Course non-completion and multiple qualifications: re-estimating the returns to education in Australia

Austen Peters *Bankwest Curtin Economics Centre*
Alfred M. Dockery *Bankwest Curtin Economics Centre*
Sherry Bawa *School of Accounting, Economics and Finance, Curtin University*

Abstract

Among Australian studies estimating returns to education, there is a consensus that education is a highly profitable investment. Conventional estimates of returns to education examine earnings conditional upon individuals' years of education, with years spent in education typically inferred from their highest qualification attained. However, this approach underestimates the actual time individuals may have spent in education as it ignores multiple qualifications obtained and time spent towards qualifications that were not completed. Using 2001-2019 panel data from the Household, Income and Labour Dynamics in Australia (HILDA) survey supplemented by several Australian Bureau of Statistics sources, we estimate the sensitivity of estimates of the returns to education from a standard wage equation to the inclusion of course non-completion and multiple qualifications. Taking account of these sources of mismeasurement, we estimate the wage premium associated with each additional year of education to be 5.5 per cent, as opposed to 6.5 per cent using the conventional approach, or around 15 per cent lower. These differences are similar by gender and broad age group. We find unaccounted for years spent accruing multiple qualifications to be the main source of overestimation of the returns to education, although we note the lack of individual-level data on incomplete qualifications may have mitigated against identifying a larger effect of accounting for this source of mismeasurement.

JEL Codes: I26, J24, J31
Keywords: Education, return to education, human capital theory, non-completion, multiple qualifications, Australia

---

Acknowledgement: This paper is based on Austen Peters' dissertation for the award of Bachelor of Commerce (Honours) at Curtin University, under supervision of the co-authors. It uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to either DSS or the Melbourne Institute.

Corresponding Authors: Austen Peters, austen.peters@curtin.edu.au; Alfred Michael Dockery, m.dockery@curtin.edu.au, ORCID: 0000-0002-4517-2442; Sherry Bawa, sherry.bawa@curtin.edu.au ORCID: 0000-0002-3346-4691

1. Introduction

Estimating the returns to education for individuals and to society is of ongoing interest to researchers and policymakers (see, for example, Psacharopoulos and Patrinos, 2018). Following Human Capital Theory (HCT), two main empirical methods guide estimation of the returns to education: Mincer's (1974) human capital earnings function and Becker's (1964) internal rate of return (IRR) method. In the case of the former, the key empirical approach is to estimate wage equations, which generate an estimate of the change in earnings associated with additional years of education or with a particular qualification level, such as a bachelor's degree.

There has been significant debate around the extent to which estimates from wage equations can be taken to infer a causal effect of education on earnings. This includes theories in which education provides a signal of workers' pre-existing abilities to allocate them to jobs, rather than directly increasing their productivity as assumed in HCT (see, for example, Arrow 1973, Spence 1973). An extensive empirical literature also attempts to control for potential bias in the estimated effect of education on earnings due to unobserved characteristics, especially ability bias and selection on other factors such as family background (Bonjour *et al*. 2003; Card 1999, 2001; Li *et al*. 2012; Heckman *et al*. 2016; Wilkins 2015).

However, one aspect of returns to education estimates that appears to have been largely unchallenged is the specification of 'years of education'. The years of education undertaken by an individual is typically inferred from their highest educational qualification attained, and taken to be the number of years that specific qualification usually takes to complete. This fails to take into account individuals who do not complete their course, either due to dropout or course switching, and individuals who hold more than one qualification at the same qualification level. These sources of under-estimation of the actual years of education individuals have accrued, hereafter referred to as course non-completion and multiple qualifications, respectively, can be expected to lead to over-estimation of the returns to education. According to Australian Bureau of Statistics (ABS) data, 2.6 million individuals in 2020 had an incomplete qualification in Australia (ABS 2020c). Whilst not directly collected in major microdata sources, we also estimate multiple qualifications to be widespread in Australia.

In the Australian context, this paper analyses the sensitivity of estimates of the returns to education in standard wage equation models to the incorporation of course non-completion and multiple qualifications into the measurement of years of education. Data from various ABS sources and the Household, Income and Labour Dynamics in Australia (HILDA) survey are used to derive a measure of 'actual' years of education, which accounts for both course non-completion and multiple qualifications.[1] Results from models using this 'actual' measure of years of education are compared with those obtained using a standard 'inferred' measure based on highest level of qualification. We also test whether accounting for course non-completion and multiple qualifications has implications for inferences on differences in the returns to education by gender and age.

1 While we use the term 'actual' years of education for convenience, we fully acknowledge that this measure also has limitations.

This paper focuses on estimates of the returns to education using the earnings function approach. However, accounting for multiple qualifications and course non-completion also has implications for estimates based on the IRR approach. We show there are potentially important differences in the effect of accounting for actual years of education under the two approaches.

2. Background

Human Capital Theory, as pioneered by Schultz (1961), Becker (1964) and Mincer (1974) suggests that the human capital an individual acquires increases their labour productivity, which in turn augments their earning capacity over their lifetime. Education facilitates the growth of human capital through the development and accumulation of knowledge and skills which enhance an individual's labour productivity. However, education comes at a cost, and the concept of the return to education has become a central focus in labour economics and for policymakers (Wei 2010). Returns to education can be considered in both the private and public context; however, this paper focuses on estimates of private returns to education.

The two main approaches to estimating returns to education are the IRR and the Mincer human capital earnings function (Wei 2010). Initially devised by Becker (1964), the IRR approach is based upon the notion that, just as with traditional capital, education increases future earnings, but at an immediate cost in the present in the form of opportunity costs (foregone earnings) and direct costs (course fees, textbooks, etc.). The IRR is the discount rate which equates the present value of the future stream of benefits to the costs. If that discount rate is relatively high, say compared to a benchmark like the long-term bond rate, then investing in education in the present can be considered to be relatively profitable and individuals will opt to undertake further education (Harmon *et al*. 2003).

Mincer's human capital earnings function is a semi-logarithmic function which expresses earnings as a function of years spent in education, and takes the following form:

$$\ln Y_i = \alpha + \beta X_i + \gamma S_i + \mu_i \qquad (1)$$

Where $\ln Y_i$ represents the log of earnings for individual i and X_i represents a vector of coefficients that affect earnings, such as age, experience, disability status and Indigenous status. S_i represents years spent in education by the individual, and μ_i an error term. The parameters to be estimated include the constant term α, the vector of coefficients β, and the private rate of return to education, γ. Under this specification, γ approximates the percentage increase in earnings associated with each additional year of education.

Major challenges to the HCT approach relate to the assumed link between workers' level of education, their productivity and their wages. Alternative theories emphasise the credential role of education, whereby educational qualifications simply act as a screen or signal to employers to allocate workers to jobs. The important implication is that productivity is attached to jobs and to fixed traits of workers; education does not increase workers' productivity *per se* (Arrow 1973, Rospigliosi *et al*. 2014, Spence 1973).

A less fundamental challenge is of unobserved ability bias. If individuals with higher ability are inherently more productive workers, and also more likely to accrue higher levels of education, the effect of education on productivity and earnings will be over-estimated if this is not fully accounted for. The growing literature on returns to education has seen increasingly sophisticated econometric techniques to account for ability bias and to establish causal effects of education on earnings, such as through natural experiments (see Card 1999, 2001; Heckman *et al*. 2006; Heckman *et al*. 2016). This has included studies utilising samples of twins to control for family background and for ability in the case of identical twins (see, for example, Bonjour *et al*. 2003 and Li *et al*. 2012; and Miller *et al*.1995 for an Australian study).

