

The persistence of overskilling and its effects on wages

Kostas Mavromaras
Stephane Mahuteau
Peter Sloane
Zhang Wei

NATIONAL INSTITUTE OF LABOUR STUDIES,
FLINDERS UNIVERSITY



NCVER



Australian Government

Department of Education, Employment
and Workplace Relations

NATIONAL VOCATIONAL EDUCATION
AND TRAINING RESEARCH AND
EVALUATION PROGRAM

RESEARCH REPORT



Australian Government

Department of Education, Employment
and Workplace Relations

The persistence of overskilling and its effects on wages

Kostas Mavromaras
Stéphane Mahuteau
Peter Sloane
Zhang Wei

National Institute of Labour Studies, Flinders University

NATIONAL VOCATIONAL EDUCATION AND
TRAINING RESEARCH AND EVALUATION PROGRAM
RESEARCH REPORT

The views and opinions expressed in this document are those of the author/project team and do not necessarily reflect the views of the Australian Government, state and territory governments or NCVER.

Any interpretation of data is the responsibility of the author/project team.

Publisher's note

To find other material of interest, search VOCED (the UNESCO/NCVER international database <<http://www.voced.edu.au>>) using the following keywords: vocational education, employment, unemployment, educational level, wage, labour market, skill, competition.

© Commonwealth of Australia, 2012



With the exception of the Commonwealth Coat of Arms, the Department's logo, any material protected by a trade mark and where otherwise noted all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia <<http://creativecommons.org/licenses/by/3.0/au>> licence.

The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence <<http://creativecommons.org/licenses/by/3.0/legalcode>>.

The Creative Commons licence conditions do not apply to all logos, graphic design, artwork and photographs. Requests and enquiries concerning other reproduction and rights should be directed to the National Centre for Vocational Education Research (NCVER).

This document should be attributed as Mavromaras, K, Mahuteau, S, Sloane, P, Wei, Z 2012, *The persistence of overskilling and its effects on wages*, NCVER, Adelaide.

COVER IMAGE: GETTY IMAGES/THINKSTOCK

ISBN 978 1 921955 95 2

TD/TNC 106.18

Published by NCVER, ABN 87 007 967 311

Level 11, 33 King William Street, Adelaide, SA 5000
PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

P +61 8 8230 8400 F +61 8 8212 3436 E ncver@ncver.edu.au W <http://www.ncver.edu.au>

About the research

The persistence of overskilling and its effect on wages

Kostas Mavromaras, Stéphane Mahuteau, Peter Sloane, and Zhang Wei, National Institute of Labour Studies, Flinders University

Overskilling is the phenomenon whereby a worker's skills are underutilised in his or her job. Overskilled workers are employed, but they are underutilised and mismatched, in that their skills and abilities are not a good match with the requirements of the job. Overskilling can lead to decreased wages and job satisfaction, which suggests that the investment in skills for that individual has been somewhat wasted.

Overskilling mismatch has been shown to be persistent; that is, present overskilling mismatch increases the probability of future overskilling mismatch. However, the previous research showing this extends back only one year. This report examines the persistence of mismatch over a longer (up to three years) time period and its effect on wages.

An obvious explanation for the persistence of overskilling is that it reflects personal unobserved characteristics (such as the person having an inflated view of their own skills). This paper exploits longitudinal data to show that persistence is more than this, with the probability of being overskilled increasing if the individual has been overskilled in the previous period, after allowing for unobserved characteristics.

Key findings

- Overskilling is persistent: overskilling mismatch is common among those who have been overskilled in the past. Persistence varies by educational level, with its being lowest among university graduates and highest among VET diploma graduates and those who did not finish high school.
- The wages of university graduates are reduced by past overskilling, more so than for any other education level.

A possible reason for the second finding is that graduates tend to be in better-paid jobs and therefore there is more at stake for them. This observation is supported by the results of quantile regressions, which differentiate the impact of overskilling by whether an individual is at the top or the bottom of the earnings distribution. With the exception of certificate III and IV graduates, workers who are better paid among their peers are more likely to suffer higher wage penalties from being overskilled.

Readers may be interested in looking at earlier research reports on overskilling: *The incidence and wage effects of overskilling among employed VET graduates* available at <http://www.ncver.edu.au/publications/2231.html> and *Over-skilling and job satisfaction in the Australian labour force* available at <http://www.ncver.edu.au/publications/2365.html>.

Contents

Tables and figures	6
Executive summary	7
Introduction	10
Mismatch and overskilling in Australia	10
Overskilling as a form of mismatch	10
The economic meaning of the problem and why it matters for policy	11
Data and descriptive statistics	13
Education	13
Overskilling	14
Over-time persistence of overskilling	15
Overskilling persistence and wages	17
Estimation methodology	19
Probability of being currently overskilled and self-persistence	19
Modelling state dependence	21
Allowing for different types of unobserved heterogeneity	22
Scenarios of skill mismatch experience	22
The effect of past overskilling on wages	23
Regression results	26
The effect of past overskilling on future overskilling: probit analysis	26
Self-persistence of mismatch: the effect of past overskilling mismatch on current overskilling mismatch	29
The effect of past overskilling on wages: linear analysis	31
The effect of past overskilling on wages: quantile regression	33
Estimating scarring	39
Conclusion	42
References	44
Appendix A	45
Appendix B	50
NVETRE Program funding	53

Tables and figures

Tables

1	The distribution of highest educational levels of paid employees	14
2	The distribution of skill utilisation	14
3	The distribution of overskilled employees by education pathway	15
4	Patterns of overskilling over time	15
5	Distribution of patterns of overskilling in four consecutive periods	16
6	Probability of being presently overskilled by past overskilling patterns, %	17
7	Hourly wages by overskilling pattern and educational level	18
8	Hourly wages penalty by overskilling pattern and educational level, %	18
9	Dynamic probit estimations of overskilling: all education levels combined 27	
10	Dynamic probit estimations of overskilling by educational level	28
11	Estimated overskilling probabilities by educational level and scenario of overskilling 29	
12	Wage estimation by education level (log hourly wage)	32
13	Quantile wage estimation by education level (log hourly wage)	34
14	Scarring predictions by education level	41
A1	Descriptive statistics	48

Figures

1	Self-persistence of overskilling mismatch	31
2	Quantile estimates of wages and overskilling, university graduates	36
3	Quantile estimates of wages and overskilling, diploma	37
4	Quantile estimates of wages and overskilling, certificate III/ IV	37
5	Quantile estimates of wages and overskilling, Year 12	38
6	Quantile estimates of wages and overskilling, below Year 12	38

Executive summary

The purpose of this research is to examine the longer-run outcomes of employee overskilling as a form of labour market mismatch. The research focuses on full-time and part-time employees in Australia between the years 2001 and 2009. Overskilling mismatch has been shown in the literature to cause losses in wages and in job satisfaction, both of which provide direct and indirect indications of reduced productivity in Australian workplaces. Overskilling mismatch occurs when someone is in paid employment, but where their skills and abilities are not fully utilised. Overskilling mismatch and other forms of on-the-job mismatch extend the conventional (job search) concept of mismatch, where workers take time to find new jobs or are between jobs (often referred to as frictional unemployment or search unemployment). The two phenomena are clearly related, in that they can both be the manifestation of underutilisation of national human capital. Overskilled workers are employed, but are underutilised and mismatched, in that their skills and abilities are not a good match with the requirements of the job.

This project builds on several relevant recent research findings on overskilling mismatch. The first such finding is that most adverse labour market outcomes stem from overskilling mismatch, where general skills and abilities are underutilised, and not from over-education, where formal qualifications are underutilised. Hence we focus on overskilling mismatch. The second such finding is that overskilling has been shown to be self-persistent; that is, present overskilling mismatch begets future overskilling mismatch. However, existing results in the literature extend back one year only. The attribute of self-persistence is common among many adverse labour market outcomes, including the very similar phenomena of long-term unemployment and underemployment. The third relevant finding that we build on is that much evidence about mismatch suggests that the workings of mismatch in the labour market are related to the educational level of the worker. This is a theme that has been running through our stream of mismatch research: we find that the way human capital is fully or less than fully utilised is intimately related to the formal qualifications of the workers concerned.

We carried out multivariate panel regression analysis using the first nine waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey data to estimate the *causal* effect of past overskilling mismatch on present overskilling mismatch, and the causal effect of having been overskilled in the past on present wages. Our analysis confirms that there are extensive differences in educational level in the effects of past mismatch on present mismatch and on present wages. We carried out several cross-section estimations as part of our usual robustness investigation procedures; however, we depart from our past practice of reporting the comparison between cross-section and panel evidence on mismatch, as we consider that enough is now known on the subject of controlling for unobserved heterogeneity through panel estimation. Up-to-date applied economic and econometric evidence suggests that we should only be looking at cross-section results when there is a specific reason for doing so, recognising that they will not provide causal evidence and that they will typically contain biases generated by cross-section estimation.

We confirm previous results on self-persistence of overskilling mismatch and extend the analysis to incorporate the effect of overskilling that occurred up to five years in the past on present overskilling and present wages. As there is no such research in the national or international

literature, we have been cautious and have experimented with many different models to discover the degree to which past overskilling shows self-persistence. We have found evidence that overskilling can be self-persistent for at least five years, but at the same time we have found evidence that the model that must be used for estimating self-persistence reaches its limits when we build five lags in the model specification and attempt to make estimations using a longitudinal dataset that covers nine waves. We therefore settle for the less ambitious but definitely more robust model which incorporates the overskilling that happened up to three years ago.

We use the multivariate regression results to predict the over-time effect of self-persistent overskilling mismatch on future overskilling by educational level. The comparisons we report are between people who differ only in terms of their overskilling mismatch self-persistence, with all their other characteristics set at the mean levels for people with the same educational level. It is important to note that this is an 'other things equal' comparison, which allows us to generalise the results to the population and which can only be achieved using multivariate regression. Reported results are accompanied by the level of their statistical significance.

We find that self-persistence in overskilling mismatch is very large among those who have been overskilled in the past and differs a great deal by educational level. A university graduate who was overskilled in all three past years has a 38% chance of being overskilled in the coming year. An equivalent university graduate who was not overskilled in any of the last three years has a chance of only 4.6% of becoming overskilled in the coming year! Note that this compares two university graduates with the same average graduate characteristics, their only difference being their past overskilling status. The difference of 33.4 percentage points is large. Overskilling mismatch self-persistence for workers who were overskilled in all three past years (and with their well-matched counterparts in brackets) has been estimated to be 59.8 (6.3)% for diploma vocational education and training (VET) graduates, 60.2 (8.4)% for certificate III and IV VET graduates, 56.3 (11)% for Year 12 school graduates, and 66.1 (12.9)% for those who did not complete school. These numbers show that overskilling mismatch is not only a self-perpetuating phenomenon, but it is also a phenomenon with a very strong 'labour market memory'. The next step of the research was to establish the damage that overskilling mismatch inflicts on those who are persistently overskilled.

Previous research suggested strongly that the effect of present overskilling mismatch on present wages can be considerable and will typically hurt university graduates. VET graduates have on most occasions not been found to suffer a strong overskilling wage penalty. The evidence on Year 12 and less than Year 12 school graduates is both weak and mixed, principally because the wage distribution for these groups is more compressed, so that if there are to be any wage penalties, they will have to be small. The effect of past overskilling mismatch on present wages is estimated by education level. We find that the wages of university graduates are reduced by past overskilling. There is some mixed and weak evidence for diploma holders. The wages of certificate III and IV VET graduates are not influenced by past overskilling. There is a mixed picture for workers without post-school qualifications, with no discernible patterns. The only unambiguous result is therefore that the present wages of university graduates suffer from past overskilling, that this effect lasts for at least three years, and that it shows no sign of diminishing in strength over time.

In order to refine the wage results, we apply the method of quantile regression to estimate the effect of past overskilling on wages at different points of the wage distribution. Quantile

regression allows us to examine whether the overskilling penalty is primarily borne by the best-paid- or the worst-paid. Especially regarding university graduates, we want to know if it is the better or the worse-paying graduate jobs that suffer more from overskilling penalties. Quantile regression estimates produce a clear result: they show that the university graduates who work in the better-paid jobs are the ones who suffer the highest overskilling self-persistence wage penalties. Quantile regression results also shed new light on the effect of overskilling on the wages of diploma VET graduates. By differentiating between the better- and worse-paying jobs, we see that workers at the top of the (diploma VET graduates) wage distribution appear to suffer overskilling wage penalties similar to those of university graduates, while those in the middle and the bottom of the distribution do not suffer any such overskilling wage penalties. Certificates III and IV VET graduates do not suffer any overskilling wage penalties. Finally, there is some evidence that among those with no post-school qualifications the better-paid suffer some discernible overskilling wage penalties. Given that workers without post-school education operate within a compressed wage distribution (especially from below through the minimum wage), this is a strong result.

The incorporation of quantile regression in the analysis allows a more general picture to emerge for all education levels, with the exception of certificates III and IV. We now find that, to a varying degree, workers who are better paid among their peers and become persistently overskilled are more likely to suffer a higher overskilling wage penalty. Although the effect is principally concentrated among the better-paid university graduates, it is worth noting that it is also present among the best-paid workers with no post-school qualifications. Quantile regression results offer further confirmation of the previously established evidence that the wages of certificates III and IV VET graduates are not influenced by the self-persistence of their overskilling mismatch.

The research suggests that overskilling can impose real costs on individuals, employers and the economy. To the degree that an overskilling wage penalty may reflect a productivity loss, its state persistence is shown to be strong; the associated wage penalties suggest that the implied productivity losses can also be large. The similarities with other forms of human capital underutilisation can be informative. In the same way that we cannot know whether frictional unemployment is at an optimal level, we cannot know whether skills underutilisation resulting from overskilling mismatch is at an optimal level. If we find that changing the circumstances surrounding underutilisation turns out to be more expensive than the losses (or foregone benefits) resulting from underutilisation itself, it would be sub-optimal to argue for policy intervention. As part of the discussion of the estimation results, we provide some calculations of the benefits that could result if we reduced the underutilisation resulting from overskilling mismatch. These calculations must be viewed as partial equilibrium results and, accordingly, treated with the necessary caution that partial equilibrium analysis warrants. Notwithstanding this caveat, given that the results we present here have been derived using a reliable nationally representative dataset and that they are based on robust econometric longitudinal analysis, they provide the best available evidence on this subject. The large estimates of the national cost of overskilling mismatch resulting from the estimated wage penalties indicate that the economy-wide losses due to overskilling should not be ignored.

Introduction

Mismatch and overskilling in Australia

Mismatch in the labour market as a form of human capital underutilisation is attracting increased attention by both academic researchers and policy practitioners. Skills underutilisation in the workplace is defined as the situation in which workers may possess skills and abilities not used in their job, which has been called overskilling. To the degree that mismatches may represent a genuine labour market imbalance, they can have a dampening effect on the growth potential of the economy.

Overskilling as a form of mismatch

Existing research on labour market mismatch caused by overskilling has extensively examined the extent of overskilling in Australia, its impact on wage levels, on job mobility, and on job satisfaction, as well as its persistence over time, using the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The research established several core results (see Mavromaras, McGuinness & Fok 2009a, 2009b, 2009c; Mavromaras et al. 2010a, 2010b, 2011). First, overskilling has a negative wage effect. The overskilled are observed to have lower wages than their well-matched comparators.¹ This has been called the 'overskilling wage penalty' and has been measured as a percentage wage difference between the wages of the overskilled and the well-matched. The overskilling wage penalties differ by level of education. The highest wage penalty is suffered by university graduates who are overskilled. The second highest wage penalty is suffered by overskilled workers without post-school qualifications, with the lowest by workers with VET certificates III/IV. In some instances empirical research has not managed to establish a statistically significant wage penalty for overskilled VET graduates.

