - \overline{S_i^a}) + \beta_o(S_{i,t}^o - \overline{S_i^o}) + \beta_u(S_{i,t}^u - \overline{S_i^u}) + \gamma(X_{i,t} - \overline{X_i}) + (\varepsilon_{i,t} - \overline{\varepsilon_i}) \qquad (2)$$

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

The second concern is that, when evaluating the educational qualifications that are necessary to meet the requirements of an employment position, avoiding or lessening measurement error is an important issue. The validity and credibility of empirical results decrease with increasing measurement error. To address measurement error, Dolton and Silles (2008) proposed the method of instrumental variables, using an alternative measure of over-education as an instrument. This method has been applied in recent studies by Iriondo and Pérez-Amaral (2016) and Verhaest and Omey (2012). We evaluate the earnings impact under alternative measures. For each measure, we use three alternative measures as instrumental variables.

5.3 Over-education and Over-skilling

To overcome the failure of controlling for unobserved ability from the measure of over-education, researchers have started to incorporate an over-skilling variable into their models. We also incorporate over-skilling[8] into our auxiliary models, following the approach used by Mavromaras et al. (2013). Over-educated and over-skilled workers are those who work in positions where both education and skill requirements are lower than the workers' actual educational attainment and skill acquisition. According to the definition of required education (Mode), and considering both education mismatch and skill mismatch, the extended model (equation (3)), includes five dummy variable categories that represent the entire sample's positioning within six mismatched groups.

$$\ln y_{i,t} = \alpha_i + \gamma_a S^a_{i,t} + \gamma_{oe,os} OEOS_{i,t} + \gamma_{oe} OE_{i,t} + \gamma_{ue,os} UEOS_{i,t} + \gamma_{os} OS_{i,t} + \gamma_{ue} UE_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \quad (3)$$
$$i = 1, \cdots, N; t = 1, \cdots, T$$

Where $OEOS_{i,t}$, $OE_{i,t}$, $UEOS_{i,t}$, $OS_{i,t}$ and $UE_{i,t}$ respectively represent worker *i*'s job mismatch status (binary variables): over-educated and over-skilled, over-educated, under-educated and over-skilled, over-skilled, and under-educated. These binary variables take the value of 1 if worker *i* falls into the corresponding category at year *t*, and a value of 0, otherwise. The coefficients of these mismatch variables examine relative earnings effects resulting from mismatch to matched positions. The rest of the variables are defined the same as in equation (1).

6. Results

The results confirm significant and positive return to years of education (about 4-9% per year of matched education across fixed effects and models with IV adjustments), highlighting the importance of the link between higher education and earnings in

[8] In detail, over-skilling is derived from the HILDA Survey by using the responses, scored on a seven-point scale, to the question "I use many of my skills and abilities in my current job", with a response of 1 corresponding with strongly disagree and 7 with strongly agree. Individuals with responses of 1, 2, 3 or 4 on the scale were classified as over-skilled and those with responses of 5, 6 or 7 as skill-matched. Sensitivity tests confirm the cut-off points for over-skilled and skill-matched are appropriate.

general. The results further provide evidence on stylised variations in the earnings returns to years of over- and under-education based on alternative measures used across OLS, and fixed effects instrumental variable (FEIV) estimations, and compared to matched years of education.

Table 5 presents the estimation results of equations (1) and (2) under alternative measures of educational mismatches for the full-time sample. For each measure, column 1 reports OLS estimates as a benchmark. Column 2 reports panel fixed effects estimates. Columns 3 to 5 report the FEIV estimates.

[Table 5 here]

6.1 Pooled ordinary least square (OLS) regressions

The OLS estimation results for full-time workers presented in Table 5 are generally consistent with the stylised facts reported in the literature (e.g., Sicherman 1991). Over-educated workers earn less than adequately educated workers who have the same years of education and work in a matched employment, and the opposite applies to the under-educated workers. This is evident by the negative sign on coefficients of over-education and the positive sign on coefficients of under-education. Those estimates are statistically significant.

Our analysis highlights that the magnitude of coefficients of over-education, actual education, and under-education systematically varies with Mode, Range-One, Range-Half, and JA. For instance, based on OLS estimates reported in column 1 of Table 5, holding other variables constant, compared to their counterparts, over-educated workers suffer a penalty ranging from 1.9% under the Mode measure to 9.4% under the Range-Half measure for each year of over-education. In contrast, under-educated workers gain a bonus from 2.3%, compared with workers with similar years of education but in matched jobs, under the Range-One measure to 6.6% under the JA measure.

6.2 Panel fixed effects and random effects

Fixed effects and random effects models are applied to evaluate the impact of unobserved heterogeneity (e.g., ability, motivation) on earnings. The Hausman test results reject the null hypothesis that individual-specific error is uncorrelated with the explanatory variables of the wage equation. Therefore, fixed effects estimates are preferred to random effects estimates and are the results reported here.

A comparison of the results in columns 1 and 2 in Table 5 supports the hypothesis that individual heterogeneity plays an important role in the return to over-education. For example, the top-left quadrant of Table 5 shows the earnings impact using the Mode measure. In a comparison of OLS estimates (column 1, Table 5) with fixed effects estimates (column 2), the magnitude of over-education effects drops dramatically from -1.9% to -0.6%. Similar results are found in the returns to actual education, which change from 5.8% to 3.1%. Likewise, the magnitude of under-education effect is 2.6% in OLS estimation, which becomes smaller (0.7%) in fixed effect estimation. The reduced size of coefficients of over-education, actual education, and under-education from OLS estimation to fixed effects estimation is also found under the Range-One, Range-Half, and JA measures (quadrants 2, 3, and 4). A large decrease in the absolute value of the coefficients of over-education is also found by Tsai (2010), Dolton and Silles (2008), Verhaest and Omey (2012), and Iriondo and Pérez-Amaral (2016), suggesting that the fixed effects models are controlling for positive bias.

