Evidence on credit constraints and university attendance.*

Buly A Cardak[†] Chris Ryan[‡]

July 2016

Abstract

The effects of credit constraints on university participation are investigated. Students most likely to face credit constraints have the same or higher probability of attending university as all other students, given their high school achievement. A novel approach to handle potential bias arising from unobservable heterogeneity is proposed and implemented, providing estimated unobservables based on post secondary plans reported during ninth grade. These estimated unobservables are important in explaining university attendance but results regarding the effects of credit constraints are robust to their inclusion in models estimated. Our new estimated unobservables approach is consistent with a related Copula estimation approach.

JEL Classification: I21, I22, I28

Keywords: University participation; Credit constraints; Unobservables; Copula estimation.

*We are grateful for comments from David Prentice and Joe Vecci. Cardak acknowledges support from a research grant provided by the Australian Research Council (DP0662909) and is grateful for the hospitality extended by the Melbourne Institute where part of this work was undertaken.

†School of Economics, La Trobe University, 3086, Victoria, Australia. Ph:+61 3 9479 3419, Fax: +61 3 9479 1654, Email: b.cardak@latrobe.edu.au.

‡Melbourne Institute of Applied Economic and Social Research, University of Melbourne, 3010, Victoria, Australia. Ph: +61 3 9035 3794, Fax: +61 3 8344 2111, Email: ryan.c@unimelb.edu.au.

1 Introduction

The role of credit constraints is a central issue in higher education. A diverse set of responses have been adopted by Governments and universities worldwide in recognition of the importance of resolving credit constraint issues and enabling students to attend university. Forms of support range from full public funding of higher education and generous income support in Scandinavian countries such as Denmark, Finland, Iceland, Norway and Sweden, through to a diverse collection of grants, loans and work-study that form the federal aid system in the US; see OECD (2013).

Despite this diversity of policies intended to support students to attend university, a socioeconomic status university attendance gradient persists and can can be thought of as an empirical regularity in higher education across the world; Blossfeld et al. (2016) and Blossfeld and Shavit (1993). The natural tendency is to attribute the SES gradient to credit constraints. In this paper we study the effects of credit constraints on university attendance in Australia. We also develop a novel approach to dealing with unobserved heterogeneity where we exploit student plans to attend university in order to obtain an estimate of unobservables affecting student decisions to attend university. This analysis is complemented with a joint estimation of models of the plans to attend university and the attendance decision with dependence modelled using a Copula specification.

In order to identify credit constrained students we exploit the fact that 35% of students attend fee charging private schools in Australia. Based on private school attendance, supplemented with socioeconomic status (SES) data, we form three categories of students, those unlikely and likely to be constrained and those that are potentially constrained; more detail is provided below. The strategy infers that families that were able and willing to pay for private education when students were in ninth grade are unlikely to face credit constraints. Over the relevant period for our data, average tuition fees in private independent secondary schools were higher than university tution charges; for example in 1998, average school fees were $6,123 while the highest university charge was $5,772 in 2000. Given the income contingent loan scheme in place, called the Higher Education Contribution Scheme (hereafter HECS), credit constraints are more likely to operate on day to day living costs while studying and the costs of books and other study materials.

We find very little difference in university participation between the different credit constraint groups. What differences we do find depend on high school achievement and are opposite to what might be expected. At low levels of high school achievement, likely constrained students are up to 11% more likely to attend university than unlikely constrained students. At higher levels of high school achievement, these differences are not present. These results suggest that short term credit constraints do not seem to deter students from attending university in Australia. The results also suggest that lower achieving likely constrained high school graduates are more optimistic about university study than otherwise similar students who are unlikely to be or potentially credit constrained.

Given these results, one might speculate that likely constrained students choosing to attend university might be more motivated or have some other unobservable characteristics that lead to our finding. We devise a novel strategy to address concerns that these results might be driven by selection on unobservables. Students are first surveyed in ninth grade and answer questions about plans for post-secondary study. These survey responses are used to estimate unobservables that might be important for the university attendance decision. The estimated unobservables approach compares student post school study plans with average plans for observably similar students. This comparison is implemented using group averages, non-parametric estimates and a parametric regression approach. Differences between expected plans and actual plans are attributed to unobservable factors that influence post school study plans. Essentially, the differences reflect how unusual it is for an individual to hold the plans they express, given the plans of otherwise similar individuals. The resulting estimated unobservables are then included in the model of university attendance to control for potentially important, yet unobserved factors. Our estimated unobservables approach provides a flexible robustness check in our context and provides another method of extracting useful information from the expectations of individuals in empirical work. Examples of other approaches to the use of student expectations in studying higher education decisions can be found in Arcidiacono et al. (2012), Dominitz and Manski (1996) and Zafar (2011).

Upon re-estimation of our baseline model with these estimated unobservables included, we find the estimated unobservables are indeed significant in explaining the university attendance probability. However, the effects of credit constraint group membership on university participation are qualitatively unchanged and robust to the inclusion of these unobservables.

We complement our estimated unobservables approach with a further robustness check which exploits the same information. This involves jointly estimating models of the university attendance decision and plans for post-secondary study while allowing for dependance between the disturbance terms of these two equations. The dependence is modelled via a Copula specification as proposed in Trivedi and Zimmer (2007). This approach identifies unobservables common to both the attendance decision and plans, and thereby corrects for any bias in the credit constraint group effects estimated in the baseline specification. A wide range of Copula specifications are considered. As before, we find allowing for this dependence is important, but the results do not change our conclusions about credit constraints. Students who are likely to face credit constraints are no less likely to attend university than all other students, conditional on their high school attainment.

These findings are consistent with a range of results about credit constraints in the US. Studies by Carneiro and Heckman (2002), Cameron and Taber (2004) and Keane and Wolpin (2001) all find little evidence of credit constraints hindering university attendance in the US, based on the NLSY 79 cohort.[1] However, Belley and Lochner (2007) study the importance of credit constraints using the more recent NLSY 97 cohort and find that youth from high income families are 16% more likely to attend university than those from low income families, after conditioning on a range of factors. After accounting for wealth along with income, the difference in attendance between high and low groups is 30 percentage points. They also find strong evidence of an ability gradient in attendance which is consistent across all income

[1] In a study using UK data, Dearden et al. (2004) also find little evidence that credit constraints affected university attendance.

levels and both NLSY cohorts.

These studies focus on parental income as a measure of credit constraints. In a related study, Brown et al. (2012) show that parental income might not be the best indicator of credit constraints if parents are unwilling or unable to contribute to college costs of the student. They exploit information on federal aid rules in the US, sibling spacing and financial gifts from parents to identify the effects of credit constraints. They find that many students who might not be thought to be constrained on the basis of income do indeed face constraints. In a similar vein, the approach adopted here does not rely on income but rather the ability and willingness of parents to invest in their child's education. Another identification strategy is employed by Coelli (2011) who studies unexpected job loss among Canadian fathers. He finds large negative effects on university attendance for those who were 16 to 17 years old when the father's job loss was experienced, indicating negative effects of short term credit constraints even in the presence of student loans.

As we find credit constraints do not play an important role in the university attendance decision, we analyse our results further, presenting a decomposition of what are the most important factors in explaining differences between the university attendance decisions of high and low SES students. The decomposition exploits observables, ninth grade test scores, the estimated unobservables described above and high school achievement. The key message is that the SES gradient is driven overwhelmingly by high school achievement rather than credit constraints, suggesting for those students that qualify for university, the funding institutions in place are able to adequately deal with any credit constraint issues that students might face.[2] These findings regarding the importance of school achievement are a common theme in the literature on disadvantaged student's access to higher education; see for example Chowdry at al. (2013) and Cunha et al. (2006) for UK and US evidence respectively. The impact of the abolition of tuition fees in Ireland is studied by Denny (2014) where, consistent with our findings, high school achievement of disadvantaged students dominates

[2] In related work by Cardak and Vecci (2016) it is found that credit constrained students exhibit a greater risk of dropout. This is in contrast with Stinebrickner and Stinebrickner (2008) who find that dropout is not affected by credit constraints for Berea College students in the US.

any impact of changes in tuition charges.

In Section 2, we provide a brief description of the educational institutions in place in Australia that are relevant to our study, including the HECS and government income support schemes for students. The empirical strategy, including our approach to controlling for the possible effects of unobservables, is presented in Section 3. We describe the data in Section 4, providing definitions of the credit constraint groups. The results, including the estimated unobservables and Copula robustness tests, are presented in Section 5. We summarise our key results and conclude in Section 6.

2 Overview of School to University Transition

University admission in Australia is similar to other countries. In order to gain a university place, students are required to (i) demonstrate academic achievement or aptitude and (ii) be willing and able to pay any tuition charges. However, there are idiosyncracies that are important for our analysis which we outline below.

In order to demonstrate academic achievement or aptitude, students need to complete high school, receiving the relevant certificate from their state body. University applications are typically handled by a state based central admissions body. Students are required to apply for an Equivalent National Tertiary Entrance Rank (ENTER) score which is supplied to the central admissions body. The ENTER score is based on their achievement in statewide examinations and other assessment tasks and reflects the student's percentile rank in the graduating cohort for their home state. It is the primary mechanism by which university places are rationed.[3] Courses of study in higher demand typically have higher ENTER score thresholds for admission, thus these ENTER scores may be considered part of the price of admission to a program.[4]

[3]The ENTER score was specific to the state of Victoria. However, each state provided graduating students with an equivalent ranking, for example a Universities Admissions Index (UAI) in New South Wales. These rankings could be used to apply for university places out of the student's home state. We use the term ENTER score as a generic name for these entrance ranks which are calibrated to a common, Australian-wide scale that ranges from 30 to 99.95. Since 2010, these percentile ranks have been renamed Australian Tertiary Admission Rank (ATAR) for all states and territories except Queensland.