Second, there is limited evidence suggesting that overskilling is a self-persistent labour market state. The meaning of 'self-persistent' is that an overskilled person today is more likely to be overskilled tomorrow, *because* they are overskilled today. That is, being overskilled is a self-perpetuating state, *over and above* all the personal and labour market reasons that make someone overskilled in the first place. Self-persistence is a common problem with adverse labour market outcomes and is typically associated with some form of scarring, which works against the longer-term prospects of the worker concerned. The investigation of self-persistence requires longitudinal data. Mavromaras, McGuinness and Fok (2009b) provided limited evidence on the self-persistence of overskilling in the Australian labour market. They found that the education pattern followed by the strength of overskilling self-persistence is similar to that followed by the wage penalties. University graduates suffer the highest wage penalties and show the highest persistence, workers without post-school qualifications are next, and certificate III/IV graduates suffer the least. However, the analysis was limited by the use of too short a longitudinal dataset (the first six waves of the Household, Income and Labour Dynamics in Australia Survey), which

1 A meaningful comparison is made between the wages of those who are overskilled against the wages of those who are well matched, keeping all other characteristics (for example, education level, number of children etc.) at the mean values in the whole sample. This should be viewed as just a (widely accepted) convention for making comparisons using the concept of an 'average person'.

restricted the dynamic analysis to one lag only. This study utilises the first nine waves of the survey, which enables us to experiment with lags of different lengths, which is particularly important for the analysis of persistence.

While previous research provides evidence on the presence of these two effects, their interaction has not been examined. Namely, we do not know how much the negative wage penalties may themselves be persistent. It is this question that this research addresses. Underlying the question is the deeper economic distinction of whether overskilling mismatch is either a scarring labour market phenomenon – that is, a phenomenon which we would expect to cause long-lasting damage to those workers who end up being overskilled in the long run – or an adjustment process, in which case we would expect the labour market negative outcomes associated with overskilling to be transient and eliminated by the market with time. To establish this we need to examine how both the self-persistence of overskilling mismatch and the wage penalties that are caused by self-persistent overskilling mismatch may work in tandem to create (i) a disadvantage that will not be time-persistent, or (ii) a disadvantage that will be time-persistent. We have identified an extreme form of scarring in this latter case, where self-persistence and wage penalties may reinforce each other. The project will establish the presence and the effect of scarring in overskilling.

The economic meaning of the problem and why it matters for policy

The main impetus behind this project is to establish whether or not overskilling mismatch is a self-perpetuating labour market phenomenon. We explain this by using the example of someone who is underutilised in their workplace. Persistent underutilisation may be the result of experienced older workers having irrelevant, unwanted, or even no qualifications. It can also be the result of inadequate management and a lack of alternative opportunities for the underutilised workers. In the case of recently arrived immigrants, underutilisation may be the result of their lack of language abilities. Typically, simple persistence will be due to reasons that we do not expect to worsen after longer periods of underutilisation. Self-persistence is the case where being underutilised today will in itself increase the chances of being underutilised in the future. For example, a researcher who is continuously allocated mundane and time-consuming projects will be underutilised but will at the same time increase their chances to remain so, as their counterparts who are allocated the challenging projects will develop their human capital more rapidly and more effectively. A journalist who is only allowed to do the local news or a floor worker who is always asked to operate the same machine could similarly develop their expertise in a job that is narrow in its conception and which only uses part of the workers' abilities. Although it is understandable that an employer may wish to have someone with spare skills and abilities on standby, it is not immediately obvious why they may want to have them in that state for long periods. However, in a situation where the researcher has spent some five to ten years without developing their research skills, their long-term promotion prospects are bound to be harmed, in that the probability of their getting a job where they are not underutilised may be lower. This is the essence of self-persistence. Underutilisation in the past increases the probability of underutilisation in the future.

The policy responses to simple persistence and to self-persistence have to be different. In the case of simple persistence, the reasons that cause underutilisation must be addressed. For example, the recently arrived immigrant must be given the chance to improve their language

ability. The older and experienced but unqualified worker should be given recognition in a training context for their experience and enabled to obtain the right qualifications. Persistence, in this sense, is a relatively straightforward policy issue.

By contrast, self-persistence is an invidious problem. First, we do not know when the underutilisation sets in and why. Once self-persistent underutilisation has set in, the deterioration of the human capital of the underutilised worker begins to reinforce the process and then (self-persistent) underutilisation becomes much harder to combat. Empirically, it is very difficult to identify self-persistence at the individual level. However, as we show below, it is feasible to identify self-persistent underutilisation in the form of overskilling mismatch in a statistical manner. That is, we can identify the types of workers who are more likely to be underutilised in their workplace. Although this knowledge is at the statistical level, it is useful as the basis of information for individual job search, better human resource management strategies by employers and job-placement activities by employment agencies. The concern about combating self-persistence in general is to have policies in place that will not allow underutilisation to begin.

This research sets out to design and estimate a framework of analysis that will allow us to establish whether or not overskilling mismatch is self-persistent, by distinguishing between the factors that may make mismatch persistent (described in the previous paragraphs as simple persistence) and the capacity of mismatch to be self-perpetuating (described as self-persistence). We then use the statistical model to examine the degree of damage that self-persistence generates in terms of lost income for the overskilled workers. The exercise is highly empirical and is based on the examination of nationally representative individual data on people who were repeatedly interviewed between 2001 and 2009. The next chapter presents the data. It describes overskilling mismatch and other pertinent attributes of the data, while the following chapter presents the estimation methodology. The results are then presented and discussed, after which conclusions are given. The appendix contains the core detailed estimations and an extended section of supplementary material that contains the full account of all estimations used in the report.

Data and descriptive statistics

The data for this research come from the first nine waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey, which began in 2001 (wave 1) with a large national probability sample of Australian households and their members.² We use an unbalanced panel — that is, a panel where certain years of observations are allowed to be missing for some individuals — of all working-age people (16–64 years for males and 16–59 years for females) in paid employment who provided complete information on the variables of interest. Self-employed and full-time students are excluded. The sample size we retain is approximately 5300 observations per wave.

The next section provides a descriptive picture of the problem of overskilling, its persistence, and outcomes in terms of wage differentials in Australia, using the survey sample. We define the indicator of overskilling mismatch used throughout this research and look at its distribution by education level. We then refine our description using the longitudinal feature of the survey sample in order to decompose the distribution of overskilling further by patterns of overskilling over time. We define scenarios of matching pathways based on individuals' experiences in the three years prior to the current interview. This allows us to look at the relationship between employees' histories of mismatch and their current situation; that is, whether or not we observe a persistence of overskilling over time and whether some education levels appear to be more affected than others. Finally, we focus on the consequences of overskilling with regard to the potential wage penalties that may be associated with the phenomenon of scarring. We cross-tabulate individuals' educational levels and overskilling histories and we analyse their wage differences as a manifestation of scarring.

Education

We split the data into the following five educational categories based on the highest education level achieved at the time of interview:

- university degrees (including bachelor degrees, graduate certificates, graduate diplomas and higher degrees)
- advanced diplomas and diplomas
- certificates III and IV
- only completed school (Year 12)
- did not complete school (below Year 12).³

² See Watson and Wooden (2004) for a detailed description of the HILDA data.

³ Certificates I and II have been subsumed into the category 'Did not complete school', depending on their highest year of schooling completed.

Table 1 reports the distribution across educational categories. The data here are pooled person–year observations across all nine waves of the survey. Table 1 shows that the sample sizes in each educational category are sufficiently large for statistical analysis, including running regressions split by educational category, with some reservations about the ‘Advanced diplomas and diplomas’ category when performing panel regressions, since they represent only 9.7% of the sample. We can see that almost 40% of the employees do not have a post-school qualification, while 28.7% are university graduates, and 21.8% have a certificate III or IV.

Table 1 The distribution of highest educational levels of paid employees

	Highest education					Total
	Did not complete school	Only completed school	Certificates III/IV	Diplomas	Degrees	
Cases	11 436	7672	10 434	4632	13 763	47 937
Per cent	24	16	22 10		29	100

Note: The sample is person years of working-age paid employees from HILDA 2001–09.

Overskilling

The overskilling variable used in this research is derived from the self-completed questionnaire of the Household, Income and Labour Dynamics in Australia Survey. Interviewees are asked to respond on a seven-point scale to the statement *I use many of my skills and abilities in my current job*, with a response of 1 corresponding to strongly disagree, up to 7, strongly agree. Individuals selecting 1, 2, 3, or 4 on the scale are classified as overskilled and those selecting 5 or higher are classified as well-matched.⁴

Table 2 reports the distribution of skill utilisation. The cut-off point we select to define overskilling (4 or less) gives us sufficient observations to support regression analysis. The way the overskilling question is asked in the survey does not allow the researcher to examine the phenomenon of underskilling and so we do not address this further. Thus, all comparisons and results in the following analysis look just at the underutilisation side of potential skills mismatch.

Table 2 The distribution of skill utilisation

	Question: ‘I use many of my skills and abilities in my current job’							Total
	1	2	3	4	5	6	7	
Cases	1409	2003	2448	4869	9245	17 090	10 888	47 952
Per cent	3	4	5	10	19	36	23	100

Note: The sample is person years of working-age paid employees from HILDA 2001–09; 1 stands for strongly disagree (strongest mismatch) and 7 for strongly agree (strongest match).

⁴ In previous research we have categorised overskilling as well-matched, moderately overskilled and severely overskilled (for example, Mavromaras et al. 2009a). For the purposes of this research it would have been too complicated to retain this three-way split of overskilling. Extensive experimentation has revealed that reducing the variable to a two-way one does not alter the core results of the analysis. The same holds for the choice of the cut-off point. The application of different cut-off points makes no qualitative difference, with lower cut-off points reducing the number of those categorised as overskilled but obviously picking from the sample the more overskilled people, thus increasing the size of the effects of overskilling, but applying it to a smaller part of the sample.

Table 3 reports the incidence of overskilling by educational level. The data show that the incidence of overskilling varies significantly across educational category. Surprisingly, overskilling is the most severe among employees with no post-school qualification, particularly for those not completing Year 12, as we would expect a positive relationship between overskilling and level of education.

Table 3 The distribution of overskilled employees by education pathway

	Highest education					Total
	Did not complete school	Only completed school	Certificates III/IV	Diplomas	Degrees	
Cases	3426	2230	2049	924	2094	10 723
Per cent	30	29.20		20	15	22

Note: The sample is person years of working-age paid employees from HILDA 2001–09.

Over-time persistence of overskilling

In the following we look at the patterns of overskilling over time and describe over-time persistence for a number of scenarios of past individual experience. Table 4 describes the patterns of overskilling over time by education level. It takes the present overskilling state (row below educational levels) and then provides the numbers found for eight distinct scenarios, beginning with '000' for never overskilled in the past three years, to '111' overskilled in all of the last three years. The largest category is those who are presently well matched and who were never overskilled in the last three years (for example, for university graduates, 4503 person years). Note that there is a sizeable minority who were continuously overskilled for all years (218 for university graduates and 1057 for the whole sample). The information in table 4 is very complex and can be presented in several ways. We have used it to derive the next two tables which describe the persistence of overskilling with more precision.

Table 4 Patterns of overskilling over time

Over-skilling at t-1, t-2, t-3	Overskilling status at time t									
	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Well matched	Over-skilled	Well matched	Over-skilled	Well matched	Over-skilled	Well matched	Over-skilled	Well matched	Over-skilled
000	4503	242	1291	96	2823	242	1425	149	1926	233
100	223	94	80	38	231	101	137	72	225	114
010	252	49	102	22	244	77	141	64	234	104
001	355	66	122	28	275	64	212	61	272	102
110	88	83	23	42	84	89	69	61	96	136
101	57	53	30	23	79	68	66	68	108	94
011	170	58	43	31	104	75	112	70	136	100
111	130	218	52	105	86	143	98	211	133	380

Note: Since three lags of overskilling are introduced, the sample here only contains working-age paid employees from HILDA 2004–09. For ease of interpreting an overskilling combination in the first column, 1 refers to a period of overskilling and 0 to a period where the employee is well matched.

Table 5 describes the distribution of the patterns highlighted in table 4, ignoring the sequence in which overskilling happened and including percentages. We find that the majority of the sample never experienced overskilling in any four consecutive periods and that this proportion is

positively associated with the educational level (68% for university graduates against 44% for those who did not complete school). This can be considered differently by looking at experiencing overskilling in any single period. Then, we see 43% of the whole sample experienced overskilling at least once in the past four years. We can see that the percentages of high school non-completers and high school graduates who are overskilled throughout the past four years are larger than those of more qualified employees. We see that the lower the qualification level, the more distributed the individuals are across all possible patterns. By summing up the percentages of the categories overskilled twice, three times and always, we can see that between 32% and 35% of the those with no post-school qualifications have experienced at least two episodes of overskilling in the past, against 16% for university graduates and 21% for VET certificate holders.

Table 5 Distribution of patterns of overskilling in four consecutive periods

	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Cases	%	Cases	%	Cases	%	Cases	%	Cases	%
Never overskilled	4503	68	1291	61	2823	59	1425	47	1926	44
Overskilled once	1072	16	400	19	992	21	639	21	964	22
Overskilled twice	524	8	184	9	509	11	444	15	660	15
Overskilled three times	324	5	148	7	318	7	297	10	463	11
Always overskilled	218	3	105	5	143	3	211	7	380	9
Total	6641	100	2128	100	4785	100	3016	100	4393	100

Table 6 presents the proportion of individuals who are currently overskilled given the path they have followed in the past three years. For instance, 5% of the university graduates who have never been overskilled before (path '000') are currently overskilled. This means that 95% of them are currently well matched. By comparison, 11% of the high school non-completers who followed the same path are currently overskilled. This table shows that more recent spells of overskilling are associated with larger proportions of individuals who are currently overskilled. Indeed, comparing the second, third, and the fourth rows of the table where one period of overskilling occurs in total, proportions are higher for the scenario involving the most recent spell, where the proportions of currently being overskilled in row 2 (past overskilling type 100) are all more than 30%. The table also shows the extent to which overskilling persists for people who have experienced it in the past, all the more so if they have experienced it several times. We can see that accumulating three consecutive years of overskilling leaves little chance of becoming well matched now.

Table 6 Probability of being presently overskilled by past overskilling patterns, %

Overskilling at t-1, t-2, t-3	Overskilling at t				
	University degrees	Diplomas	Certificates III/IV	Only completed school	Did not complete school
000	5	7	8	9	11
100	30	32	30	34	34
010	16	18	24	31	31
001	16	19	19	22	27
110	49	65	51	47	59
101	48	43	46	51	47
011	25	42	42	38	42
111	63	67	62	68	74

Note: 1 refers to a period of overskilling and 0 to a period where the employee is well matched.

Overskilling persistence and wages

We know that current overskilling is associated with lower wages. Table 7 shows how present wages are associated with past overskilling patterns by educational level. The more spells there are of overskilling, the lower the wage. This is particularly so for university graduates and, to a certain extent, also for diploma graduates and Year 12 graduates. There is no association between wages and overskilling for certificate III/IV holders and workers with Year 11 or less.

Table 8 presents the percentage reduction of the wage of presently overskilled workers compared with presently well-matched workers for different overskilling histories. It is clear that more overskilling leads to higher wage losses for better-paid workers (university graduates and diploma graduates). The same applies to Year 12 school graduates, with little evidence of wage losses being influenced in any systematic way by the historic pattern of overskilling for certificate III/IV graduates and Year 11 or less workers.

Table 7 Hourly wages by overskilling pattern and educational level, \$

Over-skilling at t-1, t-2, t-3	Overskilling at t									
	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	<i>Well matched</i>	<i>Over-skilled</i>	<i>Well matched</i>	<i>Over-skilled</i>	<i>Well matched</i>	<i>Over-skilled</i>	<i>Well matched</i>	<i>Over-skilled</i>	<i>Well matched</i>	<i>Over-skilled</i>
000	34.6	31.8	29.4	25.4	25.0	23.6	25.6	23.3	22.5	19.8
100	32.4	30.5	26.4	23.7	23.5	23.3	21.8	21.8	20.7	19.0
010	31.5	31.5	26.8	24.3	25.3	23.8	22.6	20.1	20.9	20.7
001	30.9	29.4	26.2	24.7	23.5	21.4	21.8	21.6	19.7	20.0
110	28.6	26.3	23.1	24.0	22.6	22.7	23.6	19.9	18.8	21.8
101	29.8	26.0	23.9	24.1	24.3	22.0	21.6	19.7	18.6	19.0
011	27.5	25.4	23.2	21.7	22.4	23.1	20.4	21.4	19.2	18.8
111	24.8	24.6	20.5	21.3	22.6	23.7	21.1	19.7	19.8	19.1

Note: Wages are measured in Australian dollars as nominal gross hourly wages from the main job. 1 refers to a period of overskilling and 0 to a period where the employee is well matched.