Internationally, studies based on human capital earnings functions generally find education to be a profitable investment for individuals (see reviews in Dickson and Harmon 2011, Harmon *et al*. 2003, Heckman *et al*. 2006, Psacharopoulos and Patrinos 2018). This also holds for Australia, with estimates suggesting returns in the vicinity of 10 per cent higher wages per year of education (Daly *et al*. 2015; Deloitte 2017; Dockery and Miller 2012; Kler 2005; Leigh 2008; Sinning 2014; Wei 2010; Wilkins 2015), although there is evidence that returns may be declining as the number of individuals with university qualifications in Australia rises (Corliss *et al*. 2020; Productivity Commission 2019).

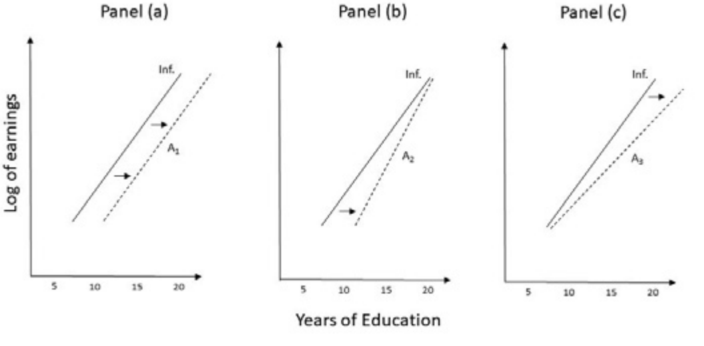
Studies to have estimated returns to education in Australia using the IRR approach similarly find education to be a profitable investment. Estimates of the return to completing a university degree typically range from between 10 to 20 per cent, with substantial variation across disciplines and over time as labour demand for graduates and non-degree holders fluctuates (see Borland *et al*. 2000; Norton 2012; Wei 2010 and a series of papers by Corliss and colleagues: Corliss *et al*. 2020, Corliss *et al*. 2013, Daly *et al*. 2015). Wei's (2010) study looked at returns to education based on each five-yearly census over a 25-year period (1981-2006), and is notable for generating estimates using both earnings functions and the IRR methods, with broadly similar results.

3. Course non-completion and multiple qualifications – theoretical expectations and some existing evidence

While the literature has paid extensive attention to potential bias in estimates of the returns to education due to ability bias or unobserved heterogeneity, very little attention has been paid to the potential bias due to mismeasurement arising from course non-completion and multiple qualifications. Our initial expectation when embarking on this study was that adding unaccounted years spent in education into the measure of years of education would lead to a roughly proportionate reduction in the estimated return to years of education. The motivation was to assess the degree of over-estimation of the returns to education, and the potential implications for decision-making by individuals and policymakers. However, in attempting to interpret our results, it became clear that the effects of this mismeasurement on estimates of the returns to education from the wage equation approach are far more nuanced.

To appreciate why, note that a wage equation of the form (1) above estimates the percentage change in wages associated with an increase in years of education, or the slope of a line of best fit to the log of earnings when plotted against years of education. This is depicted by the schedule labelled 'Inf.' (for inferred) in Figure 1. Assume all unaccounted years of education are evenly distributed across all workers. Adding these to the measure of inferred education simply leads to a rightward shift in the schedule from 'Inf' to the schedule of 'actual years' (depicted by A_1) as shown in Panel (a). It can be seen that there is no change in the slope of the earnings function with respect to years of education, and hence no change in the estimated return to an additional year of education.

Figure 1. Effects of adjusting for actual years of education



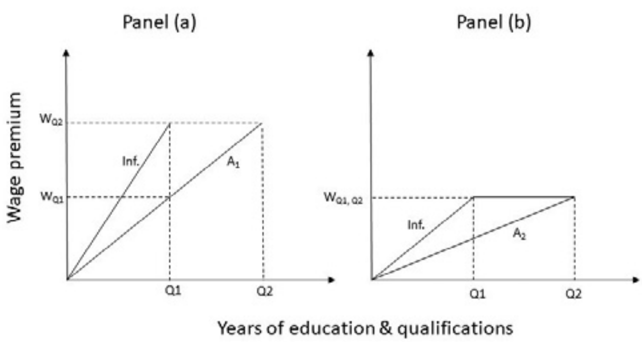
In contrast, consider the case in which unaccounted years of education are concentrated among the less educated. Relative to inferred years of education, there is lower variation for years of actual education, and a steeper education-wage gradient as shown by A_2 in Panel (b). Hence, adjusting for actual years spent in education can lead to a higher estimated rate of return. The converse, shown in Panel (c), applies if unaccounted years of education are concentrated among the more educated. Variation in measured years of education widens, reducing the education-earnings gradient.

Put another way, if the degree of mismeasurement is negatively correlated with inferred education, adjusting for actual years of education will lead to higher estimated rates of return to education. If mismeasurement is positively associated with years of inferred education, adjusting for actual years of education will reduce the estimated returns.

However, the results will also be affected by the extent to which unaccounted years of education attract a wage premium. This is illustrated using the case of multiple qualifications. Compared to the approach based on inferred years of education, a calculation based on actual years of education will necessarily reduce the return to each year of education for those with multiple qualifications. The exact nature of any bias varies depending on the returns to years of education spent attaining multiple qualifications.

This is shown in Figure 2. Panel (a) represents the case in which years spent gaining the first (Q1) and second qualification (Q2) attract exactly the same wage premium. In this case, using inferred years of education gives a higher wage gradient (Inf) for people with a second qualification, because the higher wages associated with the second qualification are attributed solely to the first qualification. Using actual years of education gives the flatter wage schedule A_1. The other extreme, in which the second qualification attracts no additional wage premium at all, is depicted in Panel (b). Again, the estimated return to years of education is higher when inferred years rather than actual years is used as the basis for calculation, but this time due only to the change in the measured number of years spent in education.

Figure 2. Estimated return to education for persons with a second qualification, based on inferred and actual years



It is difficult to say which scenario will lead to the larger divergence between estimates based on inferred and actual years of education, and the net result will depend on the values of the underlying parameters, such as the relative wage premiums, the number of additional years required to attain a second qualification, and their distributions across the population. Potentially, the case in which multiple qualifications attract a relatively high wage premium will be most sensitive to the inclusion of actual years in education, given that there is both a wage (numerator) and years (denominator) effect.

This example is related to the so-called 'sheepskin effect' that differentiates between wage premiums associated with years spent in education towards incomplete qualifications and those spent towards a qualification that is completed (Jaeger and Page 1996). If education has an effect on workers' productivity, then it is argued that the time spent towards complete and incomplete qualifications should have a similar effect on productivity and wages. In contrast, if returns accrue only for years of a completed qualification, this would suggest credentialism effects, rather than productivity effects.

As in the case of multiple qualifications, these contrasting scenarios also have implications for the potential effect of controlling for time spent in non-completed qualifications. If those years offer a low return, then accounting for actual years would result in a lower estimate of the returns to education compared to an estimate using inferred years, as those additional years are ignored in the inferred measure. On the other hand, if years in completed and non-completed qualifications generate similar wage premiums, accounting for actual years may increase the estimated returns to education relative to using inferred years. This is because the higher wages associated with those years of incomplete qualifications will be assigned to unqualified persons using the inferred approach, reducing the apparent return to years spent successfully gaining qualifications.