[4]Admission to university on the basis of ENTER scores is the dominant mode for matriculating students

Virtually all universities in Australia are public and since 1989 students have faced a tuition charge. This tuition charge may be paid up front or students may defer payment by agreeing to an income contingent loan provided by the federal government. These income contingent loans were referred to as HECS loans.[5] HECS loan balances are indexed to inflation, thereby offering students a zero real interest rate. Repayments are administered through the federal tax system and are only required when taxable income exceeds a legislated level; see Chapman (1997) for more details. In addition to HECS, students could apply for means tested government support (called AUSTUDY) while studying. This means tested support is based on parental income and available in the later years of high school and through university. It takes the form of regular government payments while studying, essentially a stipend, where the amount varies with the level of assessed need. However, only small proportions of university students are eligible for such support (34% in our sample); see Ryan (2013) for a detailed discussion of these payments.

While the institutional and policy environment in Australia seems to address short term credit constraints at least to some degree, we still observe a strong socioeconomic status (SES) gradient in university attendance. Potential students may be credit constrained if they are unable to fund costs such as books, study materials and living expenses that are not covered by the HECS scheme, or if they do not qualify for sufficient income support while studying. The SES gradient is illustrated using the Longitudinal Surveys of Australian Youth (LSAY) data in Figure 1. A nonparametric estimate of the probability of university attendance conditional on SES is presented and highlights the highest SES students are nearly 3 times more likely to attend university than the lowest SES students.[6]

In order to identify students that are credit constrained, we exploit the fact that 35% of students attend some form of private school in Australia.[7] Thus, while public education

in Australia. Other criteria are used for mature aged entrants which form a smaller part of the student body and are not the focus of our analysis.

[5]Since 2005, the HECS scheme has been renamed the HECS-Higher Education Loan Program (HECS-HELP).

[6]The conditional means shown in figures in this paper were all estimated using the *lowess* or *mlowess* programs in Stata.

[7]The Australian non-government school sector is typically divided into the Catholic sector (20%) and the

is the dominant mode of school education in Australia, a large proportion of the population is willing and able to opt out of the public system and pay for education, indicating these households might not face educational credit constraints. As the private school sector exhibits a wide variety of schools with different fee structures, we cannot rely entirely on private school attendance as an indicator of the absence of credit constraints. Instead, we complement private school attendance with other data to identify households that are less likely to face credit constraints. We use additional information about the SES of the student's parents and about the average SES of the school attended. We define three groups. The first comprises students who are least likely to face credit constraints. They have relatively high SES backgrounds and attend private schools where the average SES of the school's students is also high. The second group comprises students most likely to face credit constraints, with relatively low SES and attending schools where the average SES of the school's students is also low. The third comprises all other students and is referred to as potentially constrained. We define an indicator variable for membership of each of these three groups, using group membership to investigate the impact of credit constraints on university participation. We provide further details below when describing our empirical approach and data.

3 Empirical Method

Optimizing students making the decision to attend university face a number of constraints and trade-offs and will choose to attend university if they believe the benefits from attending will make it worthwhile relative to any costs they face. The student's assessment of this will depend on their attitude towards education, a range of personal attributes related to university study and their personal circumstances at the time they decide to undertake university study, including the likelihood that they are credit constrained. These factors comprise both 'permanent'and 'temporary'components, some possibly unobserved, that may influence the decision to attend university.

Individuals with similar observed demographic characteristics may make different deci-

independent sector (15%).

sions if their underlying unobserved attitudes towards education or their circumstances at critical times (current or past) differ. The following empirical model reflects these various features of the university participation decision:

$$u_{it}^* = \alpha + y_{it}\kappa + G_{it}'\tau + (y_{it} \times G_{it}')\varphi + X_{it}'\beta + c_i + e_{it}. \tag{1}$$

The desired level of university participation by individual i at time t is latent and given by u_{it}^*. This desired level of participation depends on a range of factors observed at time t, including a vector of demographic characteristics, X_{it}, where β is a vector of parameters reflecting the importance of these characteristics. The student's final high school achievement (ENTER score) is given by y_{it}, with the parameter κ capturing the importance of this achievement. Participation also depends on whether the student faces credit constraints. This is captured by G_{it} which indicates whether a student is likely or unlikely to face credit constraints, based on the groups described in the previous section and formally defined below. The importance of credit constraints is captured by the parameter vector τ, as well as the parameter on the interaction between group membership and ENTER φ. Along with these observed factors, participation depends on unobserved factors. We divide these unobserved factors into transitory and permanent components. Transitory unobserved factors operate at the time of the attendance decision and are denoted e_{it}. The permanent factors are denoted c_i and include unobserved attitudes towards education, individual motivation and ambition along with personal discount rates. We distinguish these two components as we outline and implement a strategy to account for permanent unobserved factors below.

The observed university participation decision is denoted u_{it} and modelled using a probit specification where university attendance, conditional on y_{it}, G_{it} and X_{it}, is given by:

$$P(u_{it} = 1 | y_{it}, G_{it}, X_{it}, c_i) = \Phi(\alpha + y_{it}\kappa + G_{it}'\tau + (y_{it} \times G_{it}')\varphi + X_{it}'\beta + c_i), \tag{2}$$

where Φ denotes the standard Normal cumulative distribution function (*cdf*). As in equation (1), this specification highlights the importance, for university participation, of high school achievement in the form of ENTER score, y_{it}, and ability to pay any costs of university education captured by the group membership dummies, G_{it}. Given the interaction term $(y_{it} \times G'_{it})$, testing the joint significance of τ and φ indicates whether credit constraint group membership affects the level of university participation.

3.1 Potential effects of unobservables

Recall that c_i captures unobserved heterogeneity in individual attitudes towards university education that are consistently held through time. Typically, equation (2) will be estimated without data on c_i. Exclusion of c_i from (2) will lead to biased parameter estimates, most notably to attenuated estimates of all parameters, though the ratios of one parameter to another are not affected; see Wooldridge (2010). In our case, however, there are grounds for more specific concerns. Suppose members of each group are differentially likely to complete school and obtain an ENTER score, possibly because their attitudes towards university education differ. In this case, parameter estimates of group effects (τ and φ) will be influenced by these differences in average group attitudes.

To illustrate the potential problem, consider all students who complete high school and obtain an ENTER score. Suppose those students from the most likely to face credit constraints group do indeed have much more positive attitudes towards university education than those from the unlikely to face credit constraints group. These different attitudes could confound estimates of the impact of group membership and thus credit constraints. The presence of such unobserved heterogeneity may lead to estimates of little difference in participation between groups, but only because the effect of heterogeneous attitudes offsets the impact of group membership on university participation. We now outline our strategy to deal with this potential problem.

3.2 Two approaches to deal with unobservables

We adopt two closely related approaches to deal with this potential problem. First, we obtain an estimate of c_i based on post school educational plans individuals reveal about themselves in Year 9 (ninth grade). We include this estimate in the university participation equation to test if the parameters on our credit constraint group variables change. We refer to this as the estimated unobservables approach. The second approach is to jointly estimate the university plans and participation equations of students, taking account of any correlation between the errors that the unobserved individual heterogeneity would induce via copula estimation as discussed in Trivedi and Zimmer (2007). Each approach is described below in turn.

To obtain an estimate of c_i, we compare student's reported post-school study plans in Year 9 with some community or group 'norm' in order to establish how 'different' a student's plans are relative this 'norm'. A simple representation is given by:

$$\hat{c}_{ih} = u_{i1}^{plan} - \tilde{u}_{h1}^{plan} \tag{3}$$

where u_{i1}^{plan} is a dichotomous variable that reflects reported university attendance plans for individual i in period 1 of the survey, when students are in year 9. The community or group 'norm' is denoted $\tilde{u}_{h1}^{plan}$ and can be thought of as the average plan, estimated over a set of all individuals in group h, or a conditional mean $\tilde{u}_{h1}^{plan} \equiv E(u_{i1}^{plan} \mid i \in h)$. Some intuitive examples include average plans by gender or average plans across all survey subjects from the same school. It is also possible to condition on multiple characteristics, like gender and school membership at the same time. These comparisons provide a measure of how students differ relative to their group average, offering an estimate of the role of factors outside group membership. Where the group 'norm' accounts for a wide range of characteristics, $\hat{c}_{ih}$ provides an estimate of unobservables, and is included in the estimation of equation (2) to test if the estimates of τ and φ are affected by these unobservables.

At the individual level, the estimated heterogeneity term from equation (3) includes both permanent and transitory factors that induce the individual to have a view that departs from the norm of the group with which their plans are being compared. The approach deals with differences in unobserved heterogeneity for our credit constraint groups that might affect the parameter estimates because the expected value of the transitory element is zero for each of the credit constrained groups, but the permanent part is not.[8]

Year 9 plans are likely to be influenced by the same kind of individual characteristics as those included in equation (2). As a consequence, a more sophisticated regression-based 'norm' is also implemented. We estimate $\tilde{u}_{h1}^{plan}$ as the predicted probability of plans to attend university, conditional on personal characteristics observed in the *first* wave of the survey, denoted X_{i1}, rather than at the time of the university attendance decision. Plans will also be influenced by academic achievement measured in the first wave, p_{i1}. Idiosyncratic factors such as motivation, ambition and aspects of credit constraints not related to SES will also influence whether young people plan to attend university. These latter factors are unrelated to X_{i1} and are expected to be reflected in c_i. This regression-based 'norm' is estimated through the following model:

$$u_{i1}^{plan} = P(u_{i1}^{plan} = 1 | X_{i1}, p_{i1}) + \epsilon_{i1} = \Phi(X_{i1}'\delta + p_{i1}\lambda) + \epsilon_{i1}, \tag{4}$$

where X_{i1} and p_{i1} are defined above, with respective parameters δ and λ. The error term ϵ_{i1} reflects the unobserved idiosyncratic factors that influence whether young people *plan* to attend university when surveyed in the first wave.

