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

DISCUSSION PAPER SERIES

IZA DP No. 13721

First in Their Families at University: Can Non-cognitive Skills Compensate for Social Origin?

Rebecca Edwards
Rachael Gibson
Colm Harmon
Stefanie Schurer

SEPTEMBER 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13721

First in Their Families at University: Can Non-cognitive Skills Compensate for Social Origin?

Rebecca Edwards
University of Sydney

Rachael Gibson
University of Sydney

Colm Harmon
University of Edinburgh and IZA

Stefanie Schurer
University of Sydney and IZA

SEPTEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

First in Their Families at University: Can Non-cognitive Skills Compensate for Social Origin?*

We study the role of non-cognitive skills in academic performance of students who are the first in their family to attend university. We collected survey data on an incoming student cohort from a leading Australian university and linked the survey with students' administrative entry and performance records. First-in-family students have lower grade point averages by about a quarter of a standard deviation than the average student. This performance penalty is larger for young men. The penalty is strongest in the first semester but disappears over time. Some non-cognitive skills (Conscientiousness, Extraversion) predict academic performance almost as strongly as standardised university admissions test scores. High levels of Conscientiousness over-compensate for the performance penalty experienced by first-in-family students, while very low levels exacerbate it. However, adjusting for extreme responses in self-assessed Conscientiousness with anchoring vignettes eliminates the performance advantage of disadvantaged, but highly conscientious students. Overall, our findings accentuate the importance of non-cognitive skills as key indicators of university readiness, and their potential for closing the socioeconomic gap in academic performance.

JEL Classification: A22, J24

Keywords: non-cognitive skills, university performance, socioeconomic gradient in education, first-in-family, linked survey and administrative data, anchoring vignettes

Corresponding author:
Stefanie Schurer
School of Economics
University of Sydney
Sydney, NSW, 2006
Australia
E-mail: stefanie.schurer@sydney.edu.au

* The authors acknowledge financial support from an Australian Research Council Early Career Discovery Program Grant to Schurer (DE140100463) and the Australian Research Council Centre of Excellence for Children and Families over the Life Course (project number CE140100027). Approval to conduct the study was obtained from the Human Research Ethics Committee (Low Risk) of the university at which the study was conducted on 11 November 2014 (Nr. 2014/921).

1. Introduction

Higher education is arguably the single most important facilitator of social and economic mobility (Breen & Muller 2020, Haveman & Smeeding 2006, Blanden, Gregg & Macmillan 2007, Blanden, Gregg & Machin 2005). However, children from disadvantaged homes find it harder to pursue higher education opportunities (Jerrim & Vignoles 2015). In OECD countries, only about 1 in 5 adults whose parents did not complete upper secondary education, complete tertiary education. In contrast, 2 in 3 adults whose parents were university educated, complete tertiary education (OECD 2018, based on PIAAC data). In some countries, socioeconomic gaps in university participation have widened in recent years (Page & Scott-Clayton 2016).

Socioeconomic inequalities in higher education are also observed in Australia, despite being considered a rich OECD country with high levels of social mobility (Leigh 2007, Mendolia & Siminski 2016) and the rapid expansion of the university sector in the past 50 years (Chesters & Watson 2013). Children from disadvantaged backgrounds make up 25 percent of the population, but only 15.7 percent enrol at university (Department of Education and Training 2017). Policy makers and university executives alike are concerned about facilitating access to university for disadvantaged students and understanding their constraints during their studies (Department of Education and Training 2016, Universities Australia 2016).

Why are students from disadvantaged homes so much less likely to attain university education? Discussions around the factors limiting access to university focus usually on the financial and opportunity costs of higher education (Page & Scott-Clayton 2016). We argue that access to university also requires the right mind-set. University education incurs psychic costs (Heckman, Stixrud & Urzua 2006, Kassenboehmer, Leung, and Schurer 2018). Even with high levels of cognitive ability, sitting exams, dealing with failure and constant deadlines is hard. Going to university requires intellectual engagement, a sincere enjoyment of challenge, as well as "willingness to accept critical feedback and to adjust based on such feedback, [and] openness to possible failures from time to time" (Conley 2003).

Thus, non-cognitive abilities are likely to play a critical role in facilitating access to and performance during university study. Previous studies have demonstrated the important role of non-cognitive abilities in the human capital accumulation process (Almlund, Duckworth, Heckman & Kautz 2011, Borghans, Duckworth, Heckman & Weel 2008, Bowles, Gintis & Osborne 2001, Lundberg 2013). Some argue that non-cognitive skills are at least as important as cognitive skills in determining life outcomes (Heckman et al. 2006, Lindqvist & Vestman 2011, Bütikofer & Peri 2020). Strong socioeconomic inequalities have been observed in non-cognitive

skills in both childhood (Attanasio, Blundell, Conti & Mason 2020, Elkins & Schurer 2020, Heckman & Mosso 2014) and adulthood (Gensowski, Goertz & Schurer 2020).