Table 8 Hourly wages penalty by overskilling pattern and educational level, %

Number of spells	Overskilling at t-1, t-2, t-3	Hourly wage penalty of being overskilled at time t				
		University degrees	Diplomas	Certificate III/IV	Only completed school	Did not complete school
0	000	-8.1	-13.6	-5.6	-9.0	-12.0
	100	-11.8	-19.4	-6.8	-14.8	-4.0
1	010	-9.0	-17.3	-4.8	-21.5	4.5
	001	-15.0	-16.0	-14.4	-15.6	1.0
2	110	-24.0	-18.4	-9.2	-22.3	10.1
	101	-24.9	-18.0	-12.0	-23.0	-4.0
3	011	-26.6	-26.2	-7.6	-16.4	-5.1
	111	-28.9	-27.6	-5.2	-23.0	-3.5

Note: The reference category is 'always well matched', which combines scenario '000' with being well matched at t, for each education level. 1 refers to a period of overskilling and 0 to a period where the employee is well matched.

Estimation methodology

While the descriptive statistics are informative in their own right, they do not allow us to understand the phenomenon of overskilling in its complexity and take account of its dynamic aspect. In assessing the extent of the scarring effect associated with the past experiences of mismatched employees, we need to account for other factors that influence overskilling and isolate those from the effect of previous overskilling mismatch. Moreover, we do not just want to know whether scarring exists, but also to gain information relating to the length of time the state of overskilling is likely to produce negative effects on the labour market outcomes and prospects of workers. Multivariate analysis allows us to disentangle these effects.

The multivariate analysis that follows uses the nine available waves of the Household, Income and Labour Dynamics in Australia Survey. Again, we deal only with people of working age and in paid employment and we exclude full-time students. Our estimations comprise three steps. First, we estimate the extent to which overskilling is persistent; that is, we investigate whether there are reasons and characteristics that make individuals persistently overskilled, and the degree to which overskilling is self-persistent over time. That is, we assess whether previous spells of overskilling make future spells more likely. We distinguish between individual workers according to their highest level of educational achievement and estimate overskilling persistence and self-persistence separately for each level, namely, university graduates, diploma, VET, Year 12, and below Year 12. Second, we look for evidence of scarring by educational level, which requires two stages of analysis. We establish first the presence and extent of overskilling mismatch wage penalties; and, second, we establish whether or not past overskilling produces lasting future wage penalties. Two types of earnings models are estimated in order to give as precise a picture as possible on the scarring effect of overskilling. We use a model that estimates the conditional mean wages (a linear model) and a model that estimates wages at different points of the wage distribution (a model of quantile regressions). The final step brings the results of all previous estimations together and computes expected wages according to individual overskilling experience. Putting all of the results together in the form of expected wages in different scenarios allows us to make meaningful comparisons of the extent of scarring and how this may differ by educational level.

Our estimation technique is set up so that we can investigate a large variety of possible scenarios, ranging from people who have been well matched throughout the time we observe them, to people who have been overskilled for the whole period of review. Drawing up such scenarios allows us, first, to give a precise picture of persistence and self-persistence of overskilling depending on the individual overskilling history, and, second, to quantify the circumstances and consequences on hourly wages.

Probability of being currently overskilled and self-persistence

The first research question we address relates to the analysis of the probability that a worker is currently overskilled in their job, with a particular emphasis on the lasting impact of previous experiences of overskilling. The explicit aim is to compare individual probabilities across a range of possible overskilling mismatch scenarios. These are: (i) always well-matched; (ii) overskilled in the previous year, but not the two years before; (iii) overskilled in the previous two years but not

in the one before; (iv) overskilled in all past three years. The objective is to identify how long a spell of overskilling exerts an influence on the probability of being overskilled and the extent to which cumulative spells of past mismatch further promote future mismatch.

The outcome variable of interest in these estimations is dichotomous; that is, it can only assume two values: 1 if the individual is overskilled in their job, and 0 if they are well matched. The appropriate estimation method for this type of outcome variable (henceforth, *dependent variable*) is non-linear estimation, typically carried out using the probit model. In the present context the probit model has many advantages over its linear counterpart, a major one being that it ensures an estimated outcome lying between 0 and 1, which is necessary for making meaningful predictions when we estimate probabilities. Furthermore, the use of the probit model in the present context is also grounded in economic theory, as the outcome can be modelled in an individual utility maximisation framework of decision-making through a latent variable interpretation of the dependent variable.

The longitudinal nature of the Household, Income and Labour Dynamics in Australia Survey allows us to address the research question of how persistent overskilling is over time by using nine waves of observations (2001–09). Panel estimation confers the added advantage of controlling for unobserved individual heterogeneity, thus allowing the estimation of direct and causal impacts of past experiences on current overskilling and wages. In order to estimate the dynamics of overskilling we use a technique known in the literature as the ‘dynamic random effects probit model’, which corresponds to the state of the art technique used to investigate dynamic panel data models involving binomial dependent variables.

A main advantage of multivariate regression is that it accounts simultaneously for observed individual heterogeneity by incorporating in the analysis a number of independent variables. In our estimations, we use variables that represent individual and workplace characteristics such as gender, firm size, industry dummy variables, hours worked per week, and tenure in the current occupation. Given our research question, we also present a number of dummy variables indicating whether the employee was overskilled or not in the last three years before the current one. We use one dummy variable for each past year, so that for three years we have three variables. We call these variables *lagged dependent variables*. We test two alternative models: one with three lags – that is, including information on past overskilling for three years – and another with five lags for five years. We display the results of the first model in the body of the text, since statistical tests lead us to accept the superiority of this model over the five-lag model. The statistical significance of each one of the lagged dependent variables indicates whether previous spells of overskilling have an effect on the estimated probability that an employee is currently overskilled. We use the estimated coefficients to compute the estimated probabilities of being currently overskilled under the four possible scenarios mentioned above, setting all other variables to equal the value of their sample means.⁵ Since we evaluate these probabilities setting all other explanatory variables to equal their sample means, the differences between scenarios give the effect of the scarring from consecutive overskilling periods. We

⁵ We remind the reader that the first scenario corresponds to employees who have been well matched in the past three waves of the survey. The second scenario corresponds to employees who were overskilled in the previous wave but not in the two before that. The third scenario corresponds to employees who were overskilled in the past two waves but not in the third wave, and, finally, the fourth scenario corresponds to employees who were mismatched for all past three waves.

perform the estimations and the subsequent scenario analysis for the whole sample of employees in the data, as well as by educational level.

Modelling state dependence

A major aim of this report is to help us understand whether past overskilling will cause future overskilling in itself, over and above the factors that have caused overskilling in the first instance. To disentangle these possible causal effects we need to use a model that controls for unobserved individual heterogeneity (through panel estimation) and which includes in the list of explanatory (independent) variables the lagged dependent variables. This method is called dynamic panel estimation and its purpose is to tell us whether past overskilling affects present overskilling, over and above all other observable and unobservable factors accounted for in the regression. There have been many names in the literature for variables where past values of the dependent variable influence their present value, including state dependent, self-persistent, self-perpetuating and others. In essence they all mean that the process at hand is self-perpetuating and its past presence causes an increase in the probability of its future presence. In this report we use the term 'self-persistent', but it is worth noting that this is just one of the many terms found in the literature.

The modelling adopted in this study is based on the Wooldridge (2005) approach and takes into account the *initial conditions* problem (Heckman 1981), which arises when a lagged dependent variable is added as a covariate in the model. Such inclusions are necessary in our study, since we are mostly interested in the persistence of overskilling over time and its consequences. The *initial conditions* problem occurs because individual overskilling in the initial period is likely to be correlated with individual-specific unobserved characteristics that affect the probability of being overskilled in later periods. Ignoring this problem would prevent us from disentangling the actual effect of self-persistence: the impact of experiencing a period of overskilling on the probability of experiencing overskilling later. Indeed, if we picked at random a sub-sample of individuals who are currently overskilled and another sample of individuals who are currently well matched, both of which are observed in sub-samples in the next survey wave, it is likely that we would observe more overskilled employees in the sample of those who were initially overskilled. It is crucial to note that this observation would not be sufficient for us to conclude that past experience in overskilling *causes* future overskilling. A part of this persistence in overskilling could be due to the fact that individuals in both initial samples have different characteristics (both observed and unobserved), some of which have an impact on both present and future overskilling. To be more precise, assume that working in the public sector (which is observed in our data) and motivation (which is unobserved in our data) are two such factors. From our estimations we know that public sector workers are less likely to be overskilled and we will, for expository purposes, assume that workers with higher motivation are also less likely to be overskilled. In the two sub-samples we have collected, we expect that the proportion of public sector workers will be lower among the overskilled, as will the proportion of above-average motivated workers. Now let us examine what happens with self-persistence. Let us assume that there is no self-persistence. That is, assume, *other things equal*, presently overskilled workers are as likely to become overskilled as are presently well-matched workers. How will the two sub-samples fare in their future overskilling? Clearly, those who were overskilled in the past will include more overskilled workers in the future, because they include fewer public sector workers and fewer above-average motivated workers. It is crucial to note that this 'apparent' or 'spurious' self-persistence is a mere

statistical artefact occurring as the result of the two characteristics of public sector employee status and motivation being non-randomly distributed in the two sub-samples. The method we use here allows us to control for all observed heterogeneity (in our example we include a ‘public sector employee status’ variable) and unobserved individual heterogeneity (through the panel estimation’s ‘individual effect’). The remainder, captured by the lagged overskilling variables, can thus be correctly attributed to self-persistence.

To be more specific, unobservable heterogeneity is tackled in our estimations through the inclusion of a random component of an individual-specific effect in the unobservable part of the model, thus splitting the latter into a *within* effect (capturing individual unobserved heterogeneity) and a *between* effect. As per the *initial conditions* problem, the specification of our model accounts for the correlation between the individual-specific effect and the lagged dependent variable through the approximation method suggested by Wooldridge (2005). It consists of modelling the relationship between the individual effect and the lagged dependent variable as a function of the initial observation of overskilling and a set of instruments.⁶ This method allows us to decompose the persistence of overskilling – state dependence or self-persistence – into what belongs to the actual self-persistence, which will be the driving force behind the scarring effects of overskilling, and the part that is due to differences among employees related to human capital.

Allowing for different types of unobserved heterogeneity

We have already pointed out that unobserved heterogeneity was an issue with respect to obtaining accurate estimates of the state dependence associated with overskilling. It remains an issue, as we want to distinguish between the *time-varying effects* and *fixed-effects* of the variables affecting overskilling, its self-persistence and its outcomes. By *fixed effects* we mean the correlation between overskilling and time-varying variables that is unchanged through time, while *time-varying effects* refer to the correlation that does vary through time. We account for these effects in our model by using a random effects specification augmented by Mundlak (1978) corrections, which consist of including a fixed effect for each time-varying variable of the model via the inclusion of individual sample means of these variables in the model. The incorporation of these extra variables enables us to control for the correlation between the random effects (capturing the individual-specific unobserved heterogeneity) and the explanatory variables of the model.

Scenarios of skill mismatch experience

In order to get a precise understanding of the effect of past overskilling mismatch on both the probability of being currently overskilled and on the level of current hourly wages, we introduce the variety of possible time scenarios that may be faced by employees. This allows us to provide interesting comparisons across employees regarding their differential experiences in terms of overskilling mismatch. As mentioned above, we introduce three lags of the dependent variable in the estimation of the probabilities to be currently mismatched. This enables us to compare the estimated probabilities for employees, based on the four past overskilling profiles we have

⁶ The instruments we use in the model are those defined under the so-called Mundlak correction. They include the sample individual-averaged time varying variables of the model.

described above. We derive the probabilities of being currently mismatched under all four scenarios. The presentation of these scenarios and the comparisons between them provide more information than we would get by simply interpreting the marginal effects (see box 1) of each lagged dependent variable. The marginal effects provided in the tables for a given lag indicate the effect on the probability of being overskilled of a change in the value of this lag from 0 to 1 (going from well matched in the given wave to being overskilled), with the value of the other lagged dependent variables set to their sample mean, which is a number between 0 and 1. In other words, these marginal effects give the change in probability for the average employee in the sample, while the scenarios contribute to isolating specific cases of individual histories. We derive the probabilities attached to each of these scenarios for each level of educational achievement. Hence, our methodology enables us to compare the scarring effect of overskilling mismatch across each educational pathway. Using the estimated probabilities attached to each scenario, we can gain a deeper understanding of the self-persistence of overskilling across educational pathways. We carry out further analysis of the various scenarios in that part of the estimations where we investigate the effect of self-persistent overskilling mismatch on hourly wages.

The effect of past overskilling on wages

In the second part of the multivariate analysis, we investigate whether wage penalties are influenced by self-persistence in overskilling mismatch. Do we observe wage penalties due to overskilling? If we do, how long does it take for an individual to recover from the penalties? What is the size of *scarring*? That is, to what extent will the self-perpetuation of overskilling also cause the self-perpetuation of lower wages? We want to know whether scarring is prevalent for all educational levels and, if not, which educational pathway leads to the most scarring and which to the least. Is this phenomenon mostly concentrated at both ends of the education spectrum, with school dropouts and university graduates being more exposed to this risk? Does the sequence of overskilling matter, or is it only the number of overskilling occurrences that matters when it comes to overskilling wage penalties?

In an attempt to answer these questions, we estimate two categories of models. The first model consists of estimating the conditional mean wages using a conventional Mincer earnings function and utilising the panel structure of the Household, Income and Labour Dynamics in Australia Survey data. We estimate the earnings equation using a random effects model augmented with a Mundlak (1978) correction as described above. The dependent variable is the logarithm of hourly wages. The lasting effect of overskilling mismatch on the log of hourly wages is captured by a set of four dummy variables indicating the employees' skill mismatch status for the current and the past three waves of the survey. The estimates for each of these 'lagged overskilling' variables indicate the magnitude of the wage penalties caused by previous spells of overskilling mismatch. A test of significance on these coefficients allows us to determine how long a spell of overskilling can be expected to have an effect on the current wage, everything else held constant. We use other relevant personal and workplace characteristics as control variables in the estimations, including age, marital status, remoteness, migration background, firm size, tenure with current employer, industry dummy variables and so on. Following the interpretation of the results in terms of the employees' histories used in the model in step 1, we derive the estimated hourly wage for the 16 possible scenarios implied by the four dummies on overskilling included in the equation. Based on these estimates, we can analyse whether wage penalties differ across

educational pathways, with the objective of finding out whether or not overskilling self-persistence has an effect on wages.

Box 1 Coefficients and marginal effects in the probability model

In conventional linear estimation it is customary to report a coefficient to represent the estimated association between each independent variable and the dependent variable. The sign of each coefficient has a ready intuitive interpretation. For a positive sign, the data suggest that there is a positive association between the dependent variable and the specific independent variable. A negative sign suggests a negative association. Simply put, a positive coefficient would suggest that subjects with high values of the independent variable are more likely to have high values also of the dependent variable. An example of a positive association is that between education and income: if we pick a person at random from our sample and they happen to have a university degree, we are more likely also to have picked someone with an above-average income. (Note that this is a probability statement. It is indeed possible that we may pick someone with a degree and a very low income, as there are people with degrees who have below-average incomes. However, we can be sure that if we keep on repeatedly looking at people with degrees, we will end up with people who have a higher-than-average income.) The linear estimation model lends itself to further interpretation. If both dependent and independent variables are measured in clearly understood units and have a relationship that we believe to be constant across the range of values of these variables, then the coefficient has a clear quantitative interpretation: a one-unit increase in the independent variable is associated with an increase in the dependent variable that equals the value of the coefficient. Simply put, if the coefficient of experience in the workforce measured in years (that is, the number of years enter as the independent variable in the right-hand side) in the estimation of hourly wage (measured in Australian dollars) is estimated to be 1.5, the result at hand says that if we pick a group of workers from our data with ten years experience and another group with 11 years experience, the latter will be paid AUD1.5 more per hour. Where the variables have been measured in logs, which is often the case with earnings estimations, the coefficients can be interpreted as elasticities, that is, as the relationship between two percentage changes.