In this case, groups are based on individual characteristics weighted by parameter estimates from equation (4). In terms of equation (3), we have $h = i$, $\tilde{u}_{i1}^{plan} = \Phi(X_{i1}'\hat{\delta}+p_{i1}\hat{\lambda})$ and $\hat{c}_i = \hat{\epsilon}_{i1}$. This estimated residual includes the unobserved permanent factors that influence the decision to attend university, c_i, that we wish to include in equation (2). Intuitively, the highest (lowest) value $\hat{\epsilon}_{i1}$ will correspond to students with poor (good) characteristics who *surprisingly* plan (not) to attend university. We include $\hat{\epsilon}_{i1}$ in the estimation of equation (2) in order to address any possible bias in parameter estimates resulting from the omis-

[8] To reduce the role of transitory factors in the calculation of the permanent heterogeneity element at the individual level, we also estimated $\hat{c}_{ih}$ from plans to attend university reported in each survey taken between Years 9-12. Averaging the heterogeneity term, $\hat{c}_{ih}$, over multiple years reduces the importance of transitory factors that might be confounded with the permanent factors of interest. Based on these averaged heterogeneity measures, results for the credit constraint group variables are qualitatively similar to those that are reported using post school plans from Year 9 only. The parameters on the heterogeneity terms increase in value, as would be expected with a reduction in what could be thought of as measurement error. However, we focus on the Year 9 estimates to ensure the unobservable heterogeneity estimates are based on the largest possible number of observations.

sion of c_i, comparing parameter estimates on the credit constraint group variables with and without the estimated unobservables.[9] It should be noted that X_{i1} does not include the constrained/unconstrained group identifiers as we do not wish to impose any group-based structure on the parameter estimates in equation (4). However, X_{i1} does include parental SES since this is an important determinant of student plans.

Our second approach to address concerns about the potential unobserved heterogeneity embodied in c_i involves jointly estimating student plans to attend university and actual attendance decisions, along with the dependence between the residuals of these equations. This joint estimation enables us to correct for any bias arising from unobservables common to both the attendance decision and stated Year 9 plans to attend university. It is similar in spirit to the estimated unobservables approach, using the regression based 'norm' outlined above, albeit in a more formal structure. We jointly estimate models of both decisions as:

$$u_{it} = P\left(\alpha_1 + y_{it}\kappa_1 + G'_{it}\tau_1 + (y_{it} \times G'_{it})\varphi_1 + X'_{it}\beta_1\right) + \left(c_i + \xi_{it}\right), \quad (5)$$

$$u_{i1}^{plan} = P\left(\alpha_2 + p_{i1}\kappa_2 + X'_{i1}\beta_2\right) + \left(c_i + \zeta_{i1}\right), \quad (6)$$

where ξ_{it} and ζ_{it} are disturbance terms and all other variables are as already defined. These discrete marginal equations are estimated via univariate Probit equations within a Copula specification that allows the dependence between the errors to be specified in a variety of ways set out below. In this case, $(c_i + \xi_{it})$ and $(c_i + \zeta_{i1})$ are the respective disturbance terms that include the unobserved permanent heterogeneity reflected in c_i. The Copula functions used here include the Gaussian (equivalent to estimation of the two equations via bivariate probit), Frank, Clayton, Gumbel and Joe Copulas. The estimation is undertaken with the Stata Copula command, *bicop*, detailed in Hernández-Alava and Pudney (2016). Equations (5) and (6) are estimated using observations of students who completed high school and obtained valid ENTER scores, y_{it}, as required in equation (5). Estimates of τ_1 and φ_1 based on various Copula functions are compared to estimates where the potential correlation is ignored. If these estimates are similar, we argue that they are robust to the omission of any

[9] In addition to actual values of the estimated error terms, we also considered $rank(\hat{\epsilon}_{i1})$, and various other transformations of the regression based estimate of c_i with little change in results.

unobserved heterogeneity from c_i.

4 Data

The data used in this paper is drawn from the Longitudinal Surveys of Australian Youth (LSAY). The similarity between the 1995 (LSAY 95) and 1998 (LSAY 98) cohorts along with the absence of any major institutional changes around the times when these cohorts made university decisions allows us to pool these two cohorts in our analysis. We include dummies in order to control for cohort effects.[10]

These cohorts are drawn from two-stage cluster samples of Australian school children. In the first stage, schools were randomly selected. In the second stage, intact classes of Year 9 students from those schools were randomly selected. The samples were stratified by school sector (government, Catholic or independent private schools). Population means in this paper are estimated with weighted data to account for this stratification along with attrition. In the first survey year, when students were in Year 9, they completed literacy and numeracy tests at their schools, along with a short questionnaire to elicit background information.[11] Participants were surveyed in subsequent years by mail and/or telephone questionnaires. In their fifth and subsequent contact years in both surveys, subjects were asked whether they had received the relevant certificate from their jurisdiction to indicate they had completed Year 12, whether they had obtained an ENTER score and whether they were studying at university.

We define three groups of students who face differing degrees of credit constraints. The first group is *unlikely* to face credit constraints and is defined to include students who in ninth

[10] Descriptive statistics for the pooled and separate LSAY 95 and LSAY 98 cohorts can be found in Table A.1 of Appendix A in Cardak and Ryan (2014). When we repeated the analysis for a more recent but smaller LSAY 06 cohort, the results and conclusions are unchanged.

[11] Student performance in ninth grade literacy and numeracy tests were used by Rothman (2002) to construct achievement scales. The individual literacy and numeracy scales were constructed to have a mean of 50 and standard deviation of 10. In this paper, we use the average of these two scales to reflect individual student achievement. This average has a standard deviation of 8.5. Where only one of the literacy and numeracy scales is available, it was used as the achievement score. This affected about 1.9% of observations used in the analysis.

grade (i) were in the top SES quartile, based on their parent's occupation;[12] (ii) attended a school in the top SES quartile of schools, based on the average of parents' occupation in schools; and (iii) attended a non-government school. The third criterion ensures that only individuals whose families had already demonstrated a preparedness to pay for at least part of their child's education were included in this group. The first two criteria are designed to pick out those students with the highest social backgrounds at the schools where such students are most concentrated. This group constitutes 8.1% of the weighted sample data.

In contrast, the second group is *likely* to face credit constraints and includes students who in ninth grade (i) were in the lowest SES quartile; and (ii) attended schools in the lowest average SES quartile of schools. This group constitutes 10.8% of the weighted sample data.

The third group comprises students who are *potentially* constrained and is defined as all students who are not members of the *unlikely* and *likely* to be credit constrained groups.[13]

The size of and summary statistics for these groups are presented in Panel A of Table 1. Consistent with the Australian and international evidence, members of the *unlikely* constrained group had the highest university participation rate and the highest average ENTER score. In turn, the middle group had a lower average ENTER score and university participation rate, while the *likely* constrained group had the lowest average ENTER score and university participation rate.

As a cross check, we provide evidence on how well these criteria have partitioned ninth grade students according to other indicators of wealth and social background in Panel B of Table 1. These other indicators include a social status index based on parental education and occupation constructed by the Australian Bureau of Statistics (1998), and whether students received government student income support (AUSTUDY) while attending university.

[12]The Australian National University (ANU) 3 scale is used for the two cohorts; Jones (1989). The scale is a status-based occupational prestige measure that lies between 0 and 100. The relevant 'parent' is the student's father unless information about his occupation is missing. In those circumstances, information on the occupation of the student's mother is used.

[13]Variations in these definitions were used to test how sensitive the results reported below are to the specific SES variable and the group selection criteria used here. Partitioning students according to alternative SES measures, such as the Australian Bureau of Statistics (ABS) neighbourhood based SES measures or measures based on average taxable incomes within postcodes, did not change the qualitative features of the results.

The ABS social status index provides an indication of the average social backgrounds of the neighbourhoods in which each student lived in the first wave of their respective survey cohorts. It shows that two thirds of the *unlikely* constrained group are in the top quartile of the ABS social status index, with one third in the top decile and 3.9% in the bottom quartile. Conversely, 57.3% of the *likely* constrained group are in the bottom quartile of the ABS social status index with only 2.2% of these students in the top quartile. The rate of receipt of AUSTUDY is also in line with expectations, with 55.3% (13.7%) of the *likely* (*unlikely*) constrained group in receipt of government support. These indicators demonstrate that members of the *unlikely* constrained group come from more privileged backgrounds than the *potentially* constrained group, who in turn have substantially higher social backgrounds than members of the *likely* constrained group.

Descriptive statistics for the set of variables used in the regression equations for the three groups are shown in Table A.1 of Appendix A. These confirm the advantaged nature of the social background of the *unlikely* constrained group relative to the other groups. Their parents are much more highly educated, and the individuals themselves are more likely to live in urban areas and less likely to be Indigenous.