6.3 Panel fixed effects with instrumental variables

Measurement error, in turn, posits a downward bias, and it is commonly corrected by using instrumental variables (e.g., Dolton and Silles 2008; Verhaest and Omey 2012; Iriondo and Pérez-Amaral 2016). We incorporate the instrumental variables method, as in the above recent studies, by using an alternative measure of over-educated (or under-educated) as instrumental variables. For example, when we use the Mode measure to define over-education, in column 3 of the top-left quadrant of Table 5 (quadrant 1), over-educated (Mode) is instrumented using over-educated (Range-One); thus this column is labelled (FEIV IV Range-One). In column 4, over-educated (Mode) is instrumented using over-educated (Range-Half); thus the column is labelled (FEIV IV Range-Half). Then we follow the same procedure for estimates in columns 4 and 5, and also under other measures. We derive earning penalties and earnings returns for each year of over-education and under-education, compared to earnings for workers with similar years of education, but in matched occupations. The ancillary under-identification test, weak identification test, and Sargan over-identification test results show that these instrumental variables satisfy both reliability and validity.

The size of the FEIV earnings penalty estimates of over-education is larger than that of OLS estimates under the Mode and Range-one measures (quadrants 1 and 2, Table 5), indicating the downward bias caused by measurement error is larger than the upward bias resulting from unobserved heterogeneity. This evidence is different from Dolton and Silles (2008), who found that both unobserved heterogeneity and measurement error effects more or less balance out. Our result is in line with findings in Verhaest and Omey (2012), who found that the upward bias in the penalty of over-education resulting from unobserved heterogeneity was smaller than the downward bias resulting from measurement error, for a sample of young Flemish graduates. Our study covers a wider Australian sample across education levels, and utilising a larger number of instrumental variable alternatives. As Table 5, quadrants 3-5 show our findings also differ from Iriondo and Pérez-Amaral's (2016) results, who find that years of under-education are rewarded similarly to matched years of education, for a sample of European workers. We find that years of under-education are rewarded positively, but at earnings rates that are less than a quarter of matched years of education.

A comparison of our results across the four quadrants of Table 5 (columns 3 and 4) confirms that the net effect of upward or downward bias depends on which instrumental variable is selected for the Range-Half and JA measures. Notably, under the Mode measure, the magnitudes of earnings penalty coefficients of over-education and under-education, as well as the coefficient of actual education, are stable across columns 3 to 5 in the top-left quadrant of Table 5, where different instruments are applied in FEIV estimations. This result indicates more reliability and less impact from measurement error when the Mode measure is applied.

We find that under the Mode measure of over-education (with Range-Half as the IV), the earnings returns to over-education are positive, but with an earnings penalty of 2.5% for each year of over-education. The comparison group is employees with similar years of education, but employed in different and matched occupations. Iriondo and Pérez-Amaral (2016) estimated a similar penalty of 2.5% and Verhaest and Omey (2012) found a slightly larger penalty of 3.0%. The results further show that the magnitude of the earnings penalty for over-education is especially volatile under the Range-One and Range-Half measures in FEIV estimations (e.g., negative earnings returns to the years of over-education, and in comparison to employees with similar years of education, but in different and matched occupations).

Finally, across all measures we find positive and significant earnings returns to actual years of education that match the required years of education in the occupation.

6.4 Over-education and Over-skilling

As discussed earlier, over-skilling can add further information for capturing unobserved individual ability, and additional insight into the quality of match (Mavromaras et al. 2010). Based on the individual's responses to the question "I use many of my skills and abilities in my current job", the derived binary variable 'over-skilled' directly evaluates if the worker utilises his skills and abilities to perform his job. In an extended version of our model in auxiliary estimations, we incorporate both skill and educational mismatches into our model.

Table 6 presents the impact of over-education and over-skilling on earnings, based on equation (3) and under the Mode measure. The results confirm that an over-educated and over-skilled worker reduces his earnings by a larger percentage of 8.1% compared with well-matched workers based on the OLS estimation, compared to an individual with similar years of education, but in a matched education and self-reported skill job. However, after correcting for omitted variables bias using fixed effects estimation, the penalty of being over-educated and over-skilled is reduced to 2.6%. Using instruments for over-education in FEIV estimates in columns 3 and 4, we find a larger penalty for over-educated and over-skilled workers, earning 9.7% to 11% less than well-matched workers after addressing measurement error. This evidence, compared to OLS and FE results, implies that the upward bias in the estimated penalty for over-education and over-skilling resulting from unobserved heterogeneity is smaller than the downward bias resulting from measurement error.

[Table 6 here]

7. Conclusion

In this study, we have used panel data to evaluate four methods that are conventionally used to measure the incidence and the earnings effects of educational mismatches. In our analysis we include the Job Analysis measure, which is less frequently available in studies. We have also provided systematised estimates of earnings penalties to over-education based on these four methods, utilising FE and FEIV estimations that allow adjustments for individual heterogeneity and measurement error.

We have made comparisons and noted the advantages and disadvantages of each of the four conventional methods based on the Mode, Mean (Range-One and Range-Half), and Job Analysis measures. For example, according to the Mode measure, we find a significant incidence rate (32.3%) of over-education. The comparative and standardised incidence rates obtained across the conventional measures we considered show a range of estimates that are in the expected direction and margins of variation as we hypothesise, in comparison to the Mode method. The Range-One method consistently reports a higher percentage of adequately-educated workers than under the other three measures due to its wide range of required years of education. On the other hand, the JA and the Mean methods tend to produce higher rates of over-education due to underlying measurement issues embedded in these measures, that lead to these comparative outcomes.