3.1 Existing evidence

There is a substantial literature looking at rates of course attrition and its contributing factors (see, for example, Bednarz (2014) and Stromback and Mahendran (2010) for Australian apprentices and trainees; Cherastidtham *et al*. (2018), Edwards and McMillan (2015) with respect to university courses). However, these generally are not in the context of estimating returns to education. Course non-completion has also been analysed in relation to the 'sheepskin effect', with that literature typically focussing on the relative returns to course non-completers compared to completers or to those who have not commenced a course. Toutkoushian *et al*. (2013) and Zeidenberg *et al*. (2015) acknowledge that the risk of non-completion is often ignored and should be taken into account in estimating returns to education. Both those studies find a much lower return to those who attend a US college but do not complete than for those who graduate. Giani *et al*. (2020) cite mixed results in the existing literature as to the benefits of attendance at a US college without attaining a qualification. Their own study finds incomplete college does confer some future benefits in terms of employment and earnings, but non-completers do not fare as well as college graduates.

In estimating the return to education in Australia, Borland *et al*. (2000) and Daly *et al*. (2015) merely acknowledged the presence of course non-completion. Long and Shah (2008) added a 10 per cent adjustment for bias due to course non-completion in their estimation, but the study only estimated returns to vocational qualifications. Amongst the other limited literature, Marks (2007) looked at non-completion and found that having a non-completed degree was associated with a lower incidence of unemployment, but similar labour market outcomes in terms of earnings, job satisfaction and occupational status compared to those who never started a degree. Those who completed a degree fared better on all counts. Norton *et al*. (2018) report mixed subjective perceptions on the value of non-complete degrees on the part of students who drop out from university.

We have not identified any previous studies of the returns to multiple qualifications or how either of these sources of mismeasurement affect overall estimates of the returns to education.

4. Allowing for course non-completion and multiple qualifications

Most studies of the returns to education in Australia that follow the earnings function approach infer the number of years of education an individual has completed from their highest level of education attained. This omits qualifications that have not been completed, as well as multiple qualifications achieved at the same or lower qualification levels (e.g. two bachelor's degrees). As set out above, if this 'inferred' years of education is significantly lower than the actual number of years individuals spent in education, existing estimates are likely to be biased.

To assess the potential magnitude of the effect of this mismeasurement, we estimate earnings functions using data from Waves 1-19 of HILDA. Initially, the model is estimated with the measure of years of education inferred from individuals' reported highest level of qualification, as is common practice. We then draw on a range of data sources to generate more accurate measures of years spent in education, accounting for course non-completion and multiple qualifications. Finally, we compare the estimated returns to years of education when the models are estimated using 'actual years' of education instead of 'inferred years'.

HILDA is a household-based panel study that tracks and interviews Australian respondents on a yearly basis. The HILDA survey collects important economic and personal information from the respondents which provides insight into economic and personal well-being, including educational attainment, labour market experience and individual earnings (Melbourne Institute 2020). HILDA follows respondents over the course of their lifetime from the age of 15. At the time of writing, data from 19 waves were available, spanning 2001 to 2019. Since the first wave in 2001, the survey has recorded responses from around 17,000 Australians from over 7,000 households annually. A top-up sample of 2,153 households (5,477 individuals) was added in 2011 to allow new population sub-groups to be represented as well as alleviate biases from non-random attrition (Watson and Wooden 2003).

For this analysis, the sample was restricted to individuals between the age of 25-65 years who are working and earning a wage.[2] This results in an unbalanced panel, with over 125,000 observations on 18,182 individuals over the period 2001-2019.

The earnings measure is the log of real hourly wages, which we derive from the "usual hours worked per week" and the "usual weekly wages" variables within the HILDA dataset. Wages for waves 1-19 were inflated using the December quarter Consumer Price Index (CPI) in Australia (ABS 2020a) to be expressed in real 2019 dollars. Below we discuss the derivation of 'inferred' and 'actual' years of education. See Appendix A for definitions of other control variables used in the modelling.

2 The intention of the age restriction is to focus the analysis on individuals' 'prime' working years following initial investments in education. This will largely abstract from part-time jobs for people under 25 whose major labour market activity is study and people aged over 65 accepting lower wages as they transition into retirement.

4.1 Inferred years of education

HILDA asks respondents their "highest year of school completed". The years of schooling are attributed to each year level, whereby Year 12[3] is equal to 12 years of schooling, Year 11 is equal to 11 years of schooling and so on until the lowest category of "attended primary school only", which is taken as 6 years of schooling.[4] Total years of education is measured as the sum of years of schooling and years spent in post-school education and training.

The conventional ('inferred') approach to determining years spent in education by an individual is through the 'highest qualification attained'. We map individuals' highest reported post-school qualification as best as possible to years of education based on 'volume of learning' descriptors given in the Australian Qualifications Framework (AQF) as in Table 1 below (Australian Qualifications Framework Council 2013).

Table 1. Inferred Years of Post-School Education

Highest education level achieved	*'Inferred' years of post-school education*
PhD	8
Master's degree	5
Graduate Diploma, Graduate Certificate	4
Bachelor's degree – Honours	4
Bachelor's degree – Pass	3.5
Advanced Diploma	2
Cert III or IV	1
Cert I or II	0.5

4.2 Accounting for multiple qualifications

HILDA records the number of qualifications individuals have obtained since leaving school at each qualification level, ranging from Certificate Level I to a PhD. Where people report having multiple qualifications at any level, they will have spent more years in education than would be inferred directly from their highest qualification attained. Figure 1 shows the proportion of individuals with a qualification at a given level who hold multiple qualifications at that same level, based on the pooled HILDA data from 2001 to 2019. For both males and females, around 15 to 20 per cent of people

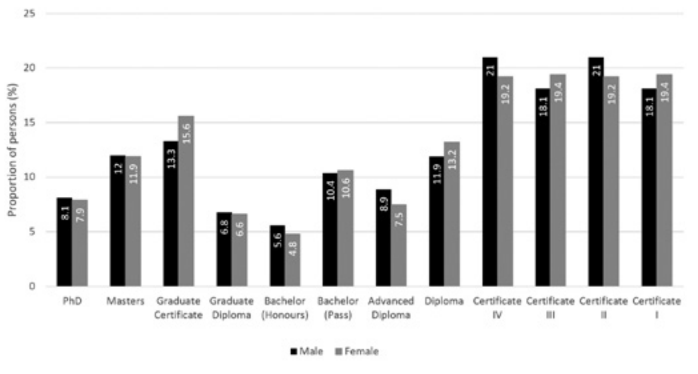
3 In the Australian education system, schooling is compulsory up to Year 10, where the vast majority of students are aged 15 to 16 years (Australian Curriculum, Assessment and Reporting Authority, 2021). Students who intend to participate in university study usually then enrol in two further years of high school study, in Year 11 and Year 12.

4 0.1% of the respondents in the sample listed their "highest year of school completed" as not finishing primary school, and were not included in the sample.

holding certificates from level I to IV have completed a second certificate at that level. Approximately 12 per cent of males and 13 per cent of females with a diploma hold multiple diplomas.

For tertiary-level education, over 10 per cent of males and females with a bachelor's (pass) degree have multiple of this qualification, but gaining multiple honours degrees is less common. Notably, 8 per cent of people with a PhD have completed multiple PhDs, requiring significant time and cost commitments.