Preliminary analysis of the role of credit constraints is presented in Figure 2 where nonparametric estimates of the probability of university attendance conditional on ENTER score are presented for each credit constraint group. The key result is that the curves for each group are virtually on top of each other. The implication is, given high school achievement as reflected in ENTER scores, students who are *likely* to face credit constraints have the same probability of attending university as students who are *unlikely* to face credit constraints. The analysis below introduces a full set of covariates and addresses unobserved heterogeneity in modelling university attendance. However, the results are consistent with those in Figure 2 and we do not find evidence that *likely* constrained students are less likely to attend university. We now proceed to the more detailed empirical analysis.

5 Results

The results of estimating equation (2), where no measures have been taken to deal with the possible effects of unobservables denoted by c_i, are presented in Table 2, with parameter estimates and marginal effects in the first and second column respectively. These are our baseline estimates and we find strong positive effects of parental education, student NESB immigrant status and a negative male gender effect, all consistent with expectations based on existing literature. We also find a small negative LSAY 98 cohort effect, consistent with the stronger labour market conditions experienced by that cohort after 2000, and unsurprising state effects which are excluded to save space.

In addition to these standard controls, linear, quadratic and cubic ENTER score terms are included. The overall effect of ENTER score is positive. The stronger the high school achievement of a student, the more likely they are to attend university. Given the nonlinear effect of ENTER score, we present the overall marginal effects evaluated at different ENTER scores in the first column of Table 3. This shows the marginal effect of ENTER score is significant at the 1% level across the full range of ENTER scores and is highest at the 60-80 range, declining at higher and lower ENTER scores. The implication is that in these higher and lower ranges, changes in ENTER score are less likely to change offers of places, or the decision to attend university.

Our research question is, are students who are likely to be credit constrained less likely to attend university? The overall marginal effects of credit constraint group membership in the second column of Table 2 imply no statistically significant average marginal effects. Since the estimated model includes an interaction term between ENTER score and group membership, the marginal effect of group membership at different ENTER scores is evaluated and presented in the second and third columns of Table 3 and Figure 3.[14] These group membership marginal effects correspond to parameter estimates in Table 2. Relative to the group that is *most likely* to be credit constrained, the marginal effects of group membership

[14]Tests of the inclusion of higher order interaction terms between ENTER score and group membership did not reject the null hypothesis that those effects were zero.

are only significant at low ENTER scores, 40-60. The *unlikely* to be constrained group is 9-11% less likely to attend university, significant at the 1% level, while the *potentially* constrained group is 5-6% less likely to attend university. The average marginal effects in Table 3 correspond to the group membership marginal effects in Table 2 and are not significantly different from zero.

Figure 3 plots predicted values for the three groups at different values of the ENTER score using average group characteristics. Consistent with Table 3, this figure shows that the difference in attendance probabilities is greatest at low ENTER scores, but in favour of the *likely* to be constrained group with the gap narrowing as high school achievement increases. This gap is particularly evident for ENTER scores up to around 60. Above that point, there are no differences between groups until the very top of the ENTER distribution, though there are very few observations from the *likely* constrained group with scores of 95 or higher. The question remains whether potentially unobserved differences in the strength of attitudes towards university education held between the groups drives our results among those who obtained an ENTER score. The next subsection presents estimates based on the approaches outlined in Section 3.2 to deal with unobserved heterogeneity in attitudes towards university study.

5.1 Sensitivity analysis

The above results imply, conditional on a range of individual characteristics, students who are expected to face credit constraints are no less likely to attend university than students in the other two credit constraint groups. In this subsection we report on the outcomes of two extensions to our main results. First, we present results of our estimated unobservables approach where we add estimates of otherwise unobserved heterogeneity to our main regression equation. These estimates of unobservables are based on student's university participation plans reported in Year 9. Second, we present estimates where we take account of unobserved heterogeneity via Copula estimation that allows for the same unobserved factors to influence both Year 9 plans and the actual university participation decision. Our key finding that likely constrained students are no less likely to attend university is robust to these alternative approaches that account for unobservables.

Based on equation (3) from Section 3.2, we generate estimates of unobservables, $\hat{c}_{ih}$, using four different group norms. Each of these group norms provide a different estimate of $\tilde{u}_{h1}^{plan}$ and include norms for: (i) all individuals of the same gender in each cohort (LSAY95 and LSAY98), providing four different values for $\tilde{u}_{h1}^{plan}$; one for each gender from each cohort. Recall that u_{h1}^{plan} is dichotomous, so $\hat{c}_{ih}$ takes up to eight values. This comprises $1 - \tilde{u}_{h1}^{plan}$ and -$\tilde{u}_{h1}^{plan}$ for each gender in each of the two cohorts; (ii) all individuals of the same gender in each cohort, for whom none, one or both of their parents attended university. In this case, $\hat{c}_{ih}$ takes a total of 24 values, twelve for each gender; (iii) all individuals with the same level of Year 9 achievement, p_{i1}. In this case, $\tilde{u}_{h1}^{plan}$ is estimated non-parametrically via the *lowess* smoothing function in STATA; (iv) all students of the same gender in each cohort at the same school. These last two group 'norms' provide estimates of unobservables that incorporate more of the individual information contained in the the LSAY data.

We also implement the regression based norm described in Section 3.2 by estimating equation (4). This approach takes into account a full range of characteristics included in the estimation of equation (2), but observed in wave one rather than when deciding to attend university, and excluding credit constraint group identifiers. Further, the ENTER variable is replaced by Year 9 achievement,

All of these alternative estimates of unobservables appear to capture important aspects of unobserved heterogeneity. We illustrate this in Table 4, where in the first column we present the average value of $\hat{c}_{ih}$, for each of the three credit constraint groups in five panels, one for each of the five norms. The averages in the first column are calculated using the the full sample of Year 9 students; this includes students that had not completed high school and those that had not obtained an ENTER score after completing high school. The negative values for *likely* constrained students for each group norm indicate that on average, *likely* constrained students tend not to plan to attend university relative to other students from the same group norm. Unsurprisingly, the more observable characteristics we take into account, such as achievement or social background through the regression based norm, the smaller the differences in $\hat{c}_{ih}$ between credit constraint groups.

The second column of Table 4 presents averages of the same departures from expectations, again measured in ninth grade, but where the averages are calculated only for students who completed high school and acquired an ENTER score. All of the average values increase. Compared to the full sample, these estimates suggest that in each group there is indeed selection into obtaining an ENTER score based on ninth grade plans to attend university. The increase is greatest for the *likely* constrained group for all measures, but is reflected most markedly in the estimated unobservables based on the school average 'norm' and the regression based measure. For these estimates, among those who obtain an ENTER score, the *likely* constrained students exhibited stronger plans to attend university than the potentially constrained students who in turn had slightly stronger plans to attend university than the *unlikely* constrained students. The interpretation is that of the students who obtain an ENTER score and qualify for university entrance, the *likely* constrained students are on average more motivated to undertake university study. In general, the table indicates that our measures of unobservables are consistent with expectations about the direction of potential biases in the baseline results.

Each of these departures from expectations, $\hat{c}_{ih}$, are included in the estimation of equation (2) to assess the potential impact of permanent unobservables on our baseline estimates. The results from estimating this equation with each alternative measure of unobservable heterogeneity included in the specification appear in Appendix Table B.1.[15] The unobservables measures are always significant at the 1% level and clearly very important in explaining university participation.[16]

[15]The standard errors reported in Appendix Table B.1 are taken from bootstrapped estimates of equation (2), based on 50 replications.

[16]Several alternative transformations of the regression based estimates of unobservables were used to correct for any potential biases in the estimation of equation (2) driven by unobservables. These include the rank of the regression based estimate of unobservables, linear, quadratic and cubic splines with parameters for positive values of $\hat{c}_{ih}$ allowed to differ from those associated with negative values, use of the Inverse Mills Ratio and allowing its effect to also be different for positive and negative values. These alternative measures did not produce qualitatively different estimates to those presented.

The relevant marginal effects on participation of being in the *potentially* and *unlikely* constrained groups across the distribution of ENTER scores appear in Table 5. These should be compared with the estimated marginal effects reported in the second and third columns of Table 3. It can be seen that the marginal effects in Table 3 are similar to those across the alternative measures of heterogeneity in Table 5. At low ENTER scores, the marginal effect of being in the *potentially* and *unlikely* constrained groups compared with the *likely* constrained group is negative. That differential disappears at higher ENTER scores and is reversed for the *unlikely* constrained group at low and mixed levels of significance. Results based on less sophisticated estimates of unobservables, using only gender or parental education, point to slightly larger negative differences for the *potentially* and *unlikely* constrained groups compared with the *likely* constrained group. However, measures that account for achievement and SES result in slightly smaller differences between the groups than the estimates that ignore the role of unobservable heterogeneity, though none of the differences are material. Across the entire distribution of ENTER scores, there were no major differences in marginal effects on university attendance between the three groups for any of the norms used to estimate unobservables. Most critical is that in all cases, the *potentially* and *unlikely* constrained students are not more likely to attend university than *likely* constrained students, implying that our baseline results are robust to the inclusion of estimated unobservables.

The second approach we use to deal with the possible effects of unobserved heterogeneity involves jointly estimating the student's plans to attend university and the actual university attendance decision via Copula estimation. Jointly estimating equations (5) and (6) using a Copula specification allows us to account for any potential dependence in the errors of these two equations, as described in Section 3.2. Results for the university attendance equation, (5), where five alternative Copula functions (Gaussian, Frank, Clayton, Gumbel and Joe) were used appear in Appendix Table B.2. For comparison, we reproduce parameter estimates where we have not corrected for any potential effects of unobservable heterogeneity from column 1 of Table 2, and where we include estimated unobservables, $\hat{c}_{ih}$, using the regression based norm, from column 5 of Table B.1. The key conclusion from these parameter estimates is that the various Copula specifications used to correct for potential biases due to unobservable heterogeneity do not markedly change the important parameter estimates for university attendance.