On earnings returns to years of over-education, we provide comparative results based on OLS (baseline), FE, and our preferred FEIV estimations that account for individual heterogeneity and measurement error. Using FEIV for the Mode measure, and Range-Half as the IV, we estimate an earnings penalty for each year of over-education to be 2.5%. The comparative estimates are 0.6% in fixed effects and 1.9% in OLS estimations. The results confirm that after controlling for these two potential errors, there is a remaining and larger earnings penalty for over-educated workers, in comparison to those with similar actual years of education who are in matched jobs. Nevertheless, our analysis with alternative mismatch measures highlights that the chosen measure of educational mismatch used has a potentially important effect on the estimated magnitudes of outcomes of both the incidence and the earning penalty of educational mismatch.

We find that the measured impact of over-education on earnings is also sensitive to which educational mismatch measure we select as an instrumental variable in this context. As such, our results point to recognising such differences, and to the advantages of applying more than a single measure in this context for a range of expected outcomes. Among the measures used, and based on these comparisons and superior FEIV estimation, we find that the Mode measure provides the most stable results across IV selections in FEIV estimations. These results support the application of the Mode measure in the over-education studies because of its lower impact from measurement error.

We extend our models by incorporating over-skilling conditions, in addition to the over-education measure, and we find larger earnings penalty effects for workers with combined over-educated and over-skilled conditions. Our analysis further highlights that over-educated and over-skilled workers are the group who experience the most significant earnings penalties from the mismatched positions, of the magnitude of 9.7-11% less than well-matched workers in FEIV estimations.

The analyses in this paper show the importance of scrutinising the impact of the chosen measure of educational mismatch and the estimation methods. This is especially important in interpreting the results of the studies in the literature based on the different measurement methods used.

Finally, the results on the significant and positive earnings return to the matched component of actual education, highlight the importance of matched positions in the link between higher education and economic performance. From a policy point of view, despite the mismatch measure used, our results point to significant education-occupation mismatch. The earnings penalty effects of educational mismatch also remain significant after adjusting for individual heterogeneity and measurement error. These results highlight the importance of policy instruments that aim at improving human capital allocation efficiency by keeping the education system up-to-date with employment needs and the pace of technological advancement in the workspace. Other such policies could aim at increasing information transparency between job markets and skill training opportunities to bridge the gap between workers' attainment and changing education requirements in the economy.

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 either FaHCSIA or the Melbourne Institute.

Appendix

Table A1 Definition of Variables

Personal Characteristics	
General	
Married	1 if married (or de facto), zero otherwise.
Without children	1 if does not have a child, zero otherwise.
Good health	1 if in good health, zero otherwise.
Female	1 if female, zero otherwise.
Qualifications	
Postgraduate	1 if highest qualification is doctorate, masters, grad Diploma, grad Certificate or Bachelor with honours; this requires 17 or more years of education, zero otherwise.
Bachelor	1 if highest qualification is Bachelor without honours; this requires 16 years of education zero otherwise.
Diploma	1 if highest qualification is Advanced Diploma or Diploma; this requires 15 years of education zero otherwise.
Certificate	1 if highest qualification is certificate I, II, III or IV; this requires 13-14 years of education, zero otherwise.
No qualification	1 if highest qualification is year12 or below, zero otherwise.
	Notes: Twenty-one years of education is required for a doctoral degree, 18 years for a master's degree, 17 years for a bachelor's honours degree, graduate diploma or graduate certificate, 16 years for a bachelor's degree without honours, 15 years for an advanced diploma or diploma, 14 years for certificate III or IV, and 13 years for certificate I or II.
State	Dummy variables *(NSW, VIC, QLD, SA, WA, TAS, NT, ACT)*
Urban	Dummy variable if urban, zero otherwise.
Country of birth	
Australian	1 if born in Australia, zero otherwise.
ESB	1 if born in an English-speaking country, zero otherwise.
NESB	1 if born in a non-English-speaking country, zero otherwise.
Job Characteristics	
Occupation Tenure	Years of tenure in the current occupation.
Occupation Tenure squared	Occupation Tenure * Occupation Tenure
Job Tenure	Years of tenure in the current job.
Hours of work	Hours per week usually worked in main job.
Weekly wage	Weekly gross wages and salary from current main job.
Hourly wage	Hourly wage from main job in 2009 dollars.
Log hourly wage	Expressed in the natural logarithm of hourly wage from main job in 2009 $.
Union membership	1 if a union member, zero otherwise.
Occupation	Dummy variables derived from the Australian and New Zealand Standard Classification of Occupation (ANZSCO).
Industry Sectors	Dummy variables derived based on the Australian and New Zealand Standard Industrial Classification (ANZSIC 2006 division).
Educational Mismatched Variables	
Actual years of education	Years of educational attainment
Years of required education	Years of adequate education
Over-educated	1 if over-educated, zero otherwise.
Under-educated	1 if under-educated, zero otherwise.
Adequately educated	1 if adequately educated, zero otherwise.
Years of Over-education (S^o)	Years of over-education

Years of Under-education (S^u)	Years of under-education

Education and Skill Mismatch Variables

Binary variables for:

Over-educated and Over-skilled	1 if Over-educated and Over-skilled, zero otherwise
Over-educated and Skill matched	1 if Over-educated and Skill-matched, zero otherwise
Under-educated and Over-skilled	1 if Under-educated and Over-skilled, zero otherwise
Well matched	1 if Education-matched and Skill-matched, zero otherwise
Only Over-skilled	1 if Education-matched and Over-skilled, zero otherwise
Only Under-educated	1 if Under-educated and Skill-matched, zero otherwise

References

Alba-Ramírez, A. (1993). Mismatch in the Spanish labor market: Overeducation? *The Journal of Human Resources*, 28(2): 259–278.

Allen, J., & van der Velden, R. (2001). Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 53(3): 434-452.

ASB (Australian Bureau of Statistics). 2016. https://www.abs.gov.au/statistics/economy/government/government-finance-statistics-education-australia/latest-release#data-download

Bauer, T. K. (2002). Educational mismatch and wages: A panel analysis. *Economics of Education Review,* 21(3): 221–229.