Unfortunately, when we need to use a non-linear model, such as a probability model, coefficients lose their intuitive interpretation. The implication of this non-linearity is that the estimated coefficients associated with each variable do not provide us with a number that can be readily interpreted in terms of the units in which the two variables are measured, as was the case in the linear regression models. The size of the coefficient and its association with the dependent variable actually change, depending on the value of the independent variable. The estimates themselves have little interpretative value beyond their sign.

To overcome this problem we calculate the so called 'marginal effects' for each of the estimated coefficients. Note that this is a calculation that contains no new information over and above what has been used to derive the original set of coefficients; it just translates these coefficients into a metric that has an intuitive interpretation. In the context of the work we present here, when we estimate the probability that someone is overskilled, the marginal effect of each independent variable states how the estimated probability of overskilling changes per a unit change in the independent variable. Using the independent variable 'experience' measured in years in this estimation, if we find a marginal effect of -0.02, we can state that each additional year of experience reduces the probability of overskilling by 2%. Similarly, for categorical variables, the marginal effect measures the difference in the estimated probability due to the change in the categorical variable from the value of 0 to the value of 1. In general, when we derive marginal effects, we set the value of all variables in the model to their sample means and vary only the value of the variable in question. However, this is not necessary, as we may wish to target our prediction and focus on individuals with particular characteristics.

The estimation of the wage equations gives us a general understanding of the average relationship between the persistence of overskilling mismatch and wages across educational pathways. However, it does not give us the full story, because wages vary widely within each educational level. Taking university graduates as our example, we would want to know whether the wage penalties of overskilling are greater for the better-paid or for the less well-paid graduates. In order to inform this question, we carry out an additional set of estimations, which complete the picture by providing information on the relationship between overskilling and wages at different points across the wage distribution. These additional models use the method of quantile regression to estimate the relationship between wages and overskilling at the three points of the distribution of wages that correspond to the three quartiles thresholds (that is 25, 50, and 75%). This estimation methodology enables us to analyse the impact of the covariates, in particular those related to individuals' history of overskilling, on both the location and scale parameters of the model. As a result, we can compare the impact of previous overskilling, not only for a given educational pathway at different locations on the wage distribution (quantile), but also across educational levels for a given quantile. For instance, quantile regression on wages enables us to know the extent to which previous experience of overskilling may affect VET graduates differently, depending on whether they belong to the lowest quartile of the (VET) wage distribution or the higher quartiles. At the same time, we can also compare the effect of past overskilling across educational levels for, say, individuals situated at the lowest quartile of the wage distribution within their educational level.

Regression results

We start our analysis by examining the scarring effect of skill mismatch. We do this by estimating the probability of employees reporting being overskilled at the current wave of the Household, Income and Labour Dynamics in Australia Survey interview. We use a dynamic specification of the random effects probit model, with separate models estimated for all educational pathways and comparisons made. We then study the scarring effect of the persistence of skill mismatches on employees' hourly wages by educational level.

The effect of past overskilling on future overskilling: probit analysis

All of the results in this section are generated using the dynamic random effects panel probit specification discussed in the methodology section. We look at the determinants of the probabilities for employees to report being currently overskilled in their job, with a particular emphasis on the effect of past experiences of skill mismatch on these probabilities. The impact of previous spells of overskilling is captured through the inclusion of lagged dependent variables in the equations. We first estimate the model on the whole sample of employees of working age (excluding full-time students), controlling for their highest level of completed education through a set of dummy variables. Results are presented in table 9. The coefficients and marginal effects given in table 9 for these educational variables must be interpreted with reference to the category that was omitted; namely, individuals who did not complete high school. Using the full sample, we test two sets of models, including, respectively, three and five lags in order to determine how far back a spell of overskilling still contributes to current overskilling propensities. The results in table 9 suggest that the model with three lags is superior to that with five lags, so that subsequent analysis is carried out exclusively with the three-lag model.

We then estimate five separate models on the sub-samples of employees by educational level; that is, university graduates, diploma, VET certificates III/IV, Year 12 completers and individuals who did not complete high school. For these educational levels we analyse the extent of the state dependence of overskilling and the depth of the scarring effect. We provide further comparisons within and between educational levels through the analysis of the different scenarios of past skill mismatch.

Table 9 displays the results of the three- and five-lag models obtained with the full sample and controlling for educational level. For each model, two sets of results are displayed. The first two columns report the estimated coefficients and marginal effects (ME) of a basic pooled cross-section probit model without a panel specification. These are used as a reference, while the next two columns of each model represent the dynamic panel random effects (RE) probit estimation, whose estimates are more consistent and efficient. We only discuss the random effects model results.

Table 9 Dynamic probit estimations of overskilling: all education levels combined

	Model I with three lags				Model II with five lags			
	Pooled probit model		Woodridge RE model		Pooled probit model		Woodridge RE model	
	Coef.	ME	Coef.	ME	Coef.	ME	Coef.	ME
	(z)	(z)	(z)	(z)	(z)	(z)	(z)	(z)
Overskilling at t-1	0.785 (30.32)	0.227 (26.24)	0.687 (21.93)	0.186 (17.18)	0.768 (20.10)	0.212 (16.94)	0.739 (16.96)	0.198 (12.89)
Overskilling at t-2	0.570 (21.64)	0.156 (19.16)	0.505 (17.75)	0.130 (14.97)	0.536 (13.69)	0.138 (11.93)	0.527 (13.19)	0.133 (11.16)
Overskilling at t-3	0.429 (16.42)	0.113 (14.88)	0.302 (10.25)	0.074 (9.27)	0.295 (7.35)	0.071 (6.76)	0.293 (7.07)	0.069 (6.56)
Overskilling at t-4	-	-	-	-	0.270 (6.77)	0.064 (6.27)	0.263 (6.24)	0.061 (5.89)
Overskilling at t-5	-	-	-	-	0.286 (7.43)	0.068 (6.88)	0.223 (5.11)	0.051 (4.84)
Complete school	-0.075 (-2.12)	-0.017 (-2.18)	-0.396 (-1.40)	-0.075 (-1.67)	-0.063 (-1.18)	-0.013 (-1.22)	-0.017 (-0.03)	-0.004 (-0.03)
Certificates III/IV	-0.119 (-3.65)	-0.027 (-3.78)	-0.394 (-2.08)	-0.078 (-2.37)	-0.082 (-1.75)	-0.018 (-1.80)	-0.434 (-1.18)	-0.081 (-1.36)
Diplomas	-0.116 (-2.73)	-0.026 (-2.88)	-0.510 (-1.60)	-0.090 (-2.11)	-0.100 (-1.65)	-0.021 (-1.73)	-0.033 (-0.05)	-0.007 (-0.05)
University degrees graduates	-0.204 (-6.08)	-0.046 (-6.34)	-0.621 (-1.84)	-0.122 (-2.09)	-0.162 (-3.36)	-0.034 (-3.47)	-0.798 (-1.20)	-0.148 (-1.38)
Sample size	21 185				11 042			

Note: Dependent variable is current overskilling. The Wooldridge method has been used to control for initial conditions. We report coefficients and marginal effects (ME), with their z-statistics in brackets. HILDA waves 1–9. The unit of observation is person years. Full estimation results are reported in appendix A.

There are two main results in table 9. First, all five lags of overskilling are significant and their marginal effects reduce in size for lags further back in the past. The message is that overskilling is a seriously long-lasting phenomenon, a new result in the literature of skills mismatch. Second, although the individual lag variables retain their significance under the five-lag specification, the whole model does not fare well. The vital statistic that represents the degree of panel information in the results becomes insignificant and many other variables show that the panel variation is completely overcome by their cross-section variation. The reason for this is that the dataset is too short, with only nine waves of information to handle a model with five lags. A model with five lags uses six waves of information simultaneously (one for the present and five for the lags), so a dataset with nine waves and five lags is equivalent to a three-wave panel without lags. The examination of the three-lag model shows that the estimation results are robust and that the data can handle the three lags. The inadequacy of the data in the five-lag model and the well-behaved results in the three-lag model suggest that we should continue the analysis using the three-lag model.

Focusing on the three-lag model in table 9, we examine the results. The coefficients of the lagged dependent variables are strongly significant, which indicates that past experiences of overskilling mismatch influence present overskilling mismatch probabilities. As expected, the more distant in the past is the occurrence of overskilling, the less is its impact on the current probability, with marginal effects diminishing as the order of the lag increases. Indeed, if we compare the marginal effects for employees who were overskilled at the previous period with those who were not, the former are expected to have an extra 18.6% chance of being

mismatched one year later than the latter (table 9, model with three lags, overskilling at t-1), 13% two years later, and 7.4% three years later. Note that as these are results on the full sample, they average the effect of overskilling over all educational levels, making them of limited use for our analysis. This observation is supported by the fact that education dummies are highly significant in the estimation. The marginal effects of the education dummies in table 9 (model 1) indicate that all educational categories above school non-completers are less likely to be overskilled. Year 12 graduates are 7.5% less likely to be overskilled, although the statistical significance of this estimate is relatively weak. VET graduates are 7.8% less likely, diploma graduates 9% less likely, and university graduates 12.2% less likely to be overskilled.⁷

Our next step is to estimate the three-lag model by educational level using the dynamic random effects probit model to highlight the differences vis-a-vis the scarring effect of overskilling. Results are in table 10. The effect of previous overskilling mismatch on present overskilling is positive and diminishes over time.

Table 10 Dynamic probit estimations of overskilling by educational level

	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Coef. (z)	ME (z)	Coef. (z)	ME (z)	Coef. (z)	ME (z)	Coef. (z)	ME (z)	Coef. (z)	ME (z)
Overskilling at t-1	0.766 (11.62)	0.150 (7.37)	0.814 (7.72)	0.210 (5.58)	0.714 (10.97)	0.198 (8.57)	0.610 (8.00)	0.188 (6.98)	0.675 (11.67)	0.231 (10.76)
Overskilling at t-2	0.333 (5.21)	0.053 (4.23)	0.575 (5.99)	0.137 (4.80)	0.574 (9.79)	0.153 (8.11)	0.488 (7.10)	0.148 (6.43)	0.575 (11.17)	0.195 (10.51)
Overskilling at t-3	0.278 (4.26)	0.043 (3.62)	0.392 (3.95)	0.088 (3.43)	0.348 (5.71)	0.087 (5.06)	0.287 (4.10)	0.083 (3.89)	0.298 (5.49)	0.098 (5.29)
Sample size	6736		2150		4819		3041		4439	

Note: Dependent variable is current overskilling. The Wooldridge method has been used to control for initial conditions. We report coefficients and marginal effects (ME), with their z-statistics in brackets. HILDA waves 1–9. The unit of observation is person years. Full estimation results are reported in appendix A.

The results shown in table 10 highlight interesting variations across graduates from different educational levels. Overskilling is the least persistent among university degree holders. Those experiencing skill mismatch one year earlier are 15% more likely than their well-matched comparators to be overskilled a year later. By contrast, the difference in probability associated with the first lag goes up to 23.1% for high school non-completers (which is about 54% higher than the university graduates). Looking three years back, the differences between university graduates and all others is even more striking. A spell of overskilling three years down the track keeps altering the probability of being currently overskilled, even for the university graduates. However, it only increases their overskilling probability by 4.3%, which is about half of the effect observed for all other educational levels (ranging from 8.3% for Year 12 graduates and 9.8% for those who did not finish high school). Altogether, our results show that, while persistence is still strong for all educational levels after three years, university graduates show the lowest self-persistence among all other workers, and those who do not finish high school show the highest. The profile of overskilling mismatch self-persistence is fairly similar for all other educational

⁷ We do not pursue the results in this model and their interpretation further and we do not suggest the reader does either, as it is clear from the subsequent estimations that estimating the whole sample with all education categories produces biased results. Splitting the sample by education level shows that most parameters are education-specific.

pathways, with estimates around 18 to 20% for the first lag, 13 to 15% for the second lag and around 9% for the third lag.

Self-persistence of mismatch: the effect of past overskilling mismatch on current overskilling mismatch

One of the main interpretational disadvantages of the results we presented in table 10 is that each estimate is calculated for the average person. While this is usually a good way to think of estimates (it makes them comparable), it confuses our interpretation when we use lagged dependent variables in the right-hand side of the equation. Although the results are mathematically correct, they do not necessarily tell us exactly what we want them to tell us. We explain this by way of example. In table 10, the marginal effect of the second overskilling lag takes the value of 5.3%. What this tells us is that if we take a person who has all their characteristics set at the average person's level (for continuous variables) and at the reference category level (for categorical variables), then the probability that this 'average' person will be overskilled in the present period is 15.3% higher if they were overskilled two years back, than if they were well matched two years back. However, note that, since we are dealing with the second lag, this also assumes that the previous lag and the subsequent lag variables are also set at their average values. While this may be mathematically correct, it is very hard to interpret. Is there an intuitively interpretable meaning for a lagged effect that is conditional on another lagged effect that takes the mean sample value? To overcome this problem we use the estimation results to calculate a number of scenarios, which we present in table 11. The objective of table 11 is to make comparisons between individuals with different past overskilling experiences and to show these comparisons by educational level. For each scenario, we compute the associated probabilities of being currently mismatched. This technique allows us to compare the effect of past spells of overskilling between two employees with the same educational pathway, and to establish whether or not self-persistence of overskilling mismatch deepens as time goes by, and, to a degree, also to form an opinion on whether this deepening differs by educational level.

Table 11 Estimated overskilling probabilities by educational level and scenario of overskilling

Scenario	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Predicted probability	Overskilling difference	Predicted probability	Overskilling difference						
000	0.046		0.063	0.084			0.110		0.129	
100	0.180	0.133	0.236	0.174	0.253	0.169	0.269	0.159	0.324	0.195
110	0.280	0.100	0.443	0.207	0.464	0.211	0.449	0.180	0.547	0.223
111	0.380	0.100	0.598	0.155	0.602	0.138	0.563	0.114	0.661	0.114

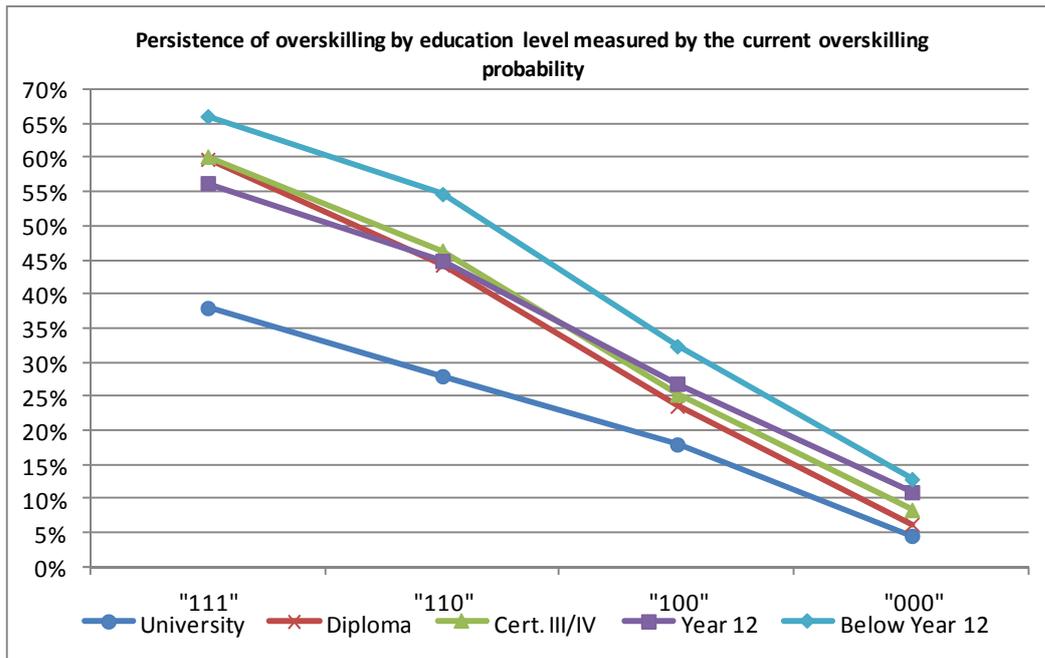
Note: Probabilities are based on the overskilling estimations presented in table 10.