Figure 3. Multiple qualifications by gender and qualification level, 2001-2019



To approximate how many additional years of education individuals had spent on multiple qualifications, we assume that completion of additional qualifications at a given level does not require any additional time spent on pre-requisite qualifications. For example, when an individual completes a bachelor's degree and then a master's degree, they will not complete another bachelor's degree to complete a second master's. This applies to all qualifications – allowance is only made for time spent on the incremental qualification.

Second, we assume that when studying toward a second qualification at a given level, people may get an exemption for completing some elements of the course in recognition of prior learning. This will be particularly so if the incremental qualification is in a related field to the learner's existing qualification(s) but, unfortunately, there is no way of assessing this with the HILDA data. We apply a discount of 25 per cent of the time normally taken for each additional multiple qualification. The justification behind the 25 per cent discount rate is that, although maximum recognition of prior learning credits varies by university and qualification level, as a general rule the limit is approximately a one-third (33 per cent) discount for qualifications from a bachelor's (honours) degree and up to a two-thirds (67 per cent) discount for qualifications below a bachelor's (honours) degree when studying in a related field (Curtin University 2016; Griffith University 2019; University of South Australia 2019). On this basis, we assumed that the average discount rate for an individual undertaking a multiple qualification is 50 per cent for a related field and 0 per cent for an unrelated field. Assuming an even split in the distribution between related and unrelated fields of study, yields the 25 per cent discount rate.[5] The resulting assumptions for time spent on multiple qualifications by level is shown in Table 2.

Table 2. Multiple qualifications – additional time spent in education

Qualification level	*Years to complete first qualification*	*Years to complete second and subsequent qualifications with 25% RPL discount*
PhD	4	3
Master's degree	2	1.5
Graduate Diploma, Graduate Certificate	1	0.75
Bachelor's degree – Honours	4	3
Bachelor's degree – Pass	3.5	2.625
Associate Degree	2	1.5
Advanced Diploma	2	1.5
Diploma	1	0.75
Cert III or IV	1	0.75
Cert I or II	0.5	0.375

4.3 Accounting for course non-completion

HILDA does not collect information on course non-completion or dropout. However, the ABS has microdata on course non-completion that can be separated by gender, age and level of qualification. ABS population estimates are used to assign rates of course non-completion to individuals in the HILDA dataset.

In order to measure the additional time spent in education as a result of course non-completion, two variables are created: average time spent in an incomplete qualification and average number of incomplete qualifications. Once created, the product of these yields the total years spent in incomplete courses.

Average time spent in incomplete qualifications

Additional time spent in education as a result of course non-completion is estimated from the Survey of Education and Work (ABS 2020b) and the Survey of Qualifications and Work (ABS 2020c). The latter records the level of most recent incomplete non-school qualification by current qualification level and age. Incomplete qualifications refer to qualifications the individual started, but stopped before completing all academic requirements – they do not include qualifications people are currently studying or training towards.

5 Section 5.2 discusses the sensitivity of the results to this assumption.

For people who failed to complete a qualification, we assume that they spent one-third of the typical years taken to complete that qualification (see Table 3). This is in line with the findings of Norton *et al*. (2018), who reported that over half of all bachelor's degree students in Australia dropped out of their course after completing the equivalent of one year of education, approximately one-third of the degree length.

Table 3. Additional Time Spent in a non-completed course

Qualification level of course	*Years to complete qualification with course requirements met*	*Average years spent in non-completed course (33%)*
Postgraduate Degree[a]	3	1
Graduate Diploma and Graduate Certificate	1	0.33
Bachelor's Degree[b]	3.5	1.16
Advanced Diploma and Diploma	1.5	0.5
Certificate III and IV	1	0.33
Certificate I and II	0.5	0.17

Notes: a. taken as the average of the typical time spent in a PhD (4 years) and a master's (2 years); b. taken as the average of the typical time spent in a bachelor's (pass) degree (3 years) and a bachelor's (honours) degree (4 years).

Based on the level of individuals' most recent incomplete qualification we then calculated the average number of years spent towards those unfinished qualifications using data from the Survey of Qualifications and Work. This average was calculated separately by gender and by highest existing qualification attained, on the basis that the more educated are likely to have attempted qualifications that take longer to complete. Table 4 shows this to be the case: persons holding a university level qualification are estimated to have spent almost one year in their most recent incomplete qualification, compared to around one-half of a year for those with certificate level qualifications.

Table 4. Average Time Spent in Incomplete Qualifications (Amongst Course Non-Completers)

Highest qualification attained	*Male*	*Female*
No post-school qualification	0.70	0.70
Postgraduate Degree	0.94	0.82
Graduate Diploma/Certificate	0.88	0.90
Bachelor's Degree	0.93	0.86
Advanced Diploma/Diploma	0.81	0.77
Certificate III/IV	0.56	0.60
Certificate I/II	0.54	0.54

Average number of incomplete qualifications

Having an estimate for the average time individuals spend in an incomplete qualification, we next needed an estimate of the number of incomplete qualifications people have commenced. ABS data from the Survey of Qualifications and Work on the average number of non-completed qualifications by gender and age were also incorporated into the data. Age was grouped by 10-year brackets across the age range of the sample to be included in the wage equations (25-65 years). The averages, which include individuals who did not have any incomplete qualification, are shown in Table 5.

Table 5. Average number of incomplete qualifications by age and gender

Age	*Male*	*Female*
25-34 years	0.20	0.26
35-44 years	0.22	0.27
45-54 years	0.20	0.20
55-64 years	0.16	0.19
55-64 years	0.54	0.54

Total Years Spent in Incomplete Qualifications

For each individual in the HILDA sample, we could impute a variable for their expected number of incomplete qualifications, based on the population average for their specific age and gender (Table 5). We could also impute a variable for the expected time they would have spent in each incomplete qualification, conditional upon their gender and level of highest qualification (Table 4). The final stage of calculation was to multiply these two variables to create an estimate of expected total years spent in incomplete qualifications. For example, for males aged 25-34 with a postgraduate degree as their highest qualification, each individual within this group had an additional (0.2 × 0.94) = 0.188 years added to their 'inferred' years of education.

This measure of time spent in incomplete qualifications and the measure of additional years for multiple qualifications are added to inferred years associated with an individual's highest level of qualification, to give the total 'actual' years of education measure.

5. Wage equation results

We estimate random effects panel models of the wage equation set out in equation (1) using data from waves 1 to 19 of HILDA. The random effects specification is chosen despite the Hausman test suggesting a fixed effects model. This is because a fixed effects approach would only estimate the returns to education for those individuals in the sample with a changing level of education, and these may be a small subset of the sample. Further, that subset may be biased towards a particular demographic, such as younger individuals, meaning the results are not representative of the population. Using fixed effects is particularly problematic for estimating the effects of course non-completion, since this is not observed for specific individuals in each year, but inferred based on the individual's age, gender and highest qualification. Hence variation in the variable occurs only with changes in these three parameters. Key results obtained using fixed effects are reported below as a robustness check.