Becker, G. S. (1964). *Human Capital*. New York: National Bureau of Economic Research.

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

Cohn, E., & Khan, S. P. (1995). The wage effects of overschooling revisited. *Labour Economics,* 2(1): 67–76.

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

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

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

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

Green, F., McIntosh S., & Vignoles, A. (1999). 'Overeducation' and skills: Clarifying the concepts. Centre for Economic Performance. *Discussion Paper No. 435*, London School of Economics and Political Science.

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

Halaby, C. N. (1994). Overeducation and skill mismatch. *Sociology of Education,* 67(1): 47–59.

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

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47 (1):153-161.

Iriondo, I., & Pérez-Amaral, T. (2016). The effect of educational mismatch on wages in Europe. *Journal of Policy Modeling*, 38(2): 304-323.

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

Kler, P. (2005). "Graduate overeducation in Australia: A comparison of the Mean and Objective Methods. *Education Economics*, 13(1), pp. 47–72.

Leuven, E. & Oosterbeek, H. (2011). Overeducation and mismatch in the labour market", *IZA Discussion Papers*, 5523, Institute for the Study of Labor (IZA).

Linsley, I. (2005). "Overeducation in the Australian labour market: Its incidence and effects. The University of Melbourne Department of Economics Research Paper. No.939.

Mavromaras, K., McGuinness, S., Nigel, O'L., Sloane, P., & Wei, Z. (2013). Job mismatches and labour market outcomes: Panel evidence on university graduates. *Economic Record*, 89(286): 382-395.

Mavromaras, K., McGuinness, S., Nigel, O'L., Sloane, P., & Fok, Y. (2010). The problem of overskilling in Australia and Britain. The Manchester School, 78(3): 219-241

Mavromaras, K., McGuinness, S., & Wooden, M. (2007). Overskilling in the Australian Labour Market. *Australian Economic Review*, 40(3): 307-312.

Mavromaras, K., Sloane, P., & Wei, Z. (2013). 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.

OECD (Organization for Economic Co-operation and Development). 2016. https://stats.oecd.org/Index.aspx?DataSetCode=MISMATCH#

Robst, J. (1994). Measurement error and the returns to excess schooling. *Applied Economics Letters,* 1(9): 142–144.

Rosen, S. (1972). Learning and Experience in the Labour Market. *Journal of Human Resources, 7*(3), 326-342.

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.

Sellami, S., Verhaest, D., Nonneman, W., & Van Trier, W. (2017). The Impact of Educational Mismatches on Wages: The Influence of Measurement Error and Unobserved Heterogeneity, *The B.E. Journal of Economic Analysis & Policy*, 17(1), 20160055.

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

Sloane, P. (2003). Much ado about nothing? What does the over-education literature really tell us? in Buchel, F., de Grip, A. and Mertens, A. (eds), Over-education in Europe, Current Issues in Theory and Policy, Cheltenham, Edward Elgar, 11–45.

Thurow, L.C. (1975). *Generating Inequality*. New York: Basic Books.

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

Tsang, Mun C., Rumberger, Russell W., & Levin, Henry M. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations, 30*(2), 209.

Van der Meer, P.H. (2006). The validity of two education requirement measures. *Economics of Education Review,* 25: 211-219.

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

Verdugo, R. R. & Verdugo, N. T. (1992). Surplus schooling and earnings. *Journal of Human Resources,* 27(4): 690-695.

Verhaest, D. & Omey, E. (2006a). Discriminating between alternative measures of over-education. *Applied Economics*, 38(18): 2113–2120.

Verhaest, D. & Omey, E. (2006b). Measuring the incidence of over- and undereducation. *Quality & Quantity,* 40(5): 783–803.

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

Verhaest, D., & Omey, E. (2012). Overeducation, undereducation and earnings: further evidence on the importance of ability and measurement error bias. *Journal of Labor Research*, 33(1): 76-90.

Voon, D. & Miller, Pa. W. (2005). Undereducation and overeducation in the Australian labour market. *Economic Record,* 81(255): 22–33.

Wen, L. & Maani, S. A. (2018). A panel study of immigrants' overeducation and earnings in Australia. *International Migration,* 56(2): 177-200.

Wen, L. & Maani, S. A. (2019). Job mismatches and career mobility. *Applied Economics,* 51 (10): 1010-1024.