The marginal effects of group membership, calculated across the ENTER distribution, are presented in Table 6, with marginal effects reproduced for the baseline specification from column's 2 and 3 of Table 3, and for the regression based estimate of unobservables, taken from the last column of 5. These results show the Copula based estimated marginal effects across ENTER scores are very similar to those estimated for the baseline specification. The Gaussian Copula estimates also provide a measure of the correlation between the disturbance terms in equations (5) and (6). The estimated correlation between these disturbance terms is $\rho = 0.23$ with standard error $s.e._{\rho} = 0.02$. This implies a strong link between these error terms, as suggested by our earlier approach of including the estimated unobservables in the university attendance equation. Thus, while unobserved heterogeneity is important in explaining university participation, it does not induce any noticeable biases in the estimated credit constraint group effects. As a consequence, our qualitative conclusions that potentially credit constrained students are no less likely to attend university than others in their cohort are unchanged.

5.2 What does matter for university attendance?

We now present the university participation decision in a slightly different way to highlight why we find the *likely* constrained group of students are at least as likely to attend university as the other credit constraint groups.

The participation decision is decomposed into four broad components. Three of these are based on the results of the estimation of equation (4). The parameters from equation (4) allow us to estimate the contributions of individual characteristics (X_{i1}), ninth grade achievement levels (p_{i1}) and estimated unobservables ($\hat{c}_i$) to student plans. We rank individuals in terms of these estimated contributions. That is, we compute:

$$\widetilde{T}_{i1} = rank(X'_{i1}\hat{\delta}), \tag{7}$$

$$\widetilde{p}_{i1} = rank(p_{i1}\hat{\lambda}) = rank(p_{i1}), \tag{8}$$

$$\widetilde{c}_i = rank(\hat{c}_i) = rank(\hat{\epsilon}_{i1}), \tag{9}$$

where $\widetilde{T}_{i1}$ is the rank of a composite index of the effects of the X_{i1} variables used to explain ninth grade plans, $\widetilde{p}_{i1}$ is the rank of the estimated effect of ninth grade academic achievement on ninth grade university attendance plans and $\widetilde{c}_i$ is the rank of the regression based norm estimates of unobservable heterogeneity. Ranking individuals on these characteristics provides a normalisation that can be used to compare the relative importance of the four factors we consider.

The fourth component is based on the rank of individual ENTER scores. We impute a value for those individuals who do not obtain an ENTER score.[17] In the imputation process we assume the estimated effects of all factors determining the ENTER score are the same for those students who did and did not obtain an ENTER score. Thus, those individuals who did not obtain an ENTER score may only receive a lower imputed score than those who did obtain an ENTER score because of poorer characteristics. Nevertheless, the average imputed ENTER score of those who did not obtain an ENTER score was 46 compared with 65 for those who did. Further, 64 percent of those who did not obtain an ENTER score were imputed to have an ENTER score of below 50, compared with 34 percent of those who did obtain an ENTER score.

We estimate the effects of these four rank variables on the probability an individual attends university using a non-parametric smoother, conditional on the other rank variables. The results are shown in Figure 4. The line for any of the individual ranks takes account of the impact of the other ranks, so it is akin to a marginal effect we would estimate from a regression equation. Panel A shows the estimated relationships when only the ranks of

[17]The imputation of missing ENTER score values was based on OLS using the following variables: group membership, achievement, parental occupational SES, gender, completed parental education, student born overseas, Indigenous, metropolitan area, state indicators, individual self-confidence and cohort.

ninth grade observables, achievement and unobservables are included, while Panel B adds the rank of individual ENTER scores.

A number of features of the two figures and their comparison are noteworthy. First, Year 9 observables and achievement predominantly affect university attendance via their impact on ENTER scores. In Panel B, the effects of moving from the bottom to the top rank of those variables is demonstrated to be much smaller once ENTER ranks are taken into account. Second, the unobservables also have an impact on attending university, though this effect is attenuated by the inclusion of the ENTER rank. Third, the ENTER rank strongly dominates the explanation of who goes to university. Some people with low ENTER scores do attend university, but marginal increases in ENTER ranks in the bottom half of the distribution do not affect the probability of attendance. However, moving from the half way point of the distribution to the top increases the probability by about 80 percentage points. Beyond the ENTER ranks of individuals, the other factors we include have very little role in explaining differences in university attendance.

The implications of these figures for our central research question is that credit constraint group membership must be less important than ENTER scores in the determination of university attendance. The observables curve in Panel B of Figure 4 includes an SES variable and reflects the information in our credit constraint group variables. Comparing the effect of observables between the top and bottom 10 percent ranked individuals, which is what our earlier analysis effectively does, the difference in university attendance is very small. Given this decomposition, it is not surprising that there is very little difference between the likely and unlikely constrained groups in university attendance once we condition on ENTER scores, as our main results suggest.

6 Conclusions

We have studied the effects of credit constraints on the university attendance decisions of students. Conditional on high school achievement, the results show that students likely to face credit constraints do not have a lower probability of university attendance than those unlikely to face constraints. Instead, we find that among those lower in the high school achievement distribution, likely credit constrained students have a slightly greater probability of attending university than students unlikely to face constraints.

An innovative approach to address any possible bias that might arise from the omission of unobservables was also devised. The approach exploited responses collected in ninth grade to questions about study plans after high school and led to the estimation of unobservable characteristics. These estimated unobservables reflected how unexpected student responses to questions about post-school plans were and likely incorporate typically unobservable characteristics such as student motivation and effort. On inclusion of these estimated unobservables in the university attendance model, we found them to be very important in explaining university attendance decisions of students. However, our key conclusions were unaffected by their inclusion. Conditional on high school achievement, coming from a likely credit constrained background does not appear to be a factor that reduces the probability of university attendance.

Our estimated unobservables approach may have application in other circumstances. The approach allows us to test for the effects of unobservables by comparing individuals against community or group norms in order to gain an estimate of unobservable heterogeneity. In our application, we were interested in the university attendance decision and extracted estimated unobservables from a ninth grade survey response to post school study plans, compared to what is expected from each individual given their characteristics. However, the approach may be applicable to other settings where information on plans or attitudes to a decision or choice are collected along with data on the actual choice, where the same unobserved factors contribute to both plans and the choice.

As a further robustness check on the role of unobservables in our results, we jointly estimated the university attendance decision and ninth grade post school study plans. This model was estimated using a Copula specification for the dependence between the disturbance terms of the respective equations. The Copula approach yielded similar results to our estimated unobservables technique. Once again, coming from a likely credit constrained background does not reduce the probability of university attendance.

Given these findings, we decompose the probability of university attendance in order to identify the most important factor affecting differences in university attendance. The dominant factor is found to be high school achievement as measured by ENTER, the key mechanism by which university places are rationed.

The Australian higher education sector has evolved since the students studied here made decisions about university attendance. One of the key changes has been the shift to a "demand driven system". Before this change, the number of university places supported by the HECS was determined by the Commonwealth on a university by university basis. The demand driven system allows each Australian university to offer as many HECS supported places as they see fit. Since this change in 2012, the proportion of undergraduate students from low SES backgrounds has grown from 16.21% in 2011 to 16.84% in 2014.[18]

These statistics suggest the demand driven system has not fundamentally changed the socioeconomic composition of Australian higher education. Consequently, our main result on the importance of high school achievement as the key determinant of participation remains. Policies to change the SES gradient in university participation are best focused on factors that affect qualifying for university admission. These factors may appear early in childhood, certainly before the age of 15 when we first see the subjects in the LSAY surveys. The role of credit constraints in a post "demand driven system" is a question for future research. The new LSAY 2015 cohort will, in years to come, provide the first opportunity to investigate if and how the behaviour of potentially credit constrained students has evolved in response to this change in higher education policy in Australia.

Compliance with Ethical Standards:

Funding: Part of this research was funded by the Australian Research Council (DP0662909).

Conflict of Interest: The authors declare that they have no conflict of interest.

[18]See 2014 Appendix 5 - Equity performance data on the Commonwealth Department of Education and Training, Selected Higher Education Statistics - 2014 Student data website at http://docs.education.gov.au/node/38151

References

[1] Arcidiacono, P., Hotz, J. and S. Kang, (2012), "Modeling College Major Choices using Elicited Measures of Expectations and Counterfactuals", *Journal of Econometrics* 166(1), 3-16.

[2] Australian Bureau of Statistics, (1998), *Information Paper: 1996 Census of Population and Housing Socio-Economic Indexes for Areas*, Cat.No, 2039.0, Canberra.

[3] Belley, P. and L. Lochner, (2007), "The Changing Role of Family Income and Ability in Determining Educational Achievement", *Journal of Human Capital* 1(1), 37-89.

[4] Blossfeld, P.N., Blossfeld, G.J. and H-P Blossfeld, (2016), Changes in Educational Inequality in Cross-National Perspective, in *Handbook of the Life Course: Volume II* (eds M.J. Shanahan, J.T. Mortimer and M. Kirkpatrick Johnson), Cham: Springer International Publishing.

[5] Blossfeld, H-P and Y. Shavit, (1993), *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, Boulder, San Francisco, Oxford: Westview Press.

[6] Brown, M., Scholz, J.K. and A. Seshadri, (2012), "A New Test of Borrowing Constraints for Education", *Review of Economic Studies* 79(2), 511-538.