The results of the random effects wage equations using the standard inferred measure of years of education and our estimate of actual years of education are reported in Table 6. All control variables have anticipated signs and are in line with existing estimates in the literature. The estimated coefficient on 'inferred' years of education is 0.0651, implying each additional year of education undertaken by an individual results, on average, in a 6.51 per cent increase in hourly wages. In comparison, the estimated coefficient on 'actual' years of education implies an increase in hourly wages of 5.52 per cent. Hence, accounting for multiple qualifications and course non-completion suggests that the return to years spent in education is a full percentage point (or 15 per cent) lower than the return to education obtained using the conventional approach.

Table 6. Random effects estimates: 'inferred' versus 'actual' years of education

	Inferred years of education		*Actual years of education*	
Variable	*Estimated coefficient*	*t-statistic*	*Estimated coefficient*	*t-statistic*
Inferred years of education	0.0651***	(53.93)	-	-
Actual years of education	-	-	0.0552***	(52.69)
Age	0.0056***	(2.92)	0.0044**	(2.27)
Age^2	-0.0083***	(-3.80)	-0.0069***	(-3.20)
Experience	0.0222***	(19.72)	0.0222***	(19.65)
Experience^2	-0.0190***	(-9.13)	-0.0197***	(-9.46)
Male	0.113***	(17.30)	0.1121***	(17.06)
Disability	-0.0155***	(-4.43)	-0.0163***	(-4.66)
Indigenous	0.0452**	(2.12)	0.0363*	(1.70)
Works part-time	0.1139***	(34.30)	0.114***	(34.45)
Marital Status:				
Married	–		–	
Never married	-0.0596***	(-10.64)	-0.0607***	(-10.84)
No longer married	-0.0152***	(-4.10)	-0.0160***	(-4.31)
Employed:				
Permanent/ongoing	–		–	
Casual	-0.0107***	(-2.81)	-0.0111***	(-2.91)
Fixed-term	0.0307***	(7.94)	0.0310***	(8.00)
Union Member	0.0398***	(11.42)	0.0395***	(11.33)
Overseas born	-0.0249***	(-8.72)	-0.0176*	(-8.82)
Socioeconomic decile of neighbourhood 6-10	0.0505***	(14.3)	0.0536***	(15.19)
Language other than English spoken at home	-0.0345***	(-3.19)	-0.0349***	(-2.26)
Overseas highest qualification	-0.0309***	(-5.42)	-0.0361***	(-6.34)
Constant	2.0394***	(53.89)	2.171***	(58.66)
R-squared	0.1467		0.1419	
N	126199		126199	

Notes: t-statistics in parenthesis; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The main driver of the difference between 'inferred' and 'actual' years of education is multiple qualifications. As shown in Table 7, multiple qualifications account for nearly the entire fall in the return to education, decreasing the measure from 6.51 per cent (inferred) to 5.53 per cent, accounting for nearly all of the percentage point difference between returns for 'inferred' and 'actual' years of education. Comparatively, accounting for course non-completion alone decreases the return to education from 6.51 per cent to 6.47 per cent, a negligible 0.04 percentage point decrease.

Table 7. Disaggregation of course non-completion and multiple qualification effects

Variable	*Inferred years of education*	*Course non-completion (Actual)*	*Multiple qualifications (Actual)*
Inferred Years of Education	0.0651***		
	(53.93)		
Actual Years of Education		0.0647***	0.0553***
		(54.10)	(52.50)
Constant	2.039***	2.039***	2.179***
	(53.89)	(53.94)	(58.92)
N	126199	126199	126199

Notes: t-statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

These figures for course non-completion are somewhat surprising. Multiple qualifications may have been expected to have a larger effect on the estimated returns to education due to a larger sample mean of 0.26 years compared to 0.15 years of course non-completion across the sample. While this is what we observe, the relative impact of accounting for multiple qualifications on the estimate of the returns to education is disproportionately large, and the effect of accounting for course non-completion disproportionately small.

As our results show reduced returns to education when actual years is used instead of inferred years, we anticipate that the extent of undercounting of years of education is positively correlated with inferred years, as shown in Panel (c) of Figure 1. This is confirmed by the data. Using the pooled data, the simple correlation between inferred years of education and our estimates of additional time spent in multiple- and non-completed qualifications is 0.253. However, this cannot explain the disproportionately large effect of accounting for years of multiple qualifications. Given the estimate of the returns to education is disproportionately sensitive to accounting for time in multiple qualifications compared to accounting for non-completion, we might anticipate that the former is more strongly and positively correlated with inferred education than years of non-completion. The reverse is in fact the case: the correlation between inferred years of education and additional years spent towards multiple qualifications is 0.225, but a much stronger 0.640 for course non-completion.

As demonstrated in Figure 2, the result will also be affected by the degree to which years spent gaining multiple qualifications or studying towards incomplete qualifications attract a wage premium. The disproportionate impact of accounting for multiple qualifications may indicate that years of education towards multiple qualifications at the same level attract a relatively high wage premium relative to years spent gaining the initial qualification. It is tempting to also impute that returns to incomplete and complete qualifications are closely aligned. However, these conjectures must be tempered by the fact that the effects of multiple qualifications could be estimated with greater precision, given that we have individual-specific measures of years spent in multiple qualifications through HILDA, and which are related to individual wage differentials observed in the data. In contrast, years of non-completion are applied as age-by-gender averages. The availability of individual-specific data on years spent in incomplete qualifications may have resulted in a larger impact on the estimated return to years of education.

The sensitivity to the inclusion of additional years of education on estimates for some of the other variables, as reported in Table 6, is also of interest. Allowing for multiple qualifications and course non-completion results in a lower age-earnings gradient, suggesting that age partly proxies for the accumulation of unobserved years of education in models following the conventional approach. This interpretation assumes those additional years of education have a positive association with hourly wages. Separately including years of incomplete qualifications and multiple qualifications reveals that it is mainly accounting for multiple qualifications that leads to the lower estimated coefficient for age, consistent with individuals accruing those multiple qualifications over time.

The penalty associated with being born overseas is also substantially reduced, and again this is driven entirely by the inclusion of time spent towards multiple qualifications. The mean of years spent in multiple qualifications is marginally higher for the overseas born, at 0.263, compared to 0.248 for Australian born workers. This may reflect the need for people born overseas to gain second qualifications in Australia due to limited- or non-recognition of qualifications accrued in their country of origin. Including those extra years of education would therefore be expected to increase the estimated penalty associated with being born overseas, not decrease it.

Equally, the correlation between inferred years of education and mismeasurement cannot explain why the estimated wage penalty associated with being born overseas becomes less pronounced when we account for time spent towards multiple qualifications. There is a weaker correlation between inferred education and time spent in multiple qualifications (0.165) for the overseas born, meaning accounting for that mismeasurement would result in a larger reduction in the estimated return to education for this group. With actual wages unchanged, this implies a higher estimated penalty associated with being born overseas, not a reduced penalty as the estimates imply. The explanation for these results may instead lie in the difference between Australian born and overseas born workers in terms of the relative pay-off to an initial versus a multiple qualification. If overseas born workers face a relatively low return to an initial qualification but relatively high return to multiple qualifications, then accounting for multiple qualifications will then explain some of the earnings penalty associated with being born overseas.

5.1 Effects by Gender and Age

To test for the possibility that the extent of mismeasurement in years of education varies between males and females, and between younger and older workers, wage equations were estimated separately for the sub-samples of men and women; and for persons aged 25-45 years and 46-65 years. Recall the estimates of additional years of education attributed to multiple qualifications are derived from HILDA for each individual, and additional years attributable to non-completion is differentiated by gender, age and highest qualification using ABS data. As shown in Table 8, for each group the difference in estimated returns to inferred and actual years of education is close to 1 percentage point per year, which equates to the returns to additional years of education being approximately 15 per cent lower once non-completion and multiple qualifications have been accounted for.