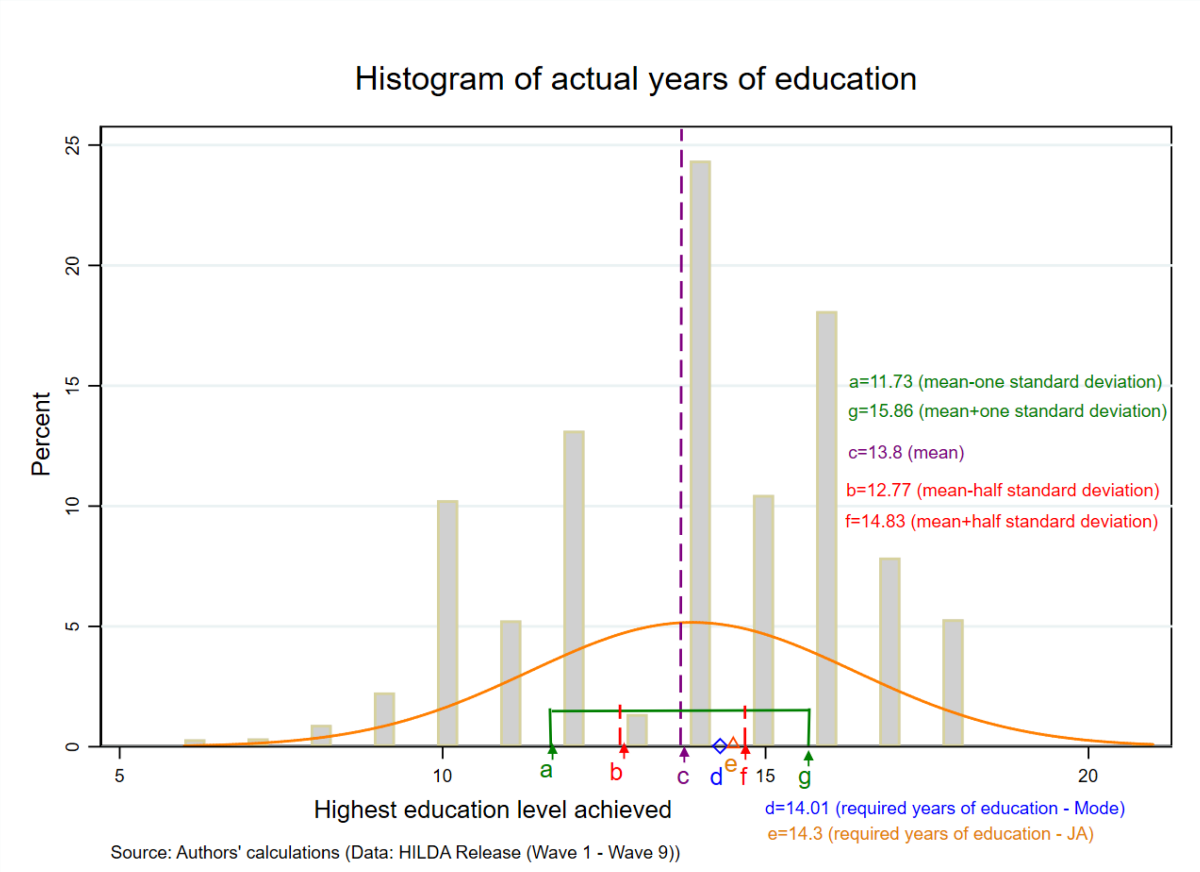

Figure 1. Histogram of actual years of education.
Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

Table 1. Correlations between educational mismatch measures.

	Alternative Over-education Measures			
Rho Significance level (P)	Mode	Range-One	Range-Half	JA
Mode	1.000			
Range-One	0.514 (0.000)	1.000 (0.000)		
Range-Half	0.703 (0.000)	0.550 (0.000)	1.000 (0.000)	
JA	0.653 (0.000)	0.519 (0.000)	0.770 (0.000)	1.000

Notes: Cross-wave correlation measures.

Numbers without brackets are correlations for full-time employed workers aged 23 to 64 years old;

Numbers in () are significance levels.

Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

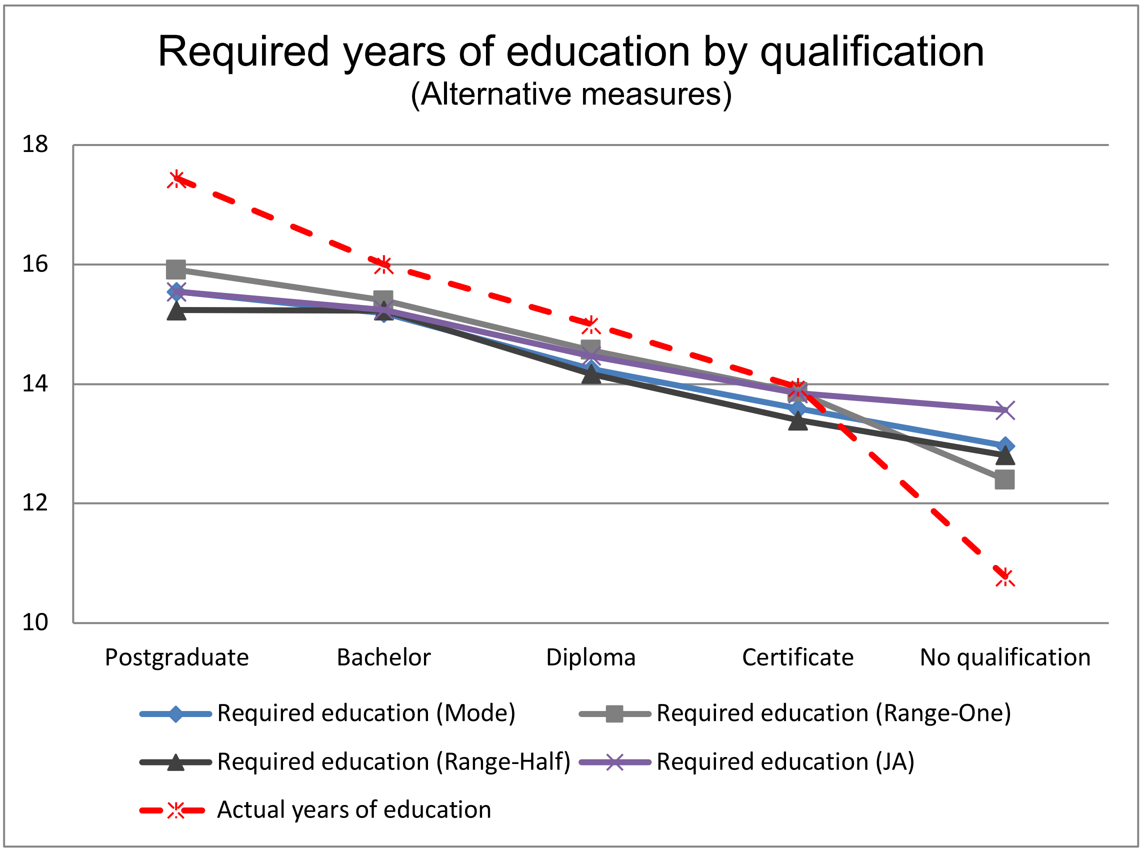

Figure 2. Required years of education by qualification (alternative measures).

Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

Table 2. Fractions with educational mismatches by qualifications (alternative measures).