[7] Cameron, S.V. and C. Taber, (2004), "Estimation of Educational Borrowing Constraints Using Returns to Schooling", *Journal of Political Economy* 112(1), 132-182.

[8] Cardak, B.A. and C. Ryan, (2014), "Evidence on credit constraints, university attendance and income contingent loans ", Melbourne Institute Working Paper No. 24/14.

[9] Cardak, B.A. and J. Vecci, (2016), "Graduates, Dropouts and Slow Finishers: The Effects of Credit Constraints on University Outcomes", *Oxford Bulletin of Economics and Statistics* 78(3), 323-346.

[10] Carneiro, P. and J.J. Heckman, (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", *The Economic Journal* 112(482), 705-734.

[11] Chapman, B., (1997), "Conceptual Issues and the Australian Experience with Income Contingent Charges for Higher Education", *The Economic Journal* 107(442), 738-751.

[12] Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and A. Vignoles, (2014), "Widening participation in higher education: analysis using linked administrative data", *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176(2), 431–457.

[13] Coelli, M.B., (2011), "Parental job loss and the education enrolment of youth", *Labour Economics* 18, 25–35.

[14] Cunha, F., Heckman, J., Lochner, L. and D. Masterov, (2006), Interpreting the evidence on life cycle skill formation, in *Handbook of the Economics of Education* (eds E. Hanushek and F. Welch), Amsterdam: North-Holland.

[15] Dearden, L., L. McGranahan and B. Sianesi, (2004), "The Role of Credit Constraints in Educational Choices: Evidence from the NCDS and BSC70", Centre for the Economics of Education, London School of Economics, Working Paper CEEDP0048, accessed at http://cee.lse.ac.uk/ceedps/ceedp48.pdf

[16] Denny, K., (2014), "The effect of abolishing university tuition costs: Evidence from Ireland", *Labour Economics* 26, 26–33.

[17] Dominitz, J. and C. Manski, (1996), "Eliciting student expectations of the returns to schooling", *Journal of Human Resources* 31(1), 1-26.

[18] Hernández-Alava, M. and S. Pudney, (2016), "bicop: A command for fitting bivariate ordinal regressions with residual dependence characterized by a copula function and normal mixture marginals", *Stata Journal* 16(1), 159-184.

[19] Keane, M and K. Wolpin, (2001), "The Effects of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economics Review* 42(4), 1051-1103.

[20] OECD, (2013), Education at a Glance 2013: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2013-en

[21] Rothman, S., (2002), *Achievement in Literacy and Numeracy by Australian 14 year-olds, 1975 - 1998*, Longitudinal Surveys of Australian Youth Research Report No. 29, Australian Council for Educational Research, Melbourne.

[22] Ryan, C., (2013), *Student income support and education and training participation in Australia*, Adelaide: NCVER.

[23] Stinebrickner, R. and T. Stinebrickner, (2008), "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using and New Panel Study", *American Economic Review* 98(5), 2163-2184.

[24] Trivedi, P.K. and D.M. Zimmer, (2007), "Copula modelling: An introduction for practitioners", *Foundations and Trends in Econometrics* 1(1), 1–111.

[25] Wooldridge, J.M., (2010), *Econometric Analysis of Cross Section and Panel Data* , 2nd Edition, Cambridge, Massachusetts: MIT Press.

[26] Zafar , B., (2011), 'Can Subjective Expectations Data Be Used In Choice Models? Evidence On Cognitive Biases', *Journal of Applied Econometrics* 26(3), 520-544.

Table 1: Descriptive statistics of key features by the three credit constraint groups.

	Groups			
	Unlikely Constrained	Potentially Constrained	Likely Constrained	Total
Panel A: Summary statistics for pooled LSAY 95 and LSAY 98 sample.				
Observations				
Unweighted	1,614	13,086	1,307	16,007
Weighted	1,323	12,861	1,770	15,954
Weighted observations with ENTER score	1,138	7,462	700	9,300
Proportion with ENTER score (%)	86.4	58.3	39.8	58.6
Proportion with ENTER score if Year 12 (%)	92.7	75.2	61.9	76.7
Proportion female (%)	54.0	48.9	48.7	49.3
University participation rate (%)	67.9	35.3	19.9	36.3
Average ENTER score (if had one)	83.1	72.0	62.4	72.6
Year 12 participation rate (%)	95.7	82.7	73.3	82.8
Proportion at Government school	0.0	70.8	93.4	67.6
Proportion at Catholic school	40.9	19.8	6.6	20.1
Proportion at Independent school	59.1	9.4	0.0	12.3
Panel B: Social background and government support for each group.				
ABS education/occupation SES index				
Proportion in top quartile(%)	63.0	20.8	2.2	
Proportion in top decile(%)	34.6	6.1	0.3	
Proportion in bottom quartile(%)	3.9	25.0	57.3	
Proportion who received AUSTUDY/Youth Allowance at university (%)	13.7	35.9	55.3	

Table 2: University participation parameter estimates and marginal effects based on equation (2) without any adjustment for unobservables, c_i.

	Parameters	Marginal Effects
Unlikely constrained	-1.243^{***}	-0.02
	(0.386)	(0.02)
Potentially constrained	-0.472^{*}	-0.02
	(0.257)	(0.02)
Unlikely constrained × ENTER	0.016^{***}	
	(0.005)	
Potentially constrained × ENTER	0.005	
	(0.004)	
ENTER	-0.087^{**}	0.01^{***}
	(0.040)	(0.00)
$(\text{ENTER})^2$	0.002^{***}	
	(0.001)	
$(\text{ENTER})^3$	-0.000^{***}	
	(0.000)	
Male	-0.083^{***}	-0.02^{***}
	(0.031)	(0.01)
Student overseas born		
English speaking country	-0.011	0.00
	(0.094)	(0.02)
Non-English speaking country	0.419^{***}	0.11^{***}
	(0.060)	(0.02)
Father has degree	0.098^{***}	0.03^{***}
	(0.037)	(0.01)
Mother has degree	0.092^{**}	0.02^{**}
	(0.040)	(0.01)
Self-confidence	0.005	0.00
	(0.005)	(0.00)
Indigenous	0.233	0.06
	(0.149)	(0.04)
Metropolitan	-0.036	-0.01
	(0.037)	(0.01)
Y98 cohort	-0.122^{***}	-0.03^{***}
	(0.034)	(0.01)
Constant	-0.700	
	(0.905)	
State indicators	**Yes**	
Observations	9,898	

Table 3: Marginal effects of ENTER scores and credit constraint group identifiers on university participation, calculated at different points in the ENTER distribution. Marginal effects for credit constraint group identifiers are relative to the likely constrained group and vary because of the interaction with ENTER scores in equation (2).

	Marginal Effects		
ENTER	ENTER	Unlikely Constrained	Potentially Constrained
40	0.007***	−0.089***	−0.053**
50	0.014***	−0.112***	−0.060**
60	0.020***	−0.105***	−0.056**
70	0.021***	−0.052	−0.036
80	0.014***	0.005	−0.012
90	0.007***	0.031	0.002
95	0.004***	0.035*	0.006
99	0.003***	0.037**	0.008
Average Marginal Effect	0.010***	−0.024	−0.023

Table 4: Estimates of unobservables, $\hat{c}_{ih}$, averaged by credit constraint group. Each panel presents estimates of unobservables based on Year 9 plans to attend university and different norms or reference groups. Estimates in the first column are averages calculated using all students reporting plans in Year 9, the second column are averages for all students with an ENTER score and the third column presents the difference between first two measures.

	All Year 9's	Only Year 9's with ENTER	Difference
Panel A: Gender			
Likely constrained	−0.16	0.14	0.30
Potentially constrained	0.00	0.20	0.20
Unlikely constrained	0.31	0.36	0.06
Total	0.01	0.21	0.21
Panel B: Parents' Education			
Likely constrained	−0.10	0.19	0.30
Potentially constrained	0.00	0.18	0.17
Unlikely constrained	0.16	0.20	0.04
Total	0.01	0.18	0.17
Panel C: Achievement			
Likely constrained	−0.07	0.15	0.22
Potentially constrained	−0.01	0.12	0.13
Unlikely constrained	0.17	0.21	0.03
Total	0.00	0.13	0.14
Panel D: School Average			
Likely constrained	−0.04	0.21	0.25
Potentially constrained	0.01	0.16	0.16
Unlikely constrained	0.06	0.10	0.04
Total	0.01	0.16	0.15
Panel E: Regression			
Likely constrained	−0.02	0.18	0.20
Potentially constrained	−0.01	0.10	0.11
Unlikely constrained	0.04	0.07	0.03
Total	0.00	0.10	0.11

Table 5: Marginal effects of credit constraint group identifiers on university participation after inclusion of estimated unobservables, $\hat{c}_{ih}$, calculated at different points in the ENTER distribution. Marginal effects are relative to the likely constrained group and vary because of the interaction with ENTER scores in equation (2).