Table 8. Returns to Education Estimates by Gender and Age

Variable	*Male*	*Female*	*Age 25-45*	*Age 46-65*
Inferred Years of Education	0.0647***	0.0651***	0.0655***	0.0624***
	(34.63)	(39.93)	(46.55)	(31.62)
Actual Years of Education	0.0539***	0.0555***	0.0561***	0.0521***
	(33.31)	(39.36)	(45.30)	(30.24)
Difference	-0.0108	-0.0096	-0.0094	-0.0103
N	64302	61897	75803	50396

Notes: t-statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5.2 Sensitivity analyses

For reasons set out above, we consider the random effects specification preferable to estimation by fixed effects. Estimation by fixed effects gives a much lower overall estimate for the returns to inferred years of education of just 1.8 per cent. In line with the results above, the estimated returns to actual years of education is lower again, at 1.6 per cent, with that difference again accounted for by additional years spent accruing multiple qualifications as opposed to additional years in incomplete qualifications. In contrast to the random effects results, sensitivity of the estimates to the inclusion of actual years is observed in the fixed effects models only for males and younger workers.

To estimate the time individuals spent gaining multiple qualifications at a given level, we assumed a discount of 25 per cent from the time normally taken to complete a qualification to allow for recognition of prior learning, as explained in section 4.2. To test the sensitivity of the key findings to this assumption, we repeated the analysis using a 33 per cent discount, something we consider to be an upper bound for the average credits that students might be awarded. Compared to the estimated 6.51 per cent returns to a year of education using the inferred approach, this higher discount rate – or assumed lower time spent gaining multiple qualifications – produces an estimated 5.69 per cent wage premium per year of actual education, as opposed to our original estimate of 5.52 per cent. Ignoring years of non-completion, the estimated effect of accounting for years in multiple qualifications alone is to reduce the estimated returns to education from 6.51 per cent to 5.71 per cent (up from 5.53 per cent originally). Thus, even with this more conservative assumption, accounting for multiple qualifications still leads to a sizeable fall in the estimated returns to education, and still dominates the effect attributable to course non-completion.

Finally, we varied the assumption that individuals who fail to complete a course typically studied for one-third of the normal completion time before dropping out (see section 4.3). Even if we assume individuals remain in the course for two-thirds of the normal completion time before dropping out (i.e. twice as long as in our base case), the impact of years of course non-completion remains very modest. The overall estimate for the returns to actual years of education drops from 5.52 per cent to 5.49 per cent; while the estimate accounting for years of non-completion only drops from 6.47 per cent to 6.44 per cent. This is in line with the findings above that the estimated rates of return to education are relatively insensitive to accounting for years spent in incomplete qualifications.

7. Conclusion

Across the 2001-2019 HILDA panel, we find that accounting for the combined effect of course non-completion and multiple qualifications has a meaningful impact on the estimated returns to education in Australia. On average, the estimated increase in hourly wages associated with each additional year of education falls from 6.5 per cent to 5.5 per cent once accounting for course non-completion and multiple qualifications. If we accept the 5.5 per cent figure as a 'truer' estimate, this represents almost a full one percentage point (or 18 per cent) overestimation of the returns to education using the conventional approach. The extent of underestimation in years of education accrued, and associated over-estimation of the returns to education, seems broadly similar by gender and age.

Our exploration highlighted that the sensitivity of estimates of the returns to education to the inclusion of years of non-completion and multiple qualifications is actually quite nuanced: it does not lead to a simple reduction in the estimated rate of returns in proportion to the underestimation of time spent in education. First, the effect is dependent upon the within-sample correlation between mismeasurement and inferred years of education. If the extent of mismeasurement is distributed evenly across the sample, accounting for 'actual' years of education within the standard wage regression model will not alter the estimated returns to education, which is typically taken to be the average increase in wages associated with one additional year of education. Accounting for actual years could potentially lead to a higher estimate of the returns to education, if mismeasurement was negatively correlated with inferred education.

Second, the effects are dependent upon the relative wage premium associated with years spent in non-completed qualifications and in multiple qualifications, compared to the returns to years spent in completed and initial qualifications. Differences in these relativities may account for observed differences in earnings for groups within the labour market, such as those born overseas or from non-English speaking backgrounds.

In total, we estimate that, in any one year, the sample of workers in the HILDA survey have accrued around 2,700 combined years of education that are unaccounted for through course non-completion and multiple qualifications. While these 2,700 years are a legitimate part of the process of human capital accumulation, their associated costs have been omitted from most previous estimates of returns to education.

The key policy implication arising from these unaccounted-for years of education is that overestimating the value of post-school education can result in policy decisions that run the risk of significantly distorting markets for human capital in Australia, with the particular threat that the returns to some degree disciplines with historically low returns to education, such as humanities (Daly *et al.* 2015), may see returns further reduced, compounding the effect. Indeed, in 2021, the Commonwealth Government increased student contribution fees by up to 28 per cent for law and commerce degrees and up to 113 per cent for humanities degrees, whilst discounting fees for science, technology, engineering and mathematics (STEM) related disciplines by around 18 per cent to 42 per cent, depending on the selected course (Department of Education, Skills and Employment 2020), resulting in vulnerable degree disciplines now facing even lower returns to education.

To the best of our knowledge, this is the first paper to assess how accounting for years spent in multiple qualifications and course non-completion affects estimates from wage equations of the overall returns to education in Australia. We find that the return to an additional year of education is nearly one-fifth lower than standard estimates suggest. The analysis has also highlighted important differences between the Mincerean wage equation and IRR approaches. Under the IRR approach, including those additional years spent in education would unambiguously reduce the estimated returns to education, as it would account for higher foregone earnings and direct costs that are not incorporated into the wage equation approach. As rates of participation in post-school education in Australia continue to rise appreciably, policymakers must take caution when setting student contribution fees in order to avoid disproportionately punishing disciplines which already have lower returns to education. This is also likely to have equity implications, as groups targeted for increased higher education participation, such as students from low socioeconomic backgrounds, Indigenous students and from rural and remote areas, are also more likely to drop out and face lower returns to those investments. Hence, there is a need for ongoing research following both approaches.

7.1 Limitations and further research

Perhaps the key limitation to this study is the lack of individual-level data on time spent in incomplete courses. As a result, course non-completion statistics were applied through the use of population averages from ABS data, rather than at the individual level in the HILDA panel data. If years of non-completion could be applied at the individual level, as was possible for multiple qualifications, the estimated effect of course non-completion on the returns to education may have been significantly higher. This is particularly relevant given increases in course switching and dropouts within Australian education (Cherastidtham *et al.* 2018). To the best of our knowledge, no longitudinal datasets allow identification of time spent studying towards incomplete qualifications at the individual level. A possible exception is the Longitudinal Surveys of Australian Youth, but these track individuals only to age 25, providing a very limited window to observe wage effects.

A related priority for further research is for estimates of the actual wage premiums associated with years spent towards non-completed qualifications and with multiple qualifications. This has implications both for the overall returns to years of education, and the degree to which these represent 'wastage' that can potentially be eliminated through better course matching and higher student retention. The high proportion of persons with multiple qualifications at the one level seems at odds with signaling theory (Arrow 1973, Spence 1973), which would suggest that there would be no advantage to individuals undertaking a second qualification at the same level. Further, there may be significant differences in returns to incomplete and multiple qualifications that help explain differences in wages between sub-groups within the working population, particularly migrants.