Mean
(Standard deviation)

	Postgraduate	Bachelor	Diploma	Certificate	No qualification
			Mode		
Over-educated	0.971	0.253	0.564	0.203	0.115
	(0.168)	(0.435)	(0.496)	(0.403)	(0.319)
Under-educated	0.000	0.018	0.427	0.175	0.729
	(0.000)	(0.132)	(0.495)	(0.380)	(0.445)
Adequately educated	0.029	0.729	0.009	0.622	0.157
	(0.168)	(0.445)	(0.096)	(0.485)	(0.363)
Required education	15.548	15.186	14.252	13.586	12.968
	(1.434)	(1.720)	(2.046)	(1.856)	(2.268)
			Range-One		
Over-educated	0.510	0.186	0.161	0.044	0.000
	(0.500)	(0.389)	(0.367)	(0.206)	(0.000)
Under-educated	0.000	0.000	0.000	0.009	0.476
	(0.000)	(0.000)	(0.000)	(0.096)	(0.499)
Adequately educated	0.490	0.814	0.839	0.946	0.524
	(0.500)	(0.389)	(0.367)	(0.225)	(0.499)
Required education	15.913	15.400	14.569	13.858	12.399
	(1.360)	(1.275)	(1.004)	(0.589)	(1.541)
			Range-Half		
Over-educated	0.999	0.276	0.453	0.387	0.000
	(0.033)	(0.447)	(0.498)	(0.487)	(0.000)
Under-educated	0.000	0.000	0.111	0.078	0.719
	(0.000)	(0.000)	(0.314)	(0.268)	(0.450)
Adequately educated	0.001	0.724	0.436	0.535	0.281
	(0.033)	(0.447)	(0.496)	(0.499)	(0.450)
Required education	15.239	15.229	14.169	13.395	12.808
	(0.988)	(1.312)	(1.289)	(1.077)	(1.235)
			JA		
Over-educated	1.000	0.403	0.407	0.389	0.000
	(0.000)	(0.490)	(0.491)	(0.487)	(0.000)
Under-educated	0.000	0.000	0.353	0.323	0.912
	(0.000)	(0.000)	(0.478)	(0.467)	(0.283)
Adequately educated	0.000	0.597	0.240	0.289	0.088
	(0.000)	(0.490)	(0.427)	(0.453)	(0.283)
Required education	15.546	15.243	14.469	13.842	13.567
	(0.919)	(1.180)	(1.489)	(1.286)	(1.378)
Observations	5,346	7,273	4,203	10,330	13,031

Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

Note: Required education is measured in years of education.

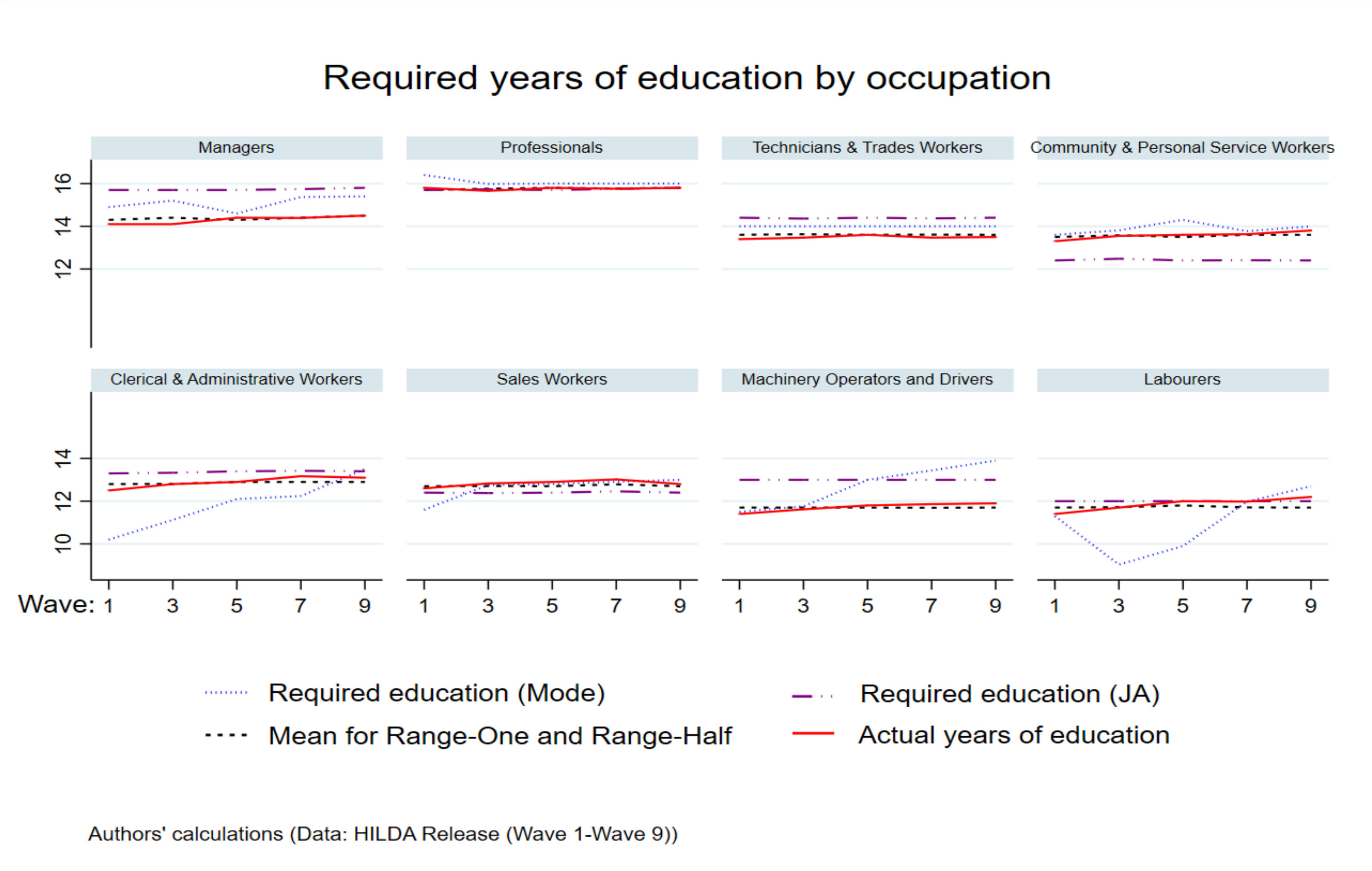

Figure 3. Required years of education by occupation (alternative measures).
Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

Table 3. Educational mismatches by occupation and survey wave

(Incidence sample fractions and years of actual and required education with alternative measures).