	Alternative Heterogeneity Measures				
ENTER	Gender	Parents' Education	Achievement	School Average	Regression
Unlikely Constrained					
40	−0.096***	−0.095***	−0.087***	−0.081***	−0.082***
50	−0.114***	−0.113***	−0.105***	−0.092***	−0.096***
60	−0.106***	−0.105***	−0.098***	−0.073**	−0.084**
70	−0.056*	−0.055*	−0.049	−0.020	−0.035
80	−0.002	−0.001	0.003	0.027	0.013
90	0.026	0.026	0.028	0.043*	0.033
95	0.032	0.032	0.032	0.044**	0.036*
99	0.036*	0.036*	0.034*	0.045**	0.038*
Average Marginal Effect	−0.028	−0.027	−0.023	−0.004	−0.015
Potentially Constrained					
40	−0.056**	−0.056**	−0.049*	−0.048*	−0.046*
50	−0.062**	−0.061**	−0.055**	−0.051*	−0.050*
60	−0.058**	−0.057**	−0.052**	−0.043	−0.045*
70	−0.039	−0.038	−0.034	−0.023	−0.027
80	−0.016	−0.016	−0.013	−0.002	−0.008
90	−0.002	−0.002	0.000	0.009	0.004
95	0.003	0.003	0.004	0.012	0.007
99	0.005	0.006	0.006	0.014	0.009
Average Marginal Effect	−0.026	−0.025	−0.022	−0.014	−0.017

Table 6: Marginal effects of credit constraint group identifiers on university participation, based on Copula estimation of equations (5) and (6) and calculated at different points in the ENTER distribution. Marginal effects are relative to the likely constrained group and vary because of the interaction with ENTER scores in equation (5). Marginal effects for baseline specification and the case of regression based estimates of unobservables, $\hat{c}_{ih}$, are included for comparison.

		$\hat{c}_{ih}$ included	Copula Specification				
ENTER	Baseline	Regression Norm	Gaussian	Frank	Clayton	Gumbel	Joe
Unlikely Constrained							
40	−0.089***	−0.082***	−0.105***	−0.105***	−0.103***	−0.107***	−0.110***
50	−0.112***	−0.096***	−0.118***	−0.117***	−0.119***	−0.119***	−0.120***
60	−0.105***	−0.084**	−0.102***	−0.100***	−0.105***	−0.101***	−0.101***
70	−0.052	−0.035	−0.047	−0.045	−0.052	−0.046	−0.044
80	0.005	0.013	0.005	0.007	0.001	0.007	0.010
90	0.031	0.033	0.030	0.031	0.027	0.031	0.033
95	0.035*	0.036*	0.035	0.036*	0.033	0.036*	0.038*
99	0.037**	0.038*	0.038*	0.039*	0.036*	0.039*	0.041
Average Marginal Effect	−0.024	−0.015	−0.023	−0.021	−0.026	−0.022	−0.020
Potentially Constrained							
40	−0.053**	−0.046*	−0.052**	−0.052**	−0.052**	−0.052**	−0.054**
50	−0.060**	−0.050*	−0.057**	−0.057**	−0.059**	−0.057**	−0.057**
60	−0.056**	−0.045*	−0.054**	−0.053**	−0.056**	−0.053**	−0.052**
70	−0.036	−0.027	−0.035	−0.034	−0.037*	−0.034	−0.032
80	−0.012	−0.008	−0.013	−0.012	−0.015	−0.012	−0.010
90	0.002	0.004	0.001	0.002	0.000	0.002	0.003
95	0.006	0.007	0.005	0.006	0.004	0.006	0.007
99	0.008	0.009	0.008	0.009	0.007	0.009	0.010
Average Marginal Effect	−0.023	−0.017	−0.021	−0.020	−0.022	−0.020	−0.019

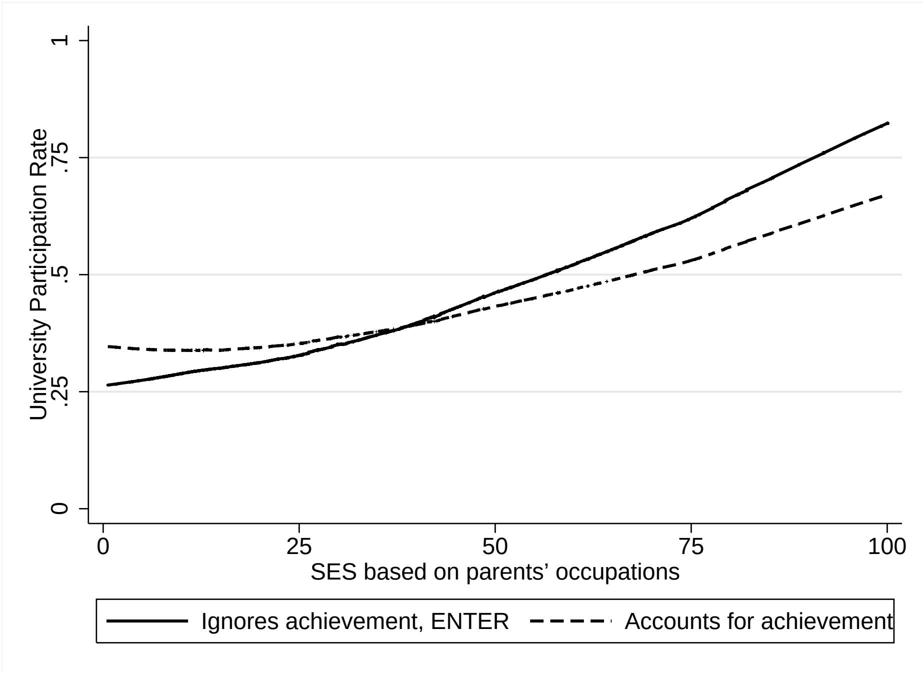

Figure 1: Probability of university participation in Australia by SES background before and after adjustment for studnet ability based on ninth grade school achievement.

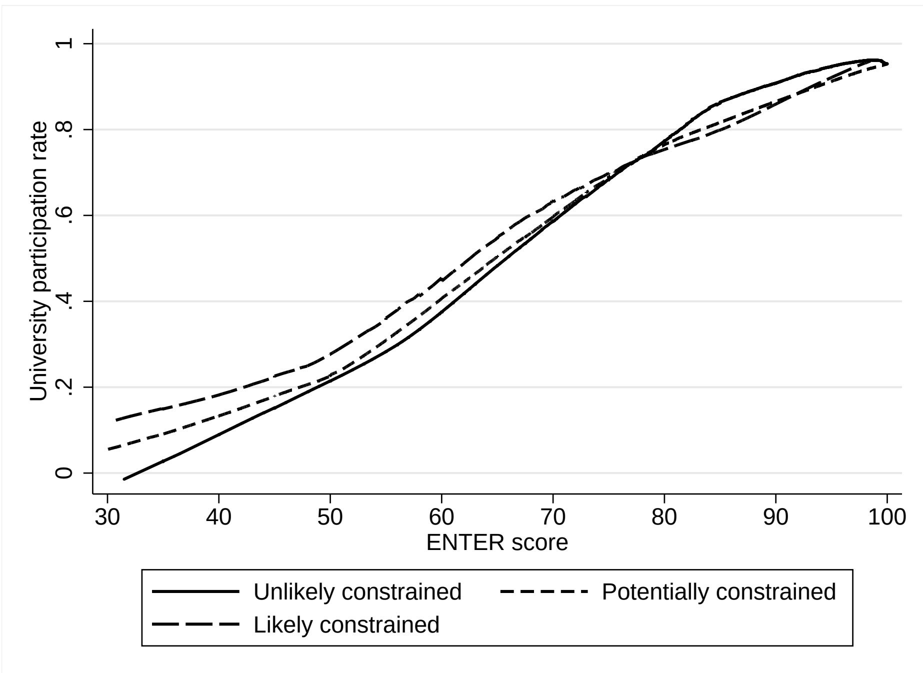

Figure 2: Non-parametric estimates of the probability of university participation by ENTER score and credit constraint group membership.

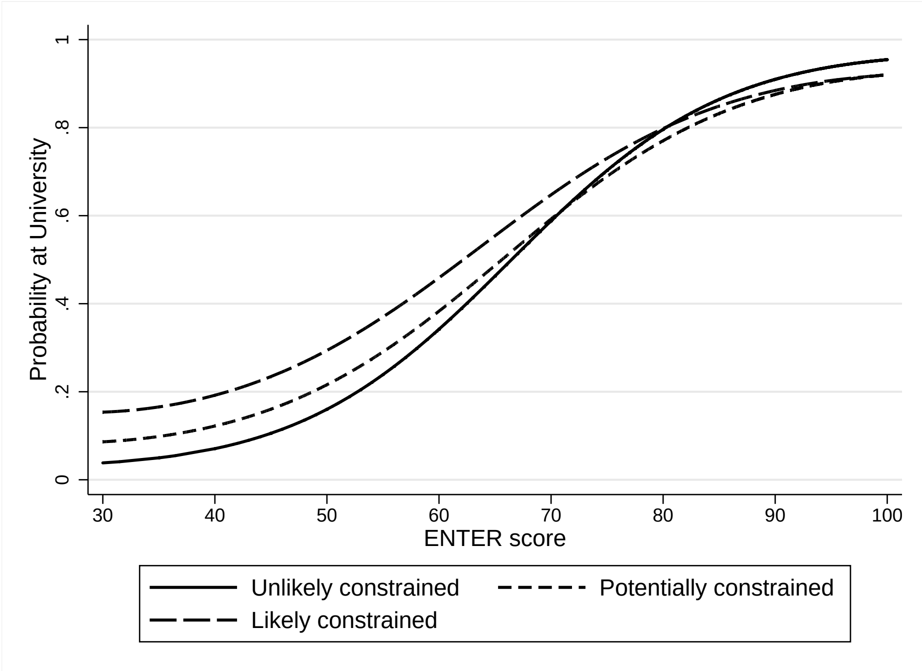

Figure 3: Predicted probabilities of university attendance for different credit constraint groups by ENTER score, based on parameter estimates in Table 2.