Table 11 contains four scenarios which correspond to a typically deepening overskilling mismatch experience in the last three years, represented by the estimated three-lag variables. The first scenario is when someone has not been overskilled in the last three years. The second is when they were overskilled only in the previous year. The third is when they were overskilled in the previous two years, and the fourth one is when they were overskilled in all three past years. The intuition of a deepening overskilling experience is obvious. There are also other overskilling combinations and we have experimented with them to find that the scenarios we present here are a good representation of the overskilling self-persistence estimation. For convenience of presentation, we shall identify these scenarios with a set of binary codes, with '1' meaning that

the employee was overskilled in the period considered and '0' if they were not. We use the three-lag model estimates to compute the probabilities attached to each scenario. Hence, the scenario '000' stands for an employee who has been well matched during the last three years. The scenario denoted by '100' represents an individual who was overskilled in the previous year but who was well matched in the two years before that, and so on. We compare probabilities for the scenarios '000', '100', '110', and '111' and we present the results alongside their change between each year for each educational level. A university graduate who was never before overskilled (the 000 type) has a 4.6% probability of becoming overskilled. By contrast, another university graduate who was overskilled in all three preceding years (the 111 type) has a 38% probability of becoming overskilled. Note that the probability difference of 33.4 (= 38 - 4.6) is conditional on both graduates having the same (average graduate) characteristics. We do not have any comparison at hand from the literature to evaluate the size of this probability, but a 33.4% probability to continue being underutilised appears high, and overskilling mismatch appears to be highly self-persistent. Still, university graduates produce by far the lowest overskilling mismatch self-persistence estimate. The highest estimate comes from those who did not finish high school, where the 000 and 111 types have a probability of 12.9 and 66.1% respectively, with a probability difference of 53.2 (= 66.1 - 12.9). The probability differences can be interpreted as the degree of self-persistence of overskilling mismatch, and they start with 33.4% for university graduates, 53.5% for VET diploma graduates, 51.8% for certificates III/IV graduates, 45.3% for Year 12 graduates, and 53.2% for those without high school completion. It is very clear that the main difference lies between university graduates and the rest. It should be noted that, while we can make comparisons within each of the educational level groups, especially in terms of the causal effect of past overskilling on future overskilling, we should not make any statistical comparisons about the absolute levels of overskilling mismatch between different educational levels, as the estimates we present do not control for any differences between any two groups. What we can compare is the way in which the different types of experiences vary by education.

Figure 1 summarises the results in table 11. It is worth noting that individuals of the type 000 are relatively similar in their overskilling probabilities, and that the deepening of overskilling (the self-persistence of overskilling mismatch) is very different for individuals of the 111 type.

Figure 1 Self-persistence of overskilling mismatch



The effect of past overskilling on wages: linear analysis

This section estimates the effect of overskilling on wages. There already exists sufficient and robust evidence on the deleterious effect of current overskilling mismatch on the current wages of the overskilled. This evidence shows that being overskilled is associated with lower wages at the population level, and becoming overskilled causes lower wages at the individual level. This section extends the evidence by investigating how having been overskilled in the past may have a lasting effect on wages. To this purpose, we present two sets of earnings models, a linear and a quantile model, both estimated by educational level, in keeping with the rest of this analysis. The first set of regressions estimate current wages (at time t) and use lagged overskilling variables to examine whether past overskilling mismatch (at time $t-1$, $t-2$ and $t-3$) may have an influence on current wages. These regressions provide a mean effect; that is, they tell us how the mean wage for all people who belong to each educational category may be influenced by overskilling. We present the linear regression results in table 12.

Table 12 shows some interesting findings. First, educational level matters vis-a-vis the effect of past overskilling on present wages. University graduates have to live for many years with the effect of their past overskilling. Even overskilling that occurred three years back appears to reduce present wages by 3.7%. Our evidence suggests that the effect of overskilling mismatch is both serious and long-lasting, considering that the loss refers to annual income and can be cumulative. It is worth noting that the overall effect of overskilling will consist of all annual effects put together. Evidence on diploma graduates is mixed and should be considered in the light of the fact that this educational category is the least populous in the dataset, so that our results may suffer small-sample problems. Note that the cross-section results (pooled ordinary least squares) suggest that, on average, past overskilling is associated with lower wages, but these results do not control for individual unobserved heterogeneity. Panel results do not show any effect of overskilling on wages, but this may be because there are not enough people in the diploma category who change their overskilling status. Estimation is not sufficiently informative

about the difference between cross-section and panel evidence for this category. By contrast, evidence on certificate III/IV graduates is relatively clear-cut: there is very little evidence to suggest that past overskilling has any effect on the present or future wages of VET graduates.

Table 12 Wage estimation by education level (log hourly wage)

	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Coef. (t ratio)		Coef. (t ratio)		Coef. (t ratio)		Coef. (t ratio)		Coef. (t ratio)	
	<i>Pooled OLS</i>	<i>Panel REM</i>	<i>Pooled OLS</i>	<i>Panel REM</i>	<i>Pooled OLS</i>	<i>Panel REM</i>	<i>Pooled OLS</i>	<i>Panel REM</i>	<i>Pooled OLS</i>	<i>Panel REM</i>
Overskilling at t	-0.065 (-3.87)	-0.030 (-1.92)	-0.072 (-3.20)	-0.029 (-1.41)	-0.037 (-2.82)	-0.009 (-0.68)	-0.030 (-1.84)	-0.037 (-2.28)	-0.051 (-3.34)	-0.010 (-0.73)
Overskilling at t-1	-0.064 (-4.08)	-0.041 (-2.56)	-0.044 (-2.11)	-0.008 (-0.38)	-0.021 (-1.57)	-0.004 (-0.40)	-0.023 (-1.53)	-0.020 (-1.29)	-0.041 (-3.18)	-0.010 (-0.82)
Overskilling at t-2	-0.067 (-4.89)	-0.031 (-2.86)	-0.051 (-2.55)	-0.016 (-0.99)	-0.003 (-0.24)	0.011 (1.11)	-0.024 (-1.49)	-0.023 (-1.62)	-0.015 (-1.27)	0.008 (0.72)
Overskilling at t-3	-0.072 (-5.03)	-0.037 (-3.18)	-0.034 (-1.66)	0.018 (1.18)	-0.022 (-1.64)	-0.006 (-0.57)	-0.042 (-2.68)	-0.024 (-1.84)	-0.050 (-3.94)	-0.024 (-2.37)
Sample size	6237		2006		4495		2754		4074	

Note: Dependent variable is the log of hourly wages. Coefficients and t ratios are reported. Full estimation results can be found in appendix A. The unit of observation is person years. HILDA waves 1–9. REM denotes random effects linear estimation with Mundlak corrections. OLS = ordinary least squares.

Moving to workers without post-school qualifications, we see that there is some evidence from the panel analysis for Year 12 graduates that suggests the presence of long-lasting overskilling wage penalties. The effect of overskilling increases with time, indicating that this may be a problem area, with strong self-persistence within a group that is generally low paid. A surprisingly strong panel coefficient for overskilling in the third lag only appears for those who did not complete high school. The statistical significance for both the pooled and the panel coefficients is too strong for us to decide to ignore this result, but its interpretation is difficult, as no other panel coefficient is even remotely statistically significant (with all t-values below 1 and all coefficients -0.01 or less, this result cannot be interpreted with ease).

The main result from the linear estimations is that there is clear self-persistence on overskilling mismatch only for graduates, and there are some indications of isolated self-persistence effects that may be difficult to interpret. These results allow us, in the next section, to study self-persistence and its scarring effect in more detail. For now we turn to the quantile regression results.

The second set of regressions refines the results of the first set of linear estimations presented in table 12 by estimating the effect of overskilling on different parts of the wage distribution for each educational category. There is good *a priori* reason for introducing this extension, which we show to be vindicated in the empirical results. We know that each educational category contains a broad distribution of wages, especially so for the better-paid educational categories such as university graduates and diploma graduates. It is important to know whether overskilling mismatch is more or less harmful to the wages of those who are better or worse paid in their own educational category. Finding, for example, that the better-paid university graduates are also those who face the maximum overskilling wage risk would make more sense than finding that the worse-paid are the ones who also bear most of the risk. Conventional human capital theory would predict the former case, as higher earnings should compensate for increased risk (Berkhout, Harthog & Webbink 2010), while the latter case would need some kind of institutional

explanation for why it is so, such as dual labour market theory (Doeringer & Piore 1971). We present the quantile regression used to obtain these results in the next section.

The effect of past overskilling on wages: quantile regression

We complement our linear regression results with further estimations of wage effects, focusing on different points of the wage distribution using quantile regression. Table 13 presents the estimated coefficients of a quantile regression model evaluated at each quartile of the distribution of wages by educational level. In addition to table 13, we present a series of graphs (figures 2–6) of the estimated association between overskilling mismatch and wages across the wage distribution and for each educational pathway. Each graph presents the estimates of the wage effect of each overskilling mismatch included in the model, namely ‘currently overskilled’, ‘overskilled a year ago’, ‘two years ago’ and ‘three years ago’.

We first interpret the results in table 13. It contains the quantile estimates at the following three points of the wage distribution of each education level: 0.25, 0.50 and 0.75. The first result is that for university graduates; the higher the wage they receive, the worse the wage penalty they suffer from self-persistence. The lowest quartile has an estimated penalty of 4.4% and the highest a penalty of 8.6%. As these estimates have been derived within the same estimation, they are comparable. In the case of diploma VET graduates, the use of quantile regression offers considerable added intuition. Those who are at the highest income quartile look a lot like university graduates (that is, they have a strong association between overskilling and wages), and those who are at the bottom income quartile look a lot like certificate III/IV graduates (that is, they do not show any association between their wage and their overskilling). Quantile regression results add further confirmation to the finding that the wages of VET graduates with certificate III/IV are not influenced by their overskilling status. Overskilling is not very prevalent and where it occurs it is of little pecuniary consequence. Finally, for those without any post-school qualifications, there are no further insights to be gained by using quantile regression.

The first striking result from the quantile estimations is that the significance levels of the lagged overskilling variables vary widely by educational level and across wage quartiles. For instance, all lags are significant for university graduates for all quartile thresholds. This means that for all quartiles, the wages of university graduates are affected by experiences of overskilling at least as far back as three years. By contrast, only current overskilling seems to affect certificate III/IV graduates, with all lags of overskilling being non-significant for this educational level. For Year 12 graduates, overskilling has a complex association with wages; however, the relatively small sample for this educational category may have contributed to this. Looking at each educational level separately, we can see interesting differences across quartiles. Indeed, the currently overskilled university graduates experience smaller wage penalties at the lower end of the wage distribution, with an estimated coefficient of -0.043 at the first quartile, against -0.077 and -0.070 at higher quartiles. For more distant experiences of overskilling, the university graduates suffer similar penalties across quartiles with the exception of the third lag of overskilling, where the penalties increase with the quartiles from -0.044 (q25) to -0.086 (q75). Altogether, a university graduate who has accumulated three years of mismatch and remains currently overskilled – and who belongs to the lowest quartile of the wage distribution – would incur a total wage penalty of 21.1% compared with a well-matched individual. The same graduate in the next quartile would total a loss of 24.8%. The figure reaches 29.1% if the employee is in the highest quartile of the wage distribution.

Table 13 Quantile wage estimation by education level (log hourly wage)

	University degrees			Diplomas			Certificates III/IV			Only completed school			Did not complete school		
	coefficient (t-ratio)			coefficient (t-ratio)			coefficient (t-ratio)			coefficient (t-ratio)			coefficient (t-ratio)		
	q25	q50	q75												
Overskilling at t	-0.043 (-1.95)	-0.077 (-4.47)	-0.070 (-4.19)	-0.042 (-1.95)	-0.080 (-3.08)	-0.063 (-2.24)	-0.041 (-3.09)	-0.036 (-2.38)	-0.041 (-2.19)	-0.015 (-0.76)	0.005 (0.26)	-0.028 (-1.52)	-0.022 (-2.00)	-0.031 (-2.85)	-0.044 (-2.85)
Overskilling at t-1	-0.054 (-3.91)	-0.040 (-3.16)	-0.069 (-3.72)	-0.032 (-1.98)	-0.037 (-1.34)	-0.038 (-1.26)	-0.001 (-0.07)	-0.018 (-1.15)	-0.025 (-1.05)	0.007 (0.26)	-0.014 (-0.80)	-0.018 (-0.64)	-0.046 (-3.39)	-0.024 (-1.71)	-0.021 (-1.30)
Overskilling at t-2	-0.070 (-3.69)	-0.065 (-3.24)	-0.066 (-3.34)	-0.017 (-0.66)	-0.029 (-1.52)	-0.054 (-2.04)	-0.003 (-0.21)	-0.005 (-0.35)	-0.013 (-0.86)	-0.040 (-2.22)	-0.025 (-1.54)	-0.030 (-1.34)	-0.016 (-1.06)	-0.033 (-2.43)	-0.019 (-1.18)
Overskilling at t-3	-0.044 (-2.23)	-0.066 (-3.16)	-0.086 (-5.30)	-0.014 (-0.40)	-0.012 (-0.55)	-0.074 (-3.30)	0.000 (-0.03)	-0.016 (-0.80)	-0.024 (-1.22)	-0.018 (-1.63)	-0.039 (-2.14)	-0.048 (-2.41)	-0.025 (-1.65)	-0.047 (-3.91)	-0.054 (-4.49)
Sample size	6237			2006			4495			2754			4074		

Note: Dependent variable is the log of hourly wages. Coefficients and t-ratios in brackets are reported. HILDA waves 1–9. The unit of observation is person years.

For diploma holders, the results must be viewed with caution because of the relatively small sample size. For this educational level, most of the wage penalty is observed at the current period of mismatch. Some between-quartile differences emerge for diploma graduates, such that the wage does not seem to be affected significantly by overskilling incurred beyond two years back in the first quartile, while it does in the highest quartile. Certificate holders seem to incur a wage penalty only when currently overskilled, and mismatch persistence seems to produce no long-term wage penalty. Moreover, current overskilling seems to produce the same wage penalty no matter on which quartile we focus. Similar to diploma graduates, the Year 12 school graduates form a relatively small part of the sample and the results may not be as robust as for the other educational categories. They should thus be interpreted with moderate caution. According to our results, mismatch occurring in a more distant past affects wages, while more recent spells do not lead to significant wage penalties. The pattern of wage penalties for workers who have not completed high school (Year 11 or less) is similar to that of university graduates, in that there are strongly persistent wage penalties at some points of the wage distribution. Employees right in the middle of the distribution (median, q50) incur wage penalties, with all lags of overskilling bringing the wage down by up to 13.5% for those who have been overskilled in the past three years and who are currently mismatched. At the lowest quartile, the coefficient of the second lag is not significantly different from zero and only the current and most distant overskilling experiences are associated with wage penalties for employees at the highest quartile. Altogether, whereas the probability of overskilling and its self-persistence are lower for graduates than for all other educational categories, the results of the wage estimations show that, when university graduates become persistently overskilled, they suffer heavy losses in terms of per-person wage penalties, especially for those at the higher end of the wage distribution. Certificate III/IV holders, by contrast, are more likely to be overskilled and experience persistence in overskilling than are university graduates, but the resulting wage penalties do not seem to give rise to long-term wage losses, in the sense that only current overskilling is associated with current wage losses.

Figures 1 to 5 illustrate the estimated impact of the current and past spells of overskilling on the hourly wage for the whole wage distribution of each educational level. Each graph displays the following two types of information.

- *OLS (ordinary least squares) information:* the dark straight dash line represents the estimated linear regression coefficient (OLS coefficient) for the spell represented. It is the mean effect of this overskilling mismatch on the log hourly wage. The dark straight dash line is surrounded by two straight dotted lines which represent the boundaries of the 95% confidence interval of this coefficient.
- *Quantile regression information:* the green irregular line represents the estimated effect of the spell at the corresponding quantile of the wage distribution. (The quantiles can be read on the x-axis.) The grey area above and below the irregular line represents the 95% confidence interval around the quantile estimates.

To get an idea of whether an estimate for a quantile of interest (belonging to the green line) is significantly different from zero, we only need to look at whether the 95% confidence interval around the quantile estimate contains zero or not. If it does not, we can accept the hypothesis that the estimate is significantly different from zero for the given quantile. For instance, looking at the first graph for the university graduates, which represents the estimated effect of a current spell of overskilling on their wages, the 95% confidence interval around the estimate for each quantile (along the green line) encompasses zero up to the second decile and does not persist further up the distribution. The estimate reaches a minimum between the sixth and eighth deciles and reaches the

value of almost -10%. As we can see in the second graph representing the effect of overskilling a year earlier, the estimates are smaller in absolute value, staying within the boundaries of the confidence interval of the OLS estimate. The estimates for the second lag are significantly different from zero for all quantiles and are very close to the OLS estimate, except for the lower end of the distribution, where the wage penalty is larger. The third lag produces significant estimates of wage penalty between the second and the ninth deciles. Outside these boundaries, the estimates are less reliable.