There is an ongoing need for evaluation of the returns to education in Australia using both the wage equation and IRR approaches, differentiated by field of study, to guide education funding policy and the setting of course fees. We also propose a more specific focus on estimates of the return to post-school education, given the high proportion of youth who now complete Year 12, and that course non-completion and multiple qualifications relate primarily to years of post-school education. An interesting avenue for future research is examining whether years of education have differing returns for school, university or TAFE in Australia.

References

Arrow, K. J. (1973), 'Higher Education as a Filter', *Journal of Public Economics*, 2(3), 193-216. https://www.sciencedirect.com/science/article/abs/pii/0047272773900133

Australian Bureau of Statistics (2017), Australians Pursuing Higher Education in Record Numbers, Canberra, ACT: ABS https://www.abs.gov.au/AUSSTATS/abs@.nsf/mediareleasesbyReleaseDate/1533FE5A8541D66CCA2581BF00362D1D

Australian Bureau of Statistics (2020a), Consumer Price Index, Australia, Catalogue No. 6401.0, Canberra, ACT: ABS. https://www.abs.gov.au/AUSSTATS/abs@.nsf/ProductsbyReleaseDate/97DE913203378356CA2583E5001D660C?OpenDocument

Australian Bureau of Statistics (2020b), Education and Work, Australia, Catalogue No. 6227.0, Canberra, ACT: ABS. https://www.abs.gov.au/statistics/people/education/education-and-work-australia/latest-release

Australian Bureau of Statistics (2020c), Qualifications and Work, Catalogue No. 4235.0, Canberra, ACT: ABS. https://www.abs.gov.au/statistics/people/education/qualifications-and-work/latest-release

Australian Curriculum, Assessment and Reporting Authority. (2021), *Schools and Schooling*. Sydney: Australian Curriculum, Assessment and Reporting Authority. https://acara.edu.au/reporting/national-report-on-schooling-in-australia/national-report-on-schooling-in-australia-2013/schools-and-schooling/school-structures

Australian Qualifications Framework Council (2013), *Australian Qualifications Framework*, second edition, January.

Becker, G. S. (1964), 'Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education', 3rd edition, The University of Chicago Press. ISBN 0-226-04119-0. https://www.nber.org/system/files/chapters/c11226/c11226.pdf

Bednarz, A. (2014), 'Understanding the Non-completion of Apprentices', Occasional Paper, National Centre for Vocational Education Research, Adelaide.

Bonjour, D., Cherkas, L., Haskel, J., Hawkes, D. and Spector, T. (2003), 'Returns to Education: Evidence from U.K. Twins', *The American Economic Review*, 93(5), 1799-1812. http://www.jstor.org/stable/3132153

Borland, J., Dawkins, P., Johnson, D. and Williams, R. (2000), 'Returns to Investment in Higher Education', Melbourne Economics of Higher Education Research Program Report No. 1, University of Melbourne.

Card, D. (1999), 'The Causal Effect of Education on Earnings', In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 3 (1801-1863), Elsevier. https://doi.org/10.1016/S1573-4463(99)03011-4

Card, D. (2001), 'Estimating the Return to Schooling: Progress on some Persistent Econometric Problems', *Econometrica*, 69(5), 1127-1160. https://doi.org/10.1111/1468-0262.00237

Cherastidtham, I., Norton, A. and Mackey, W. (2018), 'University Attrition: What Helps and What Hinders University Completion', Grattan Institute. https://apo.org.au/sites/default/files/resource-files/2018-04/apo-nid142971.pdf

Corliss, M., Daly, A., and Lewis, P. (2020), 'Is a University Degree Still a Worthwhile Financial Investment in Australia?, *Australian Journal of Education*, 64(1), 73-90. doi:10.1177/0004944120901883

Corliss, M., Lewis, P., and Daly, A. (2013), 'The rate of return to higher education over the business cycle', *Australian Journal of Labour Economics*, 16(2), 219–236.

Curtin University (2020), Credit for Recognised Learning Manual, Western Australia: Curtin University. https://policies.curtin.edu.au/local/docs/draft-policy/Draft_Credit_for_Recognised_Learning_Procedures_May2020.pdf

Daly, A., Lewis, P. Corliss, M. and Heaslip, T. (2015), 'The Private Rate of Return to a University Degree in Australia', *Australian Journal of Education*, 59(1), 97-112. https://doi.org/10.1177/0004944114565117

Deloitte Access Economics (2017), 'Estimating the Public and Private Benefits of Higher Education', https://www.dese.gov.au/higher-education-reviews-and-consultations/resources/estimating-public-and-private-benefits-higher-education

Department of Education, Skills and Employment (2020), 'Job-Ready Graduates: Higher Education Reform Package 2020'. https://www.voced.edu.au/content/ngv:87007

Dickson, M. and Harmon, C. (2011), 'Economic Returns to Education: What We Know, What We Don't Know, and Where We Are Going – Some brief pointers', *Economics of Education Review*, 30 (6), 1118-1122. https://doi.org/10.1016/j.econedurev.2011.08.003

Dockery, A.M. and Miller, P.W. (2012), 'Over-Education, Under-Education and Credentialism in the Australian Labour Market', NCVER Monograph Series 10/2012, National Centre for Vocational Education Research Ltd. PO Box 8288, Station Arcade, Adelaide, SA 5000, Australia. https://files.eric.ed.gov/fulltext/ED535568.pdf

Edwards, D. and McMillan, J. (2015), 'Completing University in a Growing Sector: Is Equity an issue?, Australian Council for Educational Research, August.

Giani, M.S., Attewell, P. and Walling, D. (2020), 'The Value of an Incomplete Degree: Heterogeneity in the Labor Market Benefits of College Non-Completion', *The Journal of Higher Education*, 91(4), 514-539. https://doi.org/10.1080/00221546.2019.1653122

Griffith University (2019), Credit and Recognition of Prior Learning Policy, Queensland: Griffith University. https://policies.griffith.edu.au/pdf/Credit%20and%20Recognition%20of%20Prior%20Learning%20Policy.pdf

Harmon, C., Oosterbeek, H. and Walker, I. (2003), 'The Returns to Education: Microeconomics', *Journal of Economic Surveys*, 17(2), 115-156. https://doi.org/10.1111/1467-6419.00191

Hausman, J. A. (1978), 'Specification Tests in Econometrics' *Econometrica*, 46(6), 1251-1271. https://www.jstor.org/stable/1913827

Heckman, J. J., Lochner, L. J. and Todd, P.E. (2006), 'Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond', In *Handbook of the Economics of Education*, 1st edition, edited by Eric A. Hanushek and F. Welch, 307–458. Amsterdam: Elsevier.