	Managers		Professionals		Technicians		Community Service workers		Clerical and Administrative		Sales workers		Operators Drivers		Labourers	
Variables	W1^	W9¬	W1	W9	W1	W9	W1	W9	W1	W9	W1	W9	W1	W9	W1	W9
Actual years of education	14.1	14.5	15.8	15.8	13.4	13.5	13.3	13.8	12.5	13.1	12.6	12.8	11.4	11.9	11.4	12.2
Mode																
Over-educated	0.31	0.27	0.26	0.31	0.15	0.16	0.36	0.35	0.62	0.31	0.53	0.36	0.37	0.09	0.39	0.27
Under-educated	0.46	0.48	0.41	0.27	0.25	0.26	0.35	0.26	0.16	0.43	0.24	0.39	0.35	0.60	0.35	0.44
Adequately educated	0.23	0.25	0.34	0.42	0.61	0.59	0.29	0.39	0.22	0.26	0.23	0.26	0.28	0.31	0.27	0.29
Required years of education	14.9	15.4	16.4	16.0	14.0	14.0	13.6	14.0	10.2	13.5	11.6	13.0	11.5	13.9	11.3	12.7
Range-One																
Over-educated	0.10	0.15	0.13	0.14	0.04	0.05	0.09	0.12	0.16	0.16	0.18	0.17	0.10	0.14	0.17	0.21
Under-educated	0.23	0.18	0.11	0.11	0.16	0.16	0.18	0.11	0.27	0.17	0.28	0.21	0.16	0.10	0.18	0.07
Adequately educated	0.67	0.67	0.76	0.76	0.80	0.79	0.74	0.76	0.58	0.67	0.53	0.62	0.74	0.76	0.65	0.72
Required years of education	14.6	14.7	15.9	15.9	13.8	13.8	13.6	13.8	12.8	13.0	12.8	12.8	11.7	11.8	11.6	11.8
Range-Half																
Over-educated	0.29	0.28	0.38	0.33	0.29	0.28	0.30	0.35	0.30	0.38	0.35	0.37	0.33	0.40	0.33	0.42
Under-educated	0.27	0.21	0.22	0.21	0.21	0.21	0.31	0.20	0.38	0.28	0.44	0.35	0.43	0.35	0.43	0.26
Adequately educated	0.44	0.51	0.40	0.47	0.51	0.50	0.39	0.46	0.32	0.34	0.21	0.28	0.24	0.25	0.24	0.33
Required years of education	14.3	14.5	15.7	15.8	13.6	13.6	13.5	13.6	12.8	12.9	12.7	12.7	11.7	11.7	11.7	11.7
JA																
Over-educated	0.22	0.24	0.45	0.41	0.09	0.10	0.53	0.67	0.33	0.43	0.44	0.49	0.30	0.39	0.33	0.43
Under-educated	0.61	0.57	0.27	0.24	0.42	0.42	0.22	0.14	0.61	0.51	0.44	0.35	0.67	0.60	0.52	0.39
Adequately educated	0.18	0.19	0.28	0.35	0.50	0.48	0.24	0.19	0.07	0.06	0.12	0.16	0.04	0.02	0.14	0.18
Required years of education	15.7	15.8	15.7	15.8	14.4	14.4	12.4	12.4	13.3	13.4	12.4	12.4	13.0	13.0	12.0	12.0
Observations	764	744	1267	1306	731	693	279	318	694	668	234	246	373	348	412	296

Notes: ^ Wave 1; ¬ Wave 9; Over-educated, Under-educated, and Adequately educated are reported in fraction of the sample. Source: Authors' calculations (Data: HILDA Release 9 (Wave 1- Wave 9)).

Table 4. Incidence of over-education.

	Alternative Over-education Measures Mean (Standard deviation)			
Panel A – This Study	Mode[1]	Range-One[2]	Range-Half[3]	JA[4]
Over-educated	0.323 (0.468)	0.130 (0.336)	0.330 (0.470)	0.348 (0.476)
Under-educated	0.329 (0.470)	0.157 (0.364)	0.265 (0.441)	0.416 (0.493)
Adequately educated	0.347 (0.476)	0.714 (0.452)	0.406 (0.491)	0.236 (0.425)
Standardised analysis by setting incidence of educational mismatch under Mode measure as 1:				
Over-educated	1	0.402	1.022	1.077
Under-educated	1	0.477	0.805	1.264
Panel B – Comparison – Existing studies	Mode[1]	Range-One[2]	Range-Half[3]	JA[4]
Kiker et al. (1997) – the Portugal labour market				
Over-educated	0.26	0.09	NA	0.33
Under-educated	0.17	0.05	NA	0.38
Standardised analysis				
Over-educated	1	0.246	NA	1.461
Under-educated	1	0.294	NA	2.235
Verhaest and Omey (2010) – the Flemish labour market				
Over-educated (adjusted measures)	0.338	NA[5]	NA	0.491
Standardised analysis				
Over-educated	1	NA	NA	1.453
Tsai (2010) - the U.S. labour market				
Over-educated	0.33	0.22	NA	NA
Under-educated	0.22	0.09	NA	NA
Standardised analysis				
Over-educated	1	0.667	NA	NA
Under-educated	1	0.409	NA	NA

Notes:
[1] Cross wave mode measure;
[2] Mean plus one standard deviation measure;
[3] Mean plus half standard deviation measure;
[4] Job Analysis measure;
[5] NA indicates not applicable;

Sample: The analysis is based on the full-time employed sample with 40,183 observations.
Data: HILDA Release 9 (Wave 1- Wave 9).