For certificate III and IV holders, the impact of current overskilling seems to affect wages with similar strength along the wage distribution, which confirms the results in the previous table, where the wage penalties were estimated at between 3.6 and 4%. The graphs for the other lags show mostly non-significant estimates across the whole distribution of wages.

As discussed earlier, the estimates for Year 12 school graduates are mostly insignificant with respect to current overskilling and its first lag. Wage penalties occur at the second lag in the lower end of the wage distribution up to the fourth decile. For employees situated in the higher end of the distribution, past overskilling produces no significant wage effect.

Like the university graduates, those in the 'Year 11 or less' educational category suffer from scarring, but the profiles of the scarring are quite different from the university graduates. Regarding the effect of current mismatch, the profile was downward sloping for the university graduates, indicating that employees at the higher end of the distribution suffered larger wage penalties. This is not the case for the Year 11 or less educational category, where the profile is smoother and the highest significant estimate is reached at the very bottom of the distribution. When overskilled in the previous year, the wage penalty is smaller in the middle of the distribution and becomes bigger at the extremities. Since the confidence intervals also widen, we cannot draw strong conclusions from this result.

Figure 2 Quantile estimates of wages and overskilling, university graduates

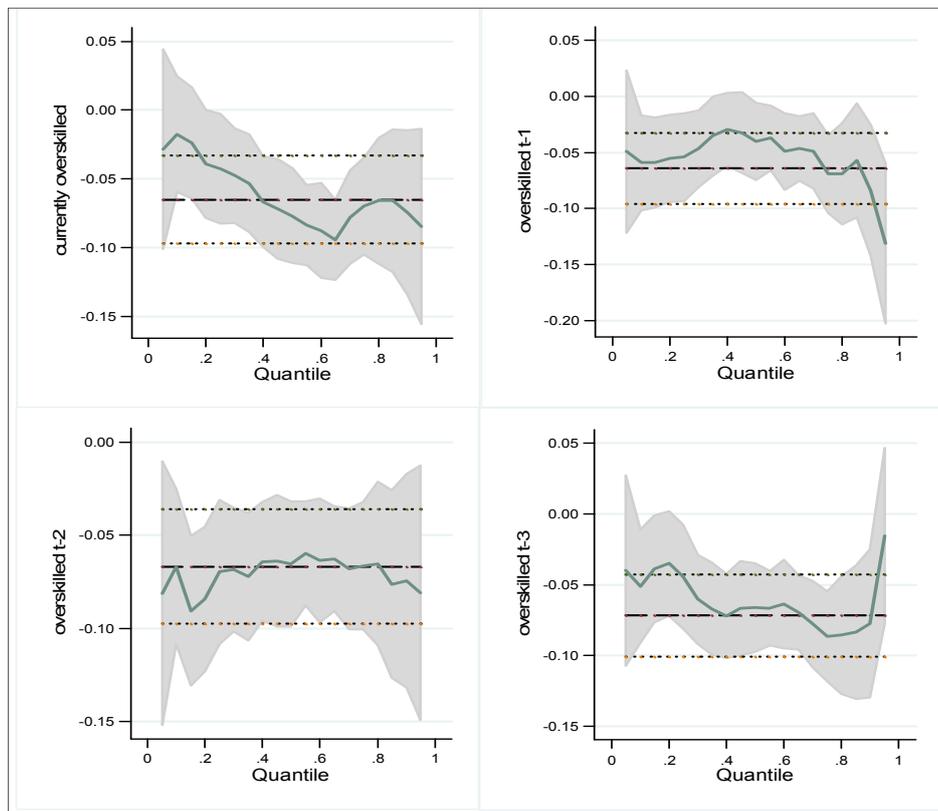


Figure 3 Quantile estimates of wages and overskilling, diploma

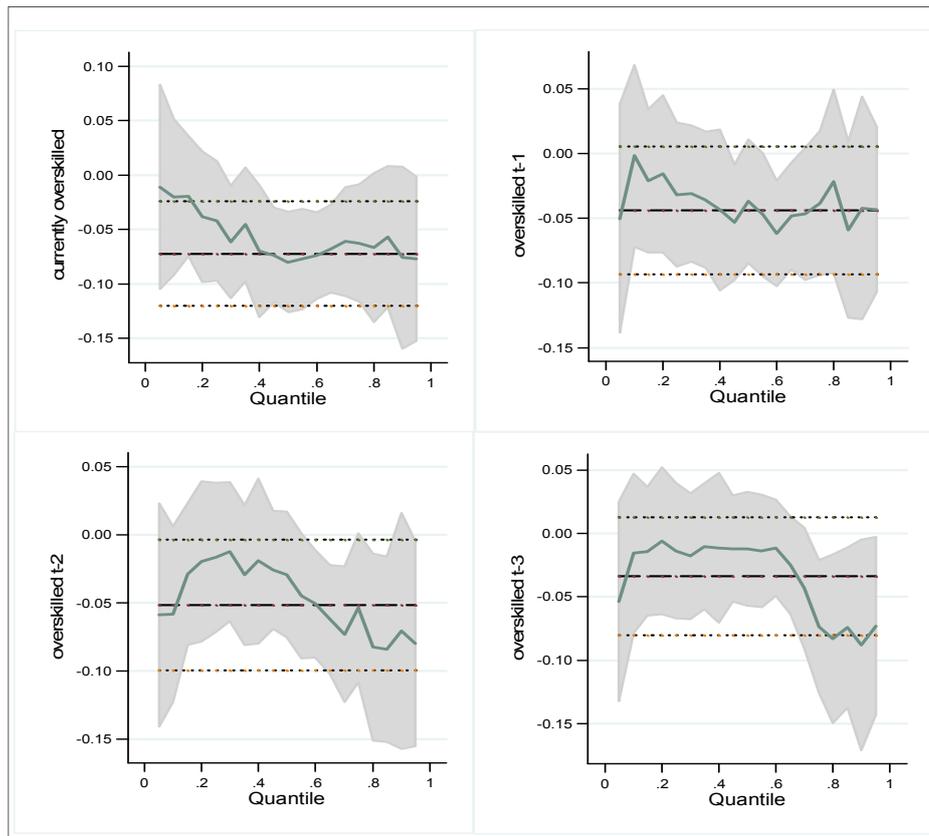


Figure 4 Quantile estimates of wages and overskilling, certificate III/ IV

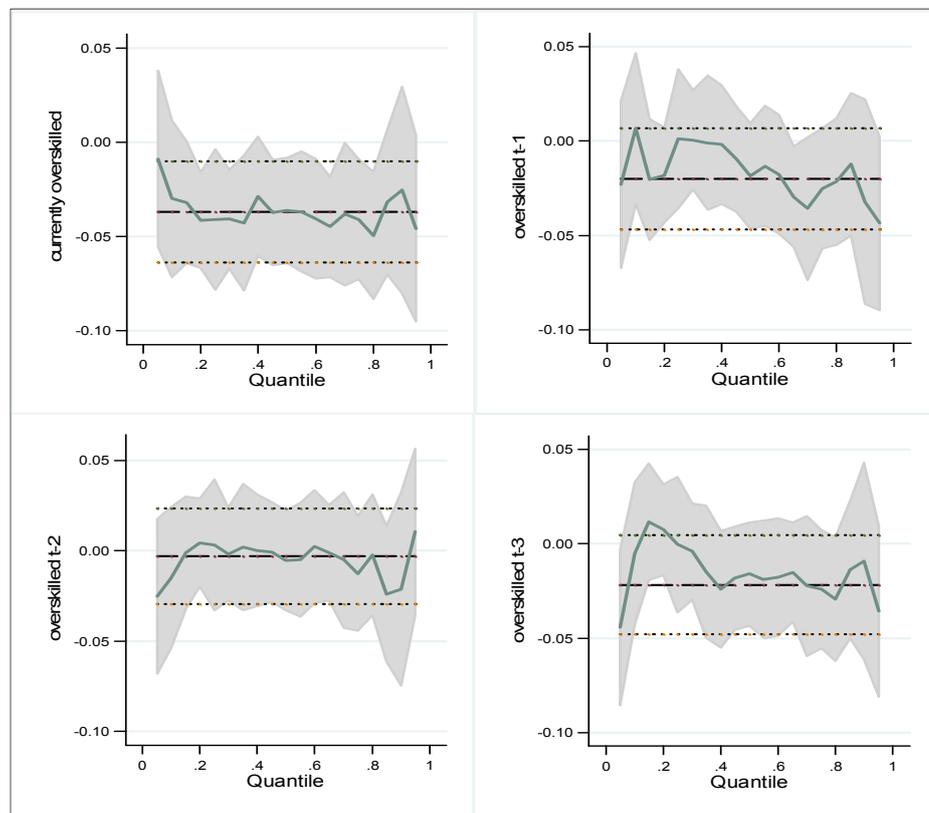


Figure 5 Quantile estimates of wages and overskilling, Year 12

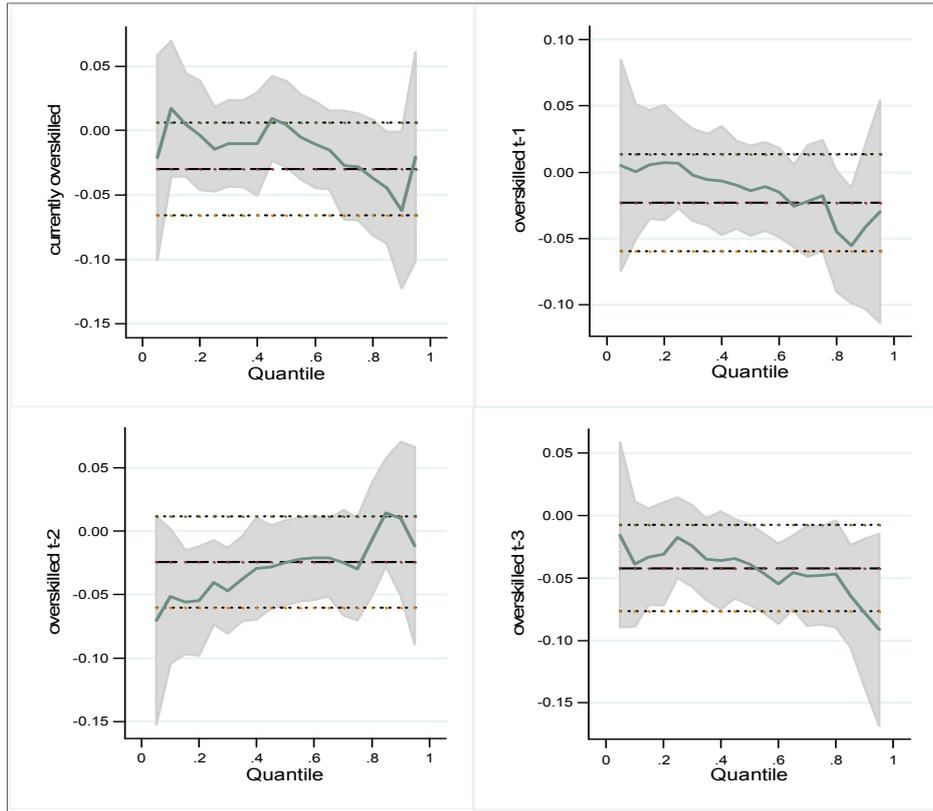
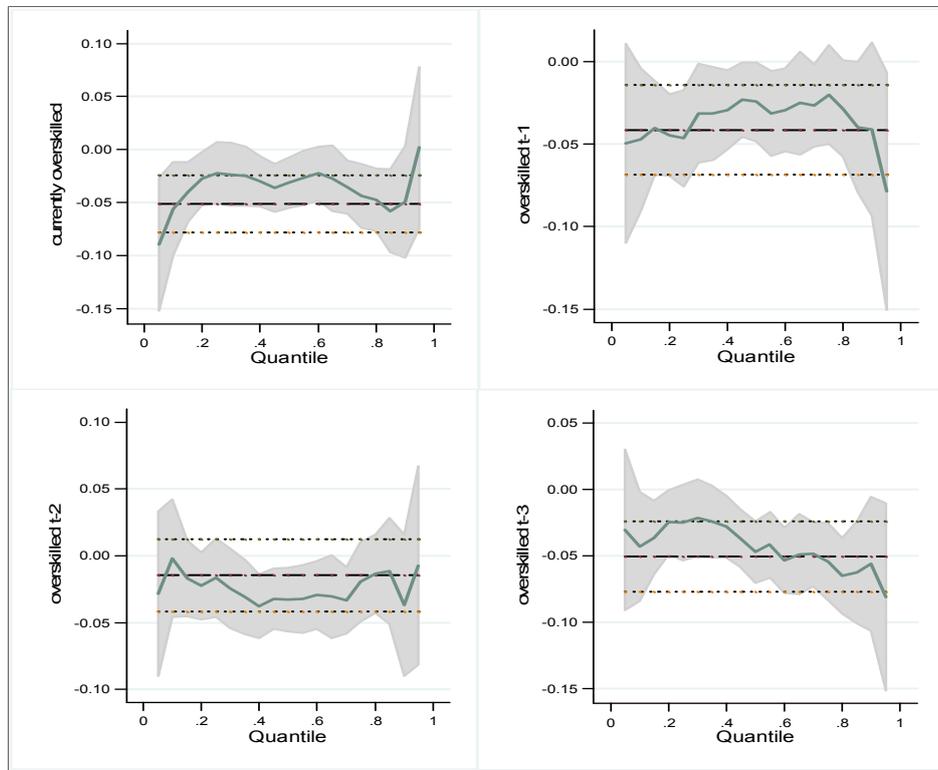


Figure 6 Quantile estimates of wages and overskilling, below Year 12



Estimating scarring

Review of the scarring concept

We begin with an explanation of the term 'scarring' in the context of labour market outcomes. In the labour economics literature there are many adverse outcomes recognised to have a scarring effect on lifetime outcomes. The economic explanation of scarring is that there are labour market experiences that harm the level and (or) the development of individual human capital and the capacity to participate in the labour market in the long run, in that those who are subjected to these experiences develop a long-term labour market disadvantage. A scarring effect is presented as a disadvantage that is self-perpetuating for the individual and is clearly over and above any positive or negative effect that their individual characteristics may play in relation to the presence or absence of this disadvantage. For example, youth underemployment and (or) unemployment may have long-lasting negative effects on the labour market participation (and subsequently general labour market performance) of those who are unfortunate enough to have these experiences (Bell & Blanchflower 2010). The increased youth unemployment that recessions cause has been shown to reduce lifetime participation and income for those who are subjected to it. Otherwise identical youth who grew up in a better macroeconomy have been shown to enjoy better lifetime labour market prospects. Another example of the phenomenon of scarring is long-term unemployment and the associated welfare dependence, which are known to cause scarring in the form of continuing long-term unemployment and stronger welfare dependence (Heckman 1978). Scarring is often referred to as state dependence.

The authors of this report have shown in numerous publications that the phenomenon of overskilling has adverse labour market consequences in the form of short-term wage and job satisfaction losses. In this research, the main question is the degree to which the overskilling phenomenon is self-perpetuating in the long-term. To arrive at an answer to this, we estimated two models which examined whether the scarring effect of overskilling lasts for three and five years, respectively. Both models suggested that overskilling is self-perpetuating and long-lasting. Overskilling scarring is analogous to unemployment and underemployment scarring. This analogy is embedded in human capital theory, in that overskilling mismatch is a form of human capital underutilisation, like underemployment and unemployment. The main difference is that unemployment and underemployment are expressed as lost production through underutilisation in terms of working too few hours, while overskilling mismatch is expressed as lost production through underutilisation in terms of using too few skills and abilities. Note that both forms of underutilisation result in lower production and lower pay. As we show in this report, overskilling is a form of underutilisation that behaves in the same self-perpetuating way as that suggested in the literature for unemployment and underemployment. We now turn to scarring and overskilling.