Heckman, J.J., Humphries, J.E. and Veramendi, G. (2016), 'Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking', Working paper 22291, DOI 10.3386/w22291

Hodgson, D. (2011), 'Policy Rationalities and Policy Technologies: A Programme for Analysing the Raised School-Leaving Age in Western Australia', *Journal of Education Policy*, 26(1), 115-130. https://doi.org/10.1080/02680939.2010.498901

Jaeger, D. A. and Page, M.E. (1996), 'Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education', *The Review of Economics and Statistics*, 78(4), 733-740. https://www.jstor.org/stable/2109960

Kler, P. (2005), 'Graduate Overeducation in Australia: A Comparison of the Mean and Objective Methods', *Education Economics*, 13(1), 47-72. https://doi.org/10.1080/0964529042000325207

Leigh, A. (2008), 'Returns to Education in Australia', Economic Papers: A *Journal of Applied Economics and Policy*, 27(3), 233-249. https://doi.org/10.1111/j.1759-3441.2008.tb01040.x

Li, H., Liu, P.K. and Zhang J. (2012), 'Estimating Returns to Education using Twins in Urban China', *Journal of Development Economics*, 97(2), 494-504.

Long, M. and Shah, C. (2008), 'Private Returns to Vocational Education and Training Qualifications. A National Vocational Education and Training Research and Evaluation Program Report', National Centre for Vocational Education Research Ltd, Adelaide, https://files.eric.ed.gov/fulltext/ED503388.pdf

Marks, G. (2007), 'Completing University: Characteristics and Outcomes of Completing and Non-completing Students', Australian Council for Educational Research (ACER). https://research.acer.edu.au/lsay_research/55

Melbourne Institute (2020), 'HILDA Survey', University of Melbourne. https://melbourneinstitute.unimelb.edu.au/hilda

Miller, P., Mulvey, C. and Martin, N. (1995), 'What Do Twins Studies Reveal about the Economic Returns to Education? A Comparison of Australian and U.S. Findings', *American Economic Review*, 85(3), 586-599.

Mincer, J. (1974), 'Schooling, Experience, and Earnings', *Human Behavior & Social Institutions No. 2*. https://EconPapers.repec.org/RePEc:nbr:nberbk:minc74-1

Montenegro, C.E., Patrinos, H.A. (2014), 'Comparable Estimates of Returns to Schooling around the World', World Bank Policy Research Working Paper Series 7020 https://documents1.worldbank.org/curated/en/830831468147839247/pdf/WPS7020.pdf

Norton, A. (2012), 'Graduate Winners: Assessing the Public and Private Benefits of Higher Education', Melbourne: Grattan Institute. https://grattan.edu.au/report/graduate-winners-assessing-the-public-and-private-benefits-of-higher-education/

Norton, A., Cherastidtham, I. and Mackey, W. (2018), 'Dropping Out: The Benefits and Costs of Trying University', Melbourne: Grattan Institute. https://grattan.edu.au/wp-content/uploads/2018/04/904-dropping-out-the-benefits-and-costs-of-trying-university.pdf

Productivity Commission (2019), 'The Demand Driven University System: A Mixed Report Card', Commission Research Paper, Canberra.

Psacharopoulos, G. and Patrinos, H.A. (2018), 'Returns to Investment in Education: A Decennial Review of the Global Literature', *Education Economics*, 26(5), 445- 458. https://doi.org/10.1080/09645292.2018.1484426

Rospigliosi, A.P., Greener, S., Bourner, T. and Sheehan, M. (2014), 'Human Capital or Signalling, Unpacking the Graduate Premium', *International Journal of Social Economics*. 41(5), 420-432. https://doi.org/10.1108/IJSE-03-2013-0056

Schultz, T. W. (1961), 'Investment in Human Capital', *The American Economic Review*, 51(1), 1-17. https://www.jstor.org/stable/1818907

Sinning, M. (2014), 'How Much is it Worth? New Estimates of Private Returns to University Education in Australia', University of Queensland, RWI and IZA. https://www.aph.gov.au/~/media/Committees/eet_ctte/estimates/add_1415/Education/Answers/SQ15-000059AttachmentB.pdf

Spence, M. (1973), 'Job Market Signaling', *The Quarterly Journal of Economics*, 87(3), 355-374. https://www.jstor.org/stable/1882010

Stromback, T. and Mahendran, A. (2010), 'An Analysis of Factors Contributing to Apprenticeship and Traineeship Completion', *International Journal of Training Research*, 8(1), 63-79.

Toutkoushian, R., Shafiq, M., and Trivette, M. (2013), 'Accounting for Risk of Non-Completion in Private and Social Rates of Return to Higher Education', *Journal of Education Finance*, 39(1), 73-95. https://eric.ed.gov/?id=EJ1082312

University of Melbourne (2020), 'HILDA Online Data Dictionary Release 18.0, DV: Highest education level achieved' https://www.online.fbe.unimelb.edu.au/HILDAodd/VariableDetails.aspx?varn=edhigh1&&varw=1

University of South Australia (2019), 'Recognition of prior learning', South Australia: University of South Australia. https://i.unisa.edu.au/siteassets/policies-and-procedures/docs/academic/a13-recognition-of-prior-learning.pdf

Watson, N. and Wooden, M. (2013), 'Adding a Top-Up Sample to the Household, Income and Labour Dynamics in Australia Survey', *Australian Economic Review*, 46 (4), 489-498. https://doi.org/10.1111/1467-8462.12027

Wei, H. (2010), 'Measuring Economic Returns to Post-School Education in Australia', Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1351.0.55.032Aug%202010?OpenDocument

Wilkins, R. (2015), 'The Household, Income and Labour Dynamics in Australia Survey: Selected Findings from Waves 1 to 12', Melbourne: Melbourne Institute of Applied Economic and Social Research, University of Melbourne. https://core.ac.uk/reader/30671650

Zeidenberg, M., Scott, M. and Belfield, C. (2015), 'What about the Non-completers? The Labor Market Returns to Progress in Community College', *Economics of Education Review*, 49, 142-156. https://doi.org/10.1016/j.econedurev.2015.09.004

Appendix A: Variable description and means

Variable	*Description/coding*	*Sample mean*
Log hourly wage (dependent variable)	See main text	3.29
'Inferred' years of education	See main text	13.03
'Actual' years of education	See main text	13.43
Years of education – multiple qualifications	See main text	0.26
Years of education – incomplete qualifications	See main text	0.15
Male	Dummy variable equal to 1 if the respondent is male, 0 if female.	0.51
Age (& age-squared)	The age of the respondent in years (and its square/100)	42.22 (18.98)
Marital status	3 dummy variables corresponding to three marital status categories: married (the omitted category), never married or no longer married	Married 0.58; Never married 0.14; No longer married 0.28
Work experience (and work experience squared)	The amount of time spent in paid work in years (and its square/100)	21.71 (5.93)
Disability status	Dummy variable equal to 1 if the respondent has a long-term health condition, disability or impairment; 0 otherwise	0.13
Language other than English spoken	Dummy variable equal to 1 if respondent speaks another language other than English at home; 0 otherwise	0.14
Employment status	3 dummy variables corresponding to employment status: permanent (the omitted category), fixed-term or casual.	Permanent 0.75; Fixed-term 0.10; Casual 0.15.
Union membership	Dummy variable equal to if 1 if respondent is a union member; 0 otherwise	0.25
Socio-economic decile of neighbourhood	Dummy variable equal to 1 if respondent is living in area in decile 6 to 10; 0 if living in area in decile 1-5.	0.55
Part-time work	Dummy variable equal to 1 if labour force status is equal to part-time; 0 if full-time	0.27
Overseas highest qualification	Dummy variable equal to 1 if respondent completed their highest qualification overseas; 0 if completed in Australia	0.28
Overseas born	Dummy variable equal to 1 if respondent was not born in Australia; 0 otherwise	0.22