Table 5. Earnings return to over-education - Alternative Over-education Measures

	Mode					Range-One				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	FE	FEIV IV Range-One	FEIV IV Range-Half	FEIV IV JA	OLS	FE	FEIV IV Mode	FEIV IV Range-Half	FEIV IV JA
Over-education	-0.019***	-0.006***	-0.031***	-0.025***	-0.028***	-0.056***	-0.040***	-0.153***	-0.059***	-0.094***
	[0.002]	[0.001]	[0.006]	[0.004]	[0.005]	[0.007]	[0.007]	[0.032]	[0.009]	[0.014]
Under-education	0.026***	0.007***	0.008***	0.007***	0.007***	0.023***	0.012***	0.020***	0.013***	0.016***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.003]	[0.002]	[0.002]
Actual Education	0.058***	0.031***	0.041***	0.039***	0.040***	0.058***	0.040***	0.063***	0.044***	0.051***
	[0.002]	[0.004]	[0.004]	[0.004]	[0.004]	[0.002]	[0.004]	[0.008]	[0.004]	[0.005]
Constant	2.117***	2.720***	2.619***	2.644***	2.631***	2.108***	2.600***	2.308***	2.552***	2.460***
	[0.031]	[0.056]	[0.061]	[0.059]	[0.059]	[0.030]	[0.060]	[0.101]	[0.062]	[0.068]
R^2	0.256	0.170	0.163	0.167	0.165	0.256	0.173	0.161	0.172	0.170

	Range-Half					JA				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	FE	FEIV IV Mode	FEIV IV Range-One	FE IV IVJA	OLS	FE	FEIV IV Mode	FEIV IV Range-One	FEIV IV Range-Half
Over-education	-0.094***	-0.040***	-0.090***	-0.042***	-0.055***	-0.072***	-0.030***	-0.090***	-0.057***	-0.044***
	[0.005]	[0.006]	[0.019]	[0.007]	[0.009]	[0.005]	[0.005]	[0.021]	[0.011]	[0.008]
Under-education	0.065***	0.022***	0.036***	0.022***	0.026***	0.066***	0.020***	0.054***	0.035***	0.028***
	[0.003]	[0.003]	[0.006]	[0.004]	[0.004]	[0.004]	[0.005]	[0.013]	[0.007]	[0.006]
Actual Education	0.103***	0.051***	0.074***	0.052***	0.058***	0.102***	0.049***	0.094***	0.069***	0.060***
	[0.003]	[0.005]	[0.010]	[0.005]	[0.006]	[0.004]	[0.005]	[0.016]	[0.009]	[0.007]
Constant	1.505***	2.454***	2.158***	2.443***	2.361***	1.594***	2.505***	1.967***	2.264***	2.378***
	[0.045]	[0.070]	[0.130]	[0.074]	[0.079]	[0.055]	[0.075]	[0.201]	[0.117]	[0.093]
R^2	0.262	0.177	0.176	0.177	0.177	0.257	0.171	0.166	0.171	0.172

Notes: Dependent variable: The natural logarithm of hourly wage of main job. Over-education and under-education are in years. The models include married, female, good health, without children, Urban, ESB, NESB, occupation tenure, occupation tenure squared, occupation categories, industry categories, states and year dummies. The Hausman test rejects the null hypothesis that individual specific error is uncorrelated with the explanatory variables of the wage equation. Therefore, fixed effects estimates are preferred to random effects estimates when analysing the context of over-education. R^2 refers to overall R^2. Sample size: 40,183 observations of 9,516 individuals. Standard errors in brackets. ***1% level of significance; **5% level of significance, *10% level of significance. Data: HILDA Release 9 (Wave 1- Wave 9).

Table 6. Earnings return to over-education and over-skilling (Mode measure)

	(1)	(2)	(3)	(4)
VARIABLES	OLS	FE	FEIV IV Range-Half	FEIV IV JA
Over-educated and Over-skilled	-0.081***	-0.026***	-0.110***	-0.097**
	[0.010]	[0.009]	[0.030]	[0.043]
Only Over-educated	-0.011	-0.005	-0.035***	-0.031*
	[0.006]	[0.007]	[0.012]	[0.017]
Under-educated and Over-skilled	-0.002	-0.006	-0.028**	-0.025*
	[0.011]	[0.010]	[0.012]	[0.015]
Only Over-skilled	-0.027***	-0.008	-0.026**	-0.023*
	[0.010]	[0.008]	[0.010]	[0.012]
Only Under-educated	0.009	-0.003	-0.020**	-0.018
	[0.007]	[0.007]	[0.009]	[0.011]
Actual Education (years)	0.041***	0.023***	0.026***	0.025***
	[0.002]	[0.004]	[0.004]	[0.004]
Constant	2.339***	2.825***	2.818***	2.819***
	[0.029]	[0.058]	[0.058]	[0.058]
R^2	0.256	0.159	0.156	0.157
Individuals	35,794	8,901	8,901	8,901
Observations	35,794	35,794	35,794	35,794

Notes: Dependent variable: The natural logarithm of hourly wage of main job. The models include married, female, good health, without children, Urban, ESB, NESB, occupation tenure, occupation tenure squared, occupation categories, industry categories, states and year dummies. Education mismatch category variables are binary variables. The Hausman test rejects the random effects result and accepts the fixed effects result. R^2 refers to overall R^2. Standard errors in brackets. ***1% level of significance; **5% level of significance, *10% level of significance. Well matched is the reference category.
Data: HILDA Release 9 (Wave 1- Wave 9).