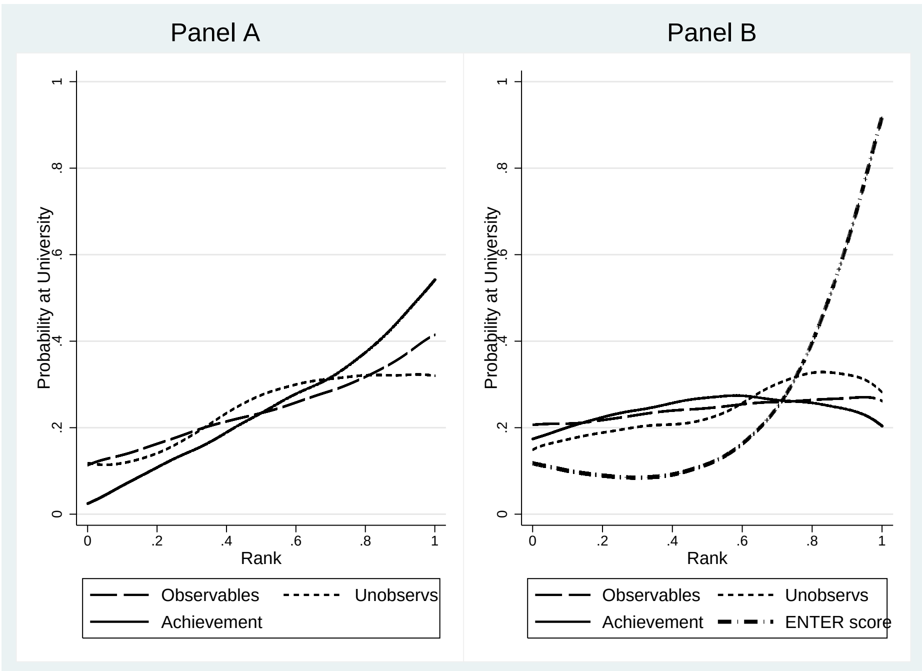

Figure 4: Estimates of the impact of ninth grade academic achievement, other observables and estimated unobservables on university participation in Panel A, and with ENTER score accounted for in Panel B.

Appendix A Supplementary Summary Statistics

Table A.1: Descriptive statistics by credit constraint group.

	Groups			
	Unlikely Constrained	Potentially Constrained	Likely Constrained	Total Mean (Std.Dev.)
University Participation	0.679	0.353	0.199	0.363 (0.481)
Proportion with an ENTER score	0.864	0.583	0.398	0.586 (0.493)
ENTER score (if had one)	83.10	72.00	62.40	72.60 (17.80)
Achievement	55.80	50.40	46.00	50.40 (8.30)
Male	0.540	0.489	0.487	0.493 (0.500)
Student born overseas				
English speaking country	0.048	0.027	0.012	0.027 (0.161)
Non-English speaking country	0.060	0.065	0.105	0.069 (0.254)
Father has degree	0.628	0.178	0.045	0.201 (0.401)
Mother has degree	0.451	0.176	0.050	0.185 (0.388)
Self-confidence	49.90	50.00	50.70	50.10 (3.19)
Indigenous	0.004	0.023	0.042	0.023 (0.151)
Metropolitan	0.813	0.504	0.598	0.540 (0.498)
New South Wales	0.348	0.331	0.308	0.330 (0.470)
Victoria	0.269	0.235	0.251	0.240 (0.427)
Queensland	0.128	0.195	0.215	0.192 (0.394)
South Australia	0.086	0.076	0.067	0.076 (0.264)
Western Australia	0.137	0.105	0.095	0.106 (0.308)
Tasmania	0.012	0.027	0.052	0.029 (0.167)
Northern Territory	0.000	0.009	0.005	0.008 (0.089)
Australian Capital Territory	0.019	0.021	0.005	0.019 (0.137)
Observations	1,614	13,086	1,307	16,007

Appendix B Supplementary Results

Table B.1: Parameter estimates for university participation model, equation (2), with the inclusion of estimated unobservable heterogeneity, $\hat{c}_{ih}$, based on the range of 'norms' outlined in Section 5.1.

	Gender	Parents Education	Achievement	School Average	Regression
Heterogeneity term, $\hat{c}_{ih}$	0.393***	0.391***	0.387***	0.337***	0.392***
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Unlikely constrained	-1.179***	-1.176***	-1.147***	-1.105***	-1.087***
	(0.382)	(0.382)	(0.384)	(0.384)	(0.383)
Potentially constrained	-0.454*	-0.451*	-0.427	-0.444*	-0.412
	(0.261)	(0.260)	(0.262)	(0.260)	(0.262)
Unlikely constrained	0.015***	0.015***	0.014***	0.015***	0.014***
$\times$ENTER	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Potentially constrained	0.005	0.005	0.005	0.005	0.005
$\times$ENTER	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ENTER	-0.097**	-0.097**	-0.093**	-0.091**	-0.095**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
$(\text{ENTER})^2$	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$(\text{ENTER})^3$	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.114***	-0.113***	-0.095***	-0.104***	-0.101***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Background of overseas born students					
English	-0.024	-0.024	-0.025	-0.018	-0.011
	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)
Non-English	0.367***	0.367***	0.344***	0.381***	0.437***
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)

Continued on next page

Table B.1 – *Continued from previous page*

	Gender	Parents Education	Achievement	School Average	Regression
Father has degree	0.070*	0.154***	0.078**	0.078**	0.131***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Mother has degree	0.063	0.144***	0.064	0.075*	0.103**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
Indigenous	0.226	0.226	0.215	0.221	0.202
	(0.152)	(0.152)	(0.151)	(0.152)	(0.152)
Metropolitan	-0.050	-0.050	-0.055	-0.015	-0.026
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Self-confidence	0.001	0.001	0.000	0.003	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
LSAY 98 cohort	-0.116***	-0.121***	-0.115***	-0.115***	-0.122***
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)
Constant	-0.247	-0.269	-0.366	-0.549	-0.610
	(0.912)	(0.911)	(0.914)	(0.907)	(0.913)
State indicators	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Observations 9,917	9,917	9,898	9,917	9,898	

Table B.2: Regression results for university participation equation, (5), jointly estimated with university study plans equation, (6), using a Copula specification for dependence between disturbance terms. Parameter estimates for baseline specification and the case of regression based estimates of unobservable heterogeneity, $\hat{c}_i$, are also presented for comparison.

		$\hat{c}_{ih}$ included	Copula Specification				
	Baseline	Regression Norm	Gaussian	Frank	Clayton	Gumbel	Joe
Likely constrained	-1.243***	-1.087***	-1.116***	-1.123***	-1.109***	-1.131***	-1.152***
	(0.386)	(0.383)	(0.373)	(0.376)	(0.375)	(0.371)	(0.366)
Potentially constrained	-0.472*	-0.412	-0.424	-0.428	-0.431	-0.427	-0.432*
	(0.257)	(0.262)	(0.264)	(0.266)	(0.266)	(0.262)	(0.258)
Likely constrained	0.016***	0.014**	0.014***	0.014***	0.014***	0.014***	0.015***
×ENTER	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Potentially constrained	0.005	0.005	0.005	0.005	0.005	0.005	0.005
×ENTER	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ENTER	-0.087**	-0.095**	-0.092**	-0.089**	-0.088**	-0.096***	-0.100***
	(0.040)	(0.040)	(0.036)	(0.037)	(0.036)	(0.036)	(0.036)
$(\text{ENTER})^2$	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$(\text{ENTER})^3$	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.083***	-0.101***	-0.087***	-0.086***	-0.085***	-0.086***	-0.084***

Continued on next page

Table B.2 – *Continued from previous page*

		$\hat{c}_{ih}$ included	Copula Specification				
	Baseline	Regression Norm	Gaussian	Frank	Clayton	Gumbel	Joe
	(0.031)	(0.032)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Background of overseas born students							
English	-0.011	-0.011	-0.019	-0.014	-0.016	-0.021	-0.025
	(0.094)	(0.094)	(0.090)	(0.090)	(0.089)	(0.090)	(0.090)
Non-English	0.419***	0.437***	0.422***	0.419***	0.424***	0.419***	0.414***
	(0.060)	(0.060)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
Father has degree	0.098***	0.131***	0.111***	0.109***	0.107***	0.110***	0.110***
	(0.037)	(0.037)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Mother has degree	0.092**	0.103***	0.096**	0.095**	0.095**	0.095**	0.093**
	(0.040)	(0.040)	(0.039)	(0.039)	(0.038)	(0.039)	(0.038)
Indigenous	0.233	0.202	0.225	0.221	0.219	0.230*	0.238*
	(0.149)	(0.152)	(0.137)	(0.138)	(0.137)	(0.136)	(0.135)
Metropolitan	-0.036	-0.026	-0.032	-0.033	-0.028	-0.035	-0.040
	(0.037)	(0.037)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Self-confidence	0.005	0.005	0.005	0.005	0.005	0.005	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
LSAY 98 cohort	-0.122***	-0.122***	-0.122***	-0.123***	-0.121***	-0.123***	-0.123***
	(0.034)	(0.034)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)

Continued on next page

Table B.2 – *Continued from previous page*

		$\hat{c}_{ih}$ included	Copula Specification				
	Baseline	Regression Norm	Gaussian	Frank	Clayton	Gumbel	Joe
Heterogeneity term, $\hat{c}_{ih}$		0.392***					
		(0.034)					
Constant	-0.700	-0.610	-0.533				
	(0.905)	(0.913)	(0.819)				
State indicators	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Observations	9,898	9,898	9,898	9,898	9,898	9,898	9,898