Scarring is a term that is used in this project and has to be explained in simple terms. We have known for some time that the strength of the labour market outcomes associated with and (or) caused by overskilling varies considerably by educational pathway. A distinction that has arisen from recent research (Mavromaras, McGuinness & Fok 2009b) that utilised panel econometrics for the first time in the context of overskilling showed that both overskilling wage penalty and overskilling persistence are concentrated at the two ends of the educational distribution. When we incorporate in the model only the very recent past (in the form of the previous year's interview), it is first and foremost the degree holders who appear to suffer the highest wage penalty and who at the same time have the strongest persistence. Then there are those who have no qualifications beyond Year 10 at school and who suffer the next highest overskilling wage penalty and persistence. In between these two groups are those

with Years 11 and 12 school completion and VET graduates, who appear to suffer the minimal wage penalty and persistence.

For degree holders, conventional wisdom suggests that 'getting a degree' will increase expected earnings by a considerable percentage, so degree-level education is an unconditionally attractive pathway to follow. This information may be correct on average but it is incomplete as it ignores the suggestion that the proportion of degree graduates who get it wrong (that is, those who end up being overskilled) will get it badly wrong (that is, they suffer a wage penalty and they can be trapped in their lower earnings status — they are 'scarred'). For Year 10 school graduates, scarring could be even worse news, as their earnings and employment circumstances fall from a much lower average level. The question regarding this group is 'how bad can it get?' when someone who is at the bottom of the education or qualifications ladder ends up being overskilled. Thus, we define scarring as the situation in which an employee may be in a long-term overskilling position and where they may be suffering long-term wage losses as a consequence. The estimations we presented in the previous sections on state dependence and scarring allow us to build some relevant scenarios for examination. We do this in two stages. The first stage examines self-persistence alone, while the second stage examines self-persistence and wages combined. We use the model which examines the effect of past overskilling up to three years earlier.

Incorporating self-persistence in the calculation of expected wages

In the previous sections we estimated the probability of overskilling mismatch self-persistence and the wage losses caused by such mismatching. This section uses the estimates from these two steps in order to predict the expected wage losses for a number of scenarios. The set of scenarios mirrors those presented in table 11 in the report, with four types of individuals (the '000' never been overskilled; '100' overskilled in the last year only; '110' overskilled in the last two years only; '111' overskilled in all past three years). We calculate expected current wages for each scenario and compare the losses in wages that can be attributed to overskilling self-persistence. The expected current wage for each type of individual is calculated as follows:

$$\begin{aligned} \text{Expected wage at } t &= \text{Predicted wage at } t \text{ if overskilled} \times \text{probability of being overskilled at } t \\ &+ \text{Predicted wage at } t \text{ if well-matched} \times \text{probability of being well-matched at } t \end{aligned}$$

It is worth noting that both probability predictions and wage predictions are made at the individual level and they depend on the estimated level of overskilling mismatch self-persistence. Thus, both predictions are made to control for observed individual differences in the sample (which means that these are accounted for in the expected wage predictions) and for unobserved individual heterogeneity in both persistence and wage estimations (through the use of panel regression, which means that unobserved heterogeneity is comprehensively accounted for in the expected wage predictions we present). Given the data at hand, the evidence we present is the closest we can get to the estimation of the *causal* effect of past and current overskilling mismatch on the earnings of the overskilled in the Australian labour market.

Table 14 offers some noteworthy findings. First, we only find evidence of deepening self-persistence for university graduates and Year 12 school graduates. Workers from all other educational levels may be experiencing overskilling mismatch, but it is not self-persistent and it does not get worse with time.

Table 14 Scarring predictions by education level

Scenario	University degrees		Diplomas		Certificates III/IV		Only completed school		Did not complete school	
	Expected wage at t	Wage penalty	Expected wage at t	Wage penalty	Expected wage at t	Wage penalty	Expected wage at t	Wage penalty	Expected wage at t	Wage penalty
000	30.59	-	25.24	-	22.59	-	21.36	-	19.63	-
100	29.26	-4.3%	24.92	-1.3%	22.46	-0.6%	20.82	-2.5%	19.39	-1.2%
110	28.28	-7.6%	24.37	-3.4%	22.67	0.4%	20.21	-5.4%	19.50	-0.7%
111	27.19	-11.1%	24.69	-2.2%	22.51	-0.4%	19.64	-8.1%	19.02	-3.1%

The main group who suffers from overskilling mismatch is university graduates. The earnings of a mismatched university graduate can be 11% lower than those of a well-matched comparator. This is a massive difference, which can have a large cumulative effect on future earnings and on lifetime income and wealth accumulation. The effect of overskilling mismatch is shown to deepen considerably from -4.3% for the '100' type, to -7.6% for the '110' type and -11.1% for the '111' type. It is clear that, while there will be fewer university graduates who will become overskilled, they will suffer considerably more *per person* than for any other educational category. A similar but weaker picture arises for Year 12 school graduates, with wage losses being -2.5% for '100', -5.4% for '110', and -8.1% for '111'. Diploma graduates show little overall effect of overskilling mismatch on wages, but the reader should be reminded that the wage estimations suggest that this is a diverse group, and that this diversity may mask some significant effects. Further, the sample of diploma graduates may be too small for the detail that is required by the estimation methodology with three lags.

VET graduates with certificate III/IV clearly do not suffer from overskilling in the long run. While there is evidence in this report and in the literature that VET graduates may suffer moderate wage losses when they become overskilled, our findings suggest that these wage losses are short-lived. This finding has been consistent throughout our estimations and presents a picture that is in sharp contrast to that for university graduates.

Finally, the findings for those without post-school qualifications are diverse. As we mentioned, Year 12 school graduates look very similar to university graduates and they suffer the second highest wage losses. However, those without Year 12 school completion provide us with no clear picture in relation to wage losses due to overskilling, with the exception of the '111' type where a strong negative effect is indicated.⁸

⁸ Presumably this is the consequence of the strong coefficient for the t-3 lag in the less than Year 12 education category. As we have already discussed, we do not have any explanation for this result; hence, we warn the reader against making any strong interpretation of this finding.

Conclusion

The research findings of this project have highlighted many hitherto unknown aspects of skills underutilisation in general and overskilling mismatch in particular. Here we discuss how these findings can help us to understand the labour market disadvantages caused by overskilling mismatch and whether and how VET may protect workers from such disadvantages.

The first advance of the project was to define scarring within the context of mismatched employed workers and then to design the appropriate estimation methodology to implement this definition empirically. This definition was based on the concept of the self-persistence of mismatch, the origins of which lie in the literature of state dependence and long-term scarring, known to be caused by long-term unemployment, participation in welfare programs, underemployment and other adverse labour market outcomes. All these phenomena are scarring in the sense that they contain elements of self-perpetuation: once they occur they are self-feeding and self-persistent and their mere presence increases the probability of their continuation. Our empirical analysis provides strong evidence for the presence of both persistence and self-persistence of overskilling mismatch. Strong persistence implies that the supply of skills and abilities by workers does not match the requirements for skills and abilities of jobs in the economy. Between 15 and 30% of employed workers report themselves as overskilled, a percentage that decreases with educational level. These high percentages suggest that the production process in the broader economy does not use the skills present in the workforce and that any means on offer through which this persistence may be reduced are not adequate for reducing persistent mismatch. The wage losses (estimated in this report) and the job satisfaction losses (Mavromaras et al. 2011) associated with overskilling mismatch indicate that overskilling reduces the welfare of workers and reflect possible productivity losses.

Further analysis reveals high levels of self-persistence as well. That is, not only do we find that there are many workers who are overskilled for longer periods of time, but that their past overskilling makes their present and future overskilling more likely. We find that, after we have controlled for the observed characteristics of the workers and their economic circumstances as well as their unobserved individual characteristics through appropriate regression analysis, overskilling mismatch is intensely self-persistent for workers at all educational levels. Even without any prior expectations, the estimated strength of self-persistence is surprising, especially for those who have not completed their high school education and for VET graduates. School graduates who completed Year 12 education show slightly lower and university graduates show much lower, but still sizeable, self-persistence. We mentioned when we presented these results in the previous sections that we do not have any comparative research results to use as a benchmark for the numbers we have estimated, so our discussion relies on our judgment and our commonsense expectations about the extent of self-persistent mismatch in the economy. We believe that the finding that a university graduate who has been overskilled three years in a row has a 38% probability of being overskilled in the next year, by comparison with another university graduate who was well matched three years in a row and has a 4.6% probability of being overskilled in the next year (a difference between the persistently overskilled and the persistently well-matched graduates of 33.4 percentage points) illustrates the importance of individuals finding a job which is an appropriate match. The difference between the persistently overskilled and the persistently well-matched rises to just below 50 percentage points for Year 12 school graduates and above 50 percentage points for all other educational categories. By any

standard, these are very high percentages, reflecting strong imbalances. But, do these imbalances matter? And if so, by how much do they matter?

This research went further than just establishing the over-time incidence of overskilling mismatch self-persistence with our investigation of whether any quantifiable scarring effects may result from it. We defined scarring as the situation where self-persistence of the probability of remaining overskilled is translated into long-run wage losses. We have used the well-established overskilling wage penalty as a measure of the wage losses that can be caused by self-persistent overskilling mismatch. By incorporating overskilling penalties in the picture, we construct a monetary measure of overskilling mismatch losses. This is important, because an argument could be advanced that someone may well be overskilled, but if they do not suffer any financial penalty from being overskilled, then we lack concrete evidence of any welfare or productivity loss and, therefore, we should not worry about workers who may be overskilled. Where we find evidence of wage losses, we argue that these losses are a quantification of welfare losses in the form of foregone production and underutilised human capital. Clearly, this argument cannot be readily transferred to the macroeconomic sphere, as the demand for skills derives from the demand for the goods that these skills can produce. Whether the necessary actual or latent product demand is present or not is a broader national question, which cannot be answered here. However, the macroeconomic conditions in Australia make the assumption of long-run excess labour supply rather improbable.

The calculation of the impact of overskilling mismatch self-persistence on wages is revealing. We find that, although university graduates are the educational category with the lowest persistence and self-persistence of overskilling mismatch, those who become overskilled sustain by far the worst per-person losses in wages amongst all other overskilled workers. The results suggest that it is the better-paid university graduates who suffer the highest overskilling per-person wage losses. This may suggest that higher graduate wages could be offered as compensation for taking higher risks. That is, while graduates can expect that their lifetime earnings on average will be higher than those of non-graduates, there is a higher variance, so that such an outcome is not guaranteed for all, and high self-persistence implies that those who end up in low-paid jobs are more likely to become trapped there. Without further evidence we can only speculate about such interpretations.

By contrast, workers who did not finish school show the highest persistence and self-persistence of overskilling, but they only show wage losses among those who have been overskilled for three years in a row. Given the compressed wage distribution in this educational group, we cannot know whether the effect of self-persistence has been masked by a lack of wage variance within this group. Year 12 school graduates show modest wage losses, which follow a similar pattern as for university graduates. Finally, VET graduates holding certificates show that their highly persistent and highly self-persistent overskilling mismatch cause no wage losses to overskilled workers. Put simply, the findings of this research are that there is an educational divide in relation to the impact of overskilling mismatch. University graduates are the least likely to experience overskilling, overskilling persistence and overskilling self-persistence. At the same time, among all overskilled workers, it is university graduates who are the most likely to sustain self-persistent overskilling wage losses. By contrast, VET graduates show high persistence but low wage losses. In this sense it could be argued that VET is less susceptible to persistent undesirable wage losses, but this advantage would have to be seen in the context of VET wages being the lowest among all workers with post-school qualifications.

References

- Arulampalam, W 1999, 'A note on estimated effects in random effect probit models', *Oxford Bulletin of Economics and Statistics*, vol.61, pp.597–602.
- Arulampalam, W & Stewart, M 2009, 'Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators', *Oxford Bulletin of Economics and Statistics*, vol.71, pp.659–81.
- Bell, D & Blanchflower, D 2010, 'Youth unemployment: déjà vu?', IZA discussion papers no.4705, Bonn.
- Berkhout, P, Hartog, J & Webbink, D 2010, 'Compensation for earnings risk under worker heterogeneity', *Southern Economic Journal*, vol.76, no.3, pp.762–90.
- Doeringer, P & Piore, M 1971, '*Internal labor markets and manpower analysis*', DC Heath, Lexington, Mass.
- Heckman, J 1978, 'Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence', *Annales de l'insee*, vol.30-1, pp.227–69.
- 1981, 'The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process', in *Structural analysis of discrete data with econometric applications*, eds C Manski & D McFadden, MIT Press, Cambridge.
- Mavromaras, K, McGuinness, S & Fok, Y 2009a, 'Assessing the incidence and wage effects of overskilling in the Australian labour market', *The Economic Record*, vol.85, no.268, pp.60–72.
- 2009b, *The incidence and wage effects overskilling among employed VET graduates*, NCVET monograph series, no.03/2009, NCVET, Adelaide.
- 2009c, 'Overskilling dynamics and education pathways', IZA discussion paper no.4321, Bonn.
- Mavromaras, K, McGuinness, S, O'Leary, N, Sloane, P & Fok, Y 2010a, 'The problem of overskilling in Australia and Britain', *The Manchester School*, vol.78, no.3, pp.219–41.
- Mavromaras, K, McGuinness, S, O'Leary, N, Sloane, P & Wei, Z 2010b, 'Job mismatches and labour market outcomes: panel evidence on Australian university graduates', IZA discussion paper no.5083, Bonn.
- Mavromaras, K, McGuinness, S, Richardson, S, Sloane, P & Wei, Z 2011, *Over-skilling and job satisfaction in the Australian labour market*, NCVET, Adelaide.
- Mundlak, Y 1978, 'On the pooling of time series and cross section data', *Econometrica*, vol.46, pp.69–85.
- Orme, CD 1997, *The initial conditions problem and two-step estimation in discrete panel data models*, mimeo, University of Manchester.
- Watson, N & Wooden, M 2004, 'The HILDA Survey four years on', *Australian Economic Review*, vol.37, no.3, pp.343–9.
- Wooldridge, J 2005, 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity', *Journal of Applied Econometrics*, vol.20, pp.39–54.

Appendix A

Definition of variables

Overskilled: dummy variable, takes the value 1 if is overskilled, zero otherwise.

Wage: log of current hourly gross wages and salary from the main job.

Female: dummy variable, takes the value 1 if is female, zero otherwise.

Age: continuous variable, expressed in years.

Age square: continuous variable, expressed in $\text{age}^2/100$.

Disability: dummy variable, takes the value 1 if an individual has disability, zero otherwise.

Married: dummy variable, takes the value 1 if an individual is married (or de facto), zero otherwise.

Urban: dummy variable, takes the value 1 if an individual domiciled within a major city, zero otherwise.

Father was a professional: dummy variable, takes the value 1 if father belonged to a professional occupation, zero otherwise.

Country of birth:

Migrant (English-speaking country): dummy variable, takes the value 1 if migrant from an English-speaking country, zero otherwise.

Migrant (non-English-speaking country): dummy variable, takes the value 1 if migrant from a non-English-speaking country, zero otherwise.

Aboriginal or Torres Strait Islander (ATSI): dummy variable, takes the value 1 if an individual is Aboriginal or Torres Strait Islander, zero otherwise.

Australian born non-ATSI is the reference category.

Hours per week usually worked in main job: continuous variable, expressed in hours.

Tenure in the current occupation: continuous variable, expressed in years.

Tenure with the current employer: continuous variable, expressed in years.

Firm size:

Less than 5 employees: dummy variable, takes the value 1 if working in a firm which has less than 5 employees, zero otherwise.

5 to 9 employees: dummy variable, takes the value 1 if working in a firm which has 5 to 9 employees, zero otherwise.

10 to 19 employees: dummy variable, takes the value 1 if working in a firm which has 10 to 19 employees, zero otherwise.

20 to 49 employees: dummy variable, takes the value 1 if working in a firm which has 20 to 49 employees, zero otherwise.

More than 49 employees is the reference category.

Children aged between 5 and 14: dummy variable, takes the value 1 if an individual has children between the ages of 5 and 14, zero otherwise.

Children aged under 5: dummy variable, takes the value 1 if an individual has children aged under 5, zero otherwise.

Per cent time spent unemployed in last financial year: continuous variable, value of which lies between 0 and 100.

Union member: dummy variable, takes the value 1 if an individual is a member of a trade union, zero otherwise.

Sector:

Agriculture, forestry and fishing: dummy variable, takes the value 1 if working in the industry of agriculture, forestry and fishing, zero otherwise.

Mining: dummy variable, takes the value 1 if working in the industry of mining, zero otherwise.

Electricity, gas, water and waste services: dummy variable, takes the value 1 if working in the industry of electricity, gas, water and waste services, zero otherwise.

Construction: dummy variable, takes the value 1 if working in the industry of construction, zero otherwise.

Wholesale trade: dummy variable, takes the value 1 if working in the industry of wholesale trade, zero otherwise.

Retail trade: dummy variable, takes the value 1 if working in the industry of retail trade, zero otherwise.

Accommodation and food services: dummy variable, takes the value 1 if working in the industry of accommodation and food services, zero otherwise.

Transport, postal and warehousing: dummy variable, takes the value 1 if working in the industry of transport, postal and warehousing, zero otherwise.

Information, media and telecommunications: dummy variable, takes the value 1 if working in the industry of information, media and telecommunications, zero otherwise.

Financial and insurance services: dummy variable, takes the value 1 if working in the industry of financial and insurance services, zero otherwise.

Rental, hiring and real estate services: dummy variable, takes the value 1 if working in the industry of rental, hiring and real estate services, zero otherwise.

Professional, scientific and technical services: dummy variable, takes the value 1 if working in the industry of professional, scientific and technical services, zero otherwise.

Administrative and support services: dummy variable, takes the value 1 if working in the industry of administrative and support services, zero otherwise.

Public administration and safety: dummy variable, takes the value 1 if working in the industry of public administration and safety, zero otherwise.

Education and training: dummy variable, takes the value 1 if working in the industry of education and training, zero otherwise.

Health care and social assistance: dummy variable, takes the value 1 if working in the industry of health care and social assistance, zero otherwise.

Arts and recreation services: dummy variable, takes the value 1 if working in the industry of arts and recreation services, zero otherwise.

Other services: dummy variable, takes the value 1 if working in the industry of other services, zero otherwise.

Manufacturing is the reference category.

Table A1 Descriptive statistics

Explanatory variable	University graduates	Diplomas	Certificates III/IV	Complete school	Not complete school
Female	0.55	0.54	0.34	0.52	0.53
Age	39.22 (10.44)	39.95 (10.47)	38.78 (11.33)	32.79 (11.85)	40.42 (12.11)
Age square/100	16.47 (8.42)	17.06 (8.46)	16.32 (8.90)	12.15 (8.64)	17.81 (9.43)
Disability	0.11	0.12	0.14	0.11	0.16
Married	0.73	0.70	0.71	0.56	0.68
Urban	0.93	0.89	0.85	0.89	0.84
Father was a professional	0.27	0.17	0.09	0.17	0.06
Migrant (English speaking country)	0.11	0.11	0.10	0.08	0.08
Migrant (non-English speaking country)	0.14	0.10	0.07	0.11	0.06
Aboriginal or Torres Strait Islander	0.01	0.01	0.02	0.02	0.03
Hours per week usually worked in main job	39.02 (12.90)	37.42 (12.92)	39.40 (12.46)	35.60 (12.95)	35.06 (13.84)
Tenure in the current occupation	9.46 (9.43)	9.46 (9.35)	9.60 (9.89)	5.96 (7.45)	8.65 (9.31)
Tenure with current employer	7.04 (7.92)	7.56 (8.31)	6.20 (7.46)	4.78 (6.23)	6.27 (7.41)
Firm has less than 5 employees	0.05	0.08	0.10	0.10	0.13
Firm has 5 to 9 employees	0.07	0.11	0.14	0.14	0.15
Firm has 10 to 19 employees	0.10	0.16	0.15	0.15	0.17
Firm has 20 to 49 employees	0.18	0.20	0.19	0.18	0.17
Have children aged between 5 and 14	0.27	0.29	0.28	0.20	0.28
Have children aged under 5	0.14	0.13	0.13	0.12	0.09
Per cent time spent unemployed in last financial year	1.30 (7.73)	1.61 (9.14)	2.06 (10.25)	2.86 (11.88)	3.42 (14.04)
Union member	0.37	0.32	0.31	0.21	0.25
Agriculture, forestry and fishing	0.01	0.02	0.02	0.02	0.04

Mining	0.01	0.01	0.04	0.01	0.02
Electricity, gas, water and waste services	0.01	0.01	0.02	0.01	0.01
Construction	0.02	0.04	0.09	0.05	0.07
Wholesale trade	0.02	0.03	0.04	0.04	0.05
Retail trade	0.03	0.06	0.08	0.14	0.14
Accommodation and food services	0.01	0.04	0.06	0.09	0.07
Transport, postal and warehousing	0.02	0.04	0.05	0.06	0.07
Information, media and telecommunications	0.03	0.04	0.02	0.03	0.02
Financial and insurance services	0.05	0.06	0.02	0.06	0.03
Rental, hiring and real estate services	0.01	0.02	0.01	0.02	0.01
Professional, scientific and technical services	0.12	0.07	0.03	0.06	0.03
Administrative and support services	0.01	0.03	0.02	0.03	0.03
Public administration and safety	0.11	0.12	0.08	0.08	0.05
Education and training	0.27	0.16	0.04	0.04	0.05
Health care and social assistance	0.18	0.14	0.14	0.10	0.11
Arts and recreation services	0.01	0.02	0.01	0.02	0.01
Other services	0.02	0.03	0.06	0.03	0.03
Observations	13763	4632	10434	7672	11436

Note: Mean (standard deviation). The sample consists of all working-age paid employees from HILDA 2001–09.

Appendix B

Estimation of individuals' probability of being overskilled in their current job and the effect of state dependence

The first research question we address relates to the analysis of individuals' probability of being currently overskilled in their job, with a particular emphasis on the lasting impacts of previous experience in overskilling, and with the explicit aim of comparing individuals across their educational achievement. Thus the outcome variable of interest in these estimations is dichotomous; it can only assume two values: 1 if the individual is overskilled in his job at the time of the Household, Income and Labour Dynamics in Australia Survey interview; 0 if they are properly matched. The appropriate estimation technique for this type of outcome variable (henceforth *dependent variable*) is a non-linear model, which offers the benefit over its linear counterpart of ensuring that the estimated outcome lies between 0 and 1, since we are estimating probabilities. It also has the noticeable benefit of allowing regressors to exert a non-constant influence on the outcome variable depending on their value. The use of these estimation techniques is also grounded in economic theory, as the outcome can be modelled in an individual utility maximisation framework of decision-making through a latent variable interpretation of the dependent variable. Through a latent variable approach to the estimation, we do not observe the individual's actual probability of being currently overskilled in their job (at time t) but we observe the outcome, that is, whether they are overskilled or not at time t . Our model aims to use the information on the outcome in order to obtain an estimate of the probability of being overskilled conditioned on a number of exogenous variables suspected to have an influence on these probabilities. Altogether we can write the model in the following manner:

$$y_i^* = v + w$$

Where y_i^* is the latent variable corresponding to the probability of being currently overskilled for individual i , while v represents the observable part of this probability composed of a number of individual and workplace characteristics and w the unobservable part which can be attributed to individuals' unobserved preferences or abilities not accounted for elsewhere. In the present estimations, due to the nature of the data discussed below, we assume w to be normally distributed with mean 0 and constant variance; thus, a model known as the probit model is estimated. The observable counterpart of the latent variable is such that:

$$y_i = \begin{cases} 1 & \text{if } y_i^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The research question and the nature of the data used imply that a number of econometric issues need to be addressed in order to obtain meaningful estimates of the probability of being currently overskilled. First of all, the data are from a longitudinal survey. For the purpose of these estimations we have access to up to nine observations per individual, corresponding to as many waves of interview. This allows us not only to address the research question whose focus is on the effect of past experience of overskilling on individuals' current overskilling, but also, through appropriate techniques, to have better control of unobserved individual heterogeneity, thus allowing derivation of the direct and causal impact of individuals' past experience on their current situation. The effect of past experience is captured in the model by introducing information on the dependent variable obtained at previous waves of interviews, which we call *lagged dependent variables*. We propose two models. The first one incorporates three lags; that is, information on whether the individual was overskilled up to three years before the current interview. The second model includes a longer history

of information, with five lags. If these lags prove to be statistically significant in the estimation, it shows, everything else held constant, that experiencing a situation of overskilling in the past has an effect on the probability of being currently overskilled. The panel data version of the probit model we use allows us to control for individual-specific unobserved heterogeneity through the inclusion of a second, individual, component in the unobservable part of the model which we assume to be random. This random component can be seen as a tool to control for time-invariant unobserved individual heterogeneity; that is, characteristics that are specific to the individual – and which we cannot observe – but which affect their propensity to be overskilled in their job. Since we include information on past experience, the appropriate methodology used in the estimation is called a dynamic random effects probit model. Altogether we can rewrite the latent equation in a more precise form as:

$$y_{it}^* = \sum_{l=1}^L \gamma_l y_{i,t-l} + X'_{it} \beta + \alpha_i + u_{it}$$

with X_{it} , a matrix of individual and workplace characteristics (including a constant) which are allowed to be both time-variant and invariant. The $y_{i,t-l}$ represent the lags of the dependent variable, with L equal to 3 in the first model and 5 in the second, and the γ_l are the coefficients associated with the lags which are to be estimated. α_i is the individual-specific random component capturing the effect of time-invariant individual unobserved heterogeneity and, finally, u_{it} is an error term which represents the unobservable component of the probability of being currently overskilled. It is assumed to be normally distributed with mean 0 and constant variance, and serially independent, $u_{it} \sim N(0, \sigma_u^2)$. We assume that, conditional on the explanatory variables of the model, the individual effect α_i is also normally distributed, $\alpha_i \sim N(0, \sigma_\alpha^2)$ and independent from the explanatory variables and the error term. This formulation of the model implies that the correlation between two successive error terms for the same individual is a constant ρ , given by $\rho = \text{corr}(w_{it}, w_{i,t-1}) = \sigma_\alpha^2 / (\sigma_\alpha^2 + \sigma_u^2)$, which we estimate in the model. Furthermore, the basic random effects model assumes that the random effect α_i is uncorrelated with the explanatory variables, whether they be the lagged dependent variables or any of the variables in X_{it} . This constitutes a particularly strong assumption that is not likely to hold, given our data. Ignoring these correlations, if they exist, would lead to inconsistent estimates of all of the coefficients of the model and the random effects probit model would tend to overestimate the effect of the lagged dependent variable. Provided that the analysis of these lagged dependent variables are of elemental importance for our research question, we need to ensure that the parameters of the model are consistently and precisely estimated.

We resolve the issue of the potential correlation between the individual effect and the explanatory variables by assuming a relationship between this individual effect and the means of the time-varying variables for each individual as suggested by Mundlak (1978) and as commonly used in the economic literature.

With regard to the assumption of no correlation between the lagged dependent variable and the individual effect, we can show that it amounts to assuming that there is no correlation between the latter and the initial observation of the dependent variable for each individual. This again is a strong assumption, which calls, at least, for testing whether it holds and the introduction of an appropriate correction if it does not. The issue of the correlation between the individual effect and the initial observation of the dependent variable for each individual arises because the observation of the individuals' situation with regard to overskilling does not coincide with the start of the stochastic process that generates the individuals' experiences with respect to overskilling. This issue is known as the *initial conditions problem* and was discussed by Heckman (1981). Heckman (1981) proposed an alternative estimator, incorporating a linear approximation of the latent dependent variable at the

initial period, which is used to express the joint probability of the observed sequence of individuals' experiences, given the individual effect α_i . Alternative, less computation-intensive estimators have been proposed, notably by Orme (1997), Arulampalam and Stewart (2009) and Wooldridge (2005)⁹. In the present research, we follow the Wooldridge (2005) method, which consists of modelling the relationship between the individual effect α_i and the initial observation of the dependent variable through the conditional distribution of α_i given the initial value of the overskilling variable. The method involves expressing the individual effect as a function of the initial observation of the dependent variable and a set of instruments, which can be the same as those normally used for the Mundlak correction; that is, the averages of the time-varying individual and workplace characteristics over time per individual. Altogether, the individual effect α_i is expressed as:

$$\alpha_i = \bar{X}_i' \delta + \theta y_{i1} + \varepsilon_i$$

With \bar{X}_i representing the time-varying explanatory variables of the model found in X_{it} , averaged over each individual. y_{i1} represents the first observation of the binary dependent variable for the individual, indicating whether the individual was overskilled in their job the first time they are observed in the survey. ε_i is a random, unobserved individual component assumed to be normally distributed with constant variance σ_ε^2 and uncorrelated with any of the explanatory variables and u_{it} . The full model for the latent variable of interest, which accounts for the correlation of the individual effect with both the lagged dependent variable and the explanatory variables, can then be written as¹⁰:

$$y_{it}^* = \sum_{l=1}^L \gamma_l y_{i,t-l} + X_{it}' \beta + \bar{X}_i' \delta + \theta y_{i1} + \varepsilon_i + u_{it}$$

With the observable counterpart defined as:

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The parameters of the model are estimated by conditional maximum likelihood following Wooldridge (2005) and the estimated probability to be overskilled for individual i is therefore given by:

$$\Pr(y_{it} = 1 | y_{i,t-l}, (l=1..L), X_{it}, \bar{X}_i, y_{i1}, \varepsilon_i) = \Phi[\sum_{l=1}^L \hat{\gamma}_l y_{i,t-l} + X_{it}' \hat{\beta} + \bar{X}_i' \hat{\delta} + \hat{\theta} y_{i1} + \hat{\varepsilon}_i]$$

with $\Phi(\cdot)$ representing the cumulative normal distribution and the 'hats' placed atop the parameters taken as meaning that they are the estimated parameters of the model. It is worth noting that we do not directly estimate $\hat{\varepsilon}_i$, the individual effect, but its variance in the model. A way to actually account for the individual effects when computing estimates of individuals' probabilities is proposed by Arulampalam (1999) and consists of using the estimate (constant) of the correlation between two successive error terms for an individual (denoted by ρ above) in order to weight the estimated parameters of the model. Accordingly, the estimated parameters of the dynamic random effects probit model should be multiplied by a factor of $\sqrt{1-\rho}$. However, as we see in the estimation results, the values of ρ we obtain for both the five lags and three lags models are very small and so the correction is not warranted. Therefore, in the result tables illustrating the estimated probabilities for the different scenarios attached to individuals' past experiences of overskilling, the estimated probabilities are obtained as follows:

$$\Pr(y_{it} = 1 | y_{i,t-l}, (l=1..L), X_{it}, \bar{X}_i, y_{i1}) = \Phi[\sum_{l=1}^L \hat{\gamma}_l y_{i,t-l} + X_{it}' \hat{\beta} + \bar{X}_i' \hat{\delta} + \hat{\theta} y_{i1}]$$

9 Arulampalam and Stewart (2009) put Heckman's and the other estimators cited above to the test. They emphasise the benefits obtained from using the Mundlak correction and point out that all estimators provide similar results. Consequently, we made the choice of the Wooldridge (2005) method for the purpose of this research.

10 We assume that the control for the initial conditions on the dependent variable via the introduction of the first observation of this variable per individual controls for the correlation between the individual effect and all of the lags of the dependent variable introduced in the model.

NVETRE Program funding

This work has been produced by NCVET under the National Vocational Education and Training Research and Evaluation (NVETRE) Program, which is coordinated and managed by NCVET on behalf of the Australian Government and state and territory governments. Funding is provided through the Department of Education, Employment and Workplace Relations.

The NVETRE program is based upon priorities approved by ministers with responsibility for vocational education and training (VET). This research aims to improve policy and practice in the VET sector. For further information about the program, go to the NCVET website <<http://www.ncvet.edu.au>>.

The author/project team was funded to undertake this research via a grant under the NVETRE Program. These grants are awarded to organisations through a competitive process, in which NCVET does not participate.



National Centre for Vocational Education Research Ltd
Level 11, 33 King William Street, Adelaide, South Australia
PO Box 8288, Station Arcade, SA 5000 Australia
Telephone +61 8 8230 8400 Facsimile +61 8 8212 3436
Website www.ncver.edu.au Email ncver@ncver.edu.au