In this study, we examine the role of non-cognitive skills in shaping the socioeconomic gradient in university performance. We conducted a unique survey on around 1,000 incoming students at a leading Australian university located in New South Wales, who started their degree in 2015. We collected information on students' non-cognitive skills, high-school achievement tests, detailed socioeconomic background, and the people who inspired them to pursue a university education. Non-cognitive skills were measured with widely accepted personality instruments, such as the Big-Five traits (Conscientiousness, Openness to Experience and Extraversion), Locus of Control, and Grit (see Almlund et al. 2011 for an overview of these instruments). With student permission, we linked this survey to administrative student records on applications and university grades in the first four semesters of study.

To address potential biases produced by the subjective nature of personality assessment, we designed and collected anchoring vignettes for a subset of personality traits collected in the survey. Vignettes have been used in studies which compare education outcomes based on self-assessed measures across cultures (He, Buchholz & Klieme 2017) and in the context of personality assessment (Mõttus, Allik, Realo, Rossier, Zecca & Ah-Kion et al. 2012, Bolt, Lu & Kim 2014, Primi, Zanon, Santos, De Fruyt & John 2016). The vignettes are designed to trigger a similar subjective reaction from all respondents; thus, any variability in respondent ratings is attributable to response style heterogeneity (Bolt et al. 2014). Under the assumptions of response consistency and vignette equivalence, the vignettes are used to correct for person-specific heterogeneity in the interpretation of response categories (King et al. 2004, King & Wand 2007). We adapted eight vignettes from Mõttus et al. (2012) to describe the personality traits of Conscientiousness and Openness to Experience of fictional characters. Survey respondents were asked to rate the personality of hypothetical individuals described in short vignettes, using the same scale as for their own personality assessment. Our vignette experiment exploits repeated assessments and variation in the vignettes by gender and sociodemographic characteristics. Under the assumption that the response style is fixed over assessment tasks – both the vignettes and the self-assessment – we can clean out any individual-specific response heterogeneity in the personality self-assessments.

We focus on students' achievement during their first two years of university study, from March 2015 to November 2016. Our data allow us to construct two measures of socioeconomic disadvantage. First, a measure of whether the student is the first in their family to attend university. We refer to such students as 'first-generation' or 'first-in-family' interchangeably. This is an

important measure of socioeconomic status as it not only incorporates potential family wealth, but also family attitudes towards learning and familiarity with how the university system works. Second, we constructed a socioeconomic status indicator based on residential location. The latter measure is used by the university administration to flag students as disadvantaged, which entitles students to additional services and scholarships.

We seek to answer three critical questions. First, are first-generation students less ready for university life in terms of their cognitive and non-cognitive skills than students from more privileged families? Second, do first-generation students have lower grade point averages (GPAs)?, and if yes, can they be explained by variation in cognitive and non-cognitive skills? Third, are skills compensating for disadvantaged produced by social origin? Answering these questions in the Australian higher education context provides valuable insights. Not least because more than three quarters of students benefit from public loans or scholarship grants. This a comparable proportion to Norway, a country considered to heavily invest in human capital and in moderating social disadvantage (Jongbloed & Vossensteyn 2016). Thus, access to university in Australia is less constrained by financial considerations than in other English-speaking countries, allowing for other factors, including skills to play a potentially important role.

We find that first-generation students have lower grade point averages by about a quarter of a standard deviation. The performance penalty is larger for young men. We also find that the penalty is strongest in the first semester but disappears over time. Both pre-study non-cognitive skills (conscientiousness and extraversion) and achievement tests are strong predictors of academic performance. Conscientiousness helps compensate for the performance penalty experienced by first-in-family students. Correcting for reporting heterogeneity in personality traits using the anchoring vignettes only leads to small changes in the personality trait distribution. However, the adjustment suggests that highly conscientious first-generation students do not outperform their more advantaged peers, a result which we find without reporting error adjustment. Nevertheless, our findings accentuate the importance of non-cognitive skills as key indicators of university readiness, and their potential for closing the socioeconomic gap in academic performance.

The remainder of this paper proceeds as follows. We present a theoretical framework for the decision to pursue a university education and an overview of recent literature in Section 2. Section 3 describes our linked survey and administrative data. Section 4 outlines the empirical framework. Section 5 presents and discusses the estimation results. Section 6 discusses the implications of our findings for policies that promote upward mobility. Section 7 concludes.

2. Background

2.1. Theoretical framework

Our study focuses on the relationship between socioeconomic status and academic performance. Specific attention is paid to the role of cognitive and non-cognitive skills in moderating the socioeconomic gradient in academic performance. We assume that children from disadvantaged backgrounds have higher psychic and opportunity costs of university education and that these costs are a function of their cognitive and non-cognitive abilities.

A typical model of post-secondary education choice (*e*) models the decision to attend university as a function of its net benefits (see Heckman et al. 2006, Carneiro, Hansen & Heckman 2003, Cunha, Heckman & Navarro 2005). Students choose the level of education, D_e which maximises net benefits I_e:

$$D_e = \arg\max\{I_e\},$$

where the net benefit of education is:

$$I_e = X_e'\beta + \alpha_e^c f^c + \alpha_e^N f^N + e_e \text{ for } e = 1, \dots, \bar{E}.$$

X_e is a vector of observable characteristics which includes perceived wage returns, perceived costs including psychic and opportunity costs associated with each level of education, and sociodemographics. Factor loadings α_e^c and α_e^N are associated with cognitive and non-cognitive latent abilities f^c and f^N respectively, influencing the potential education benefit. Students with high cognitive and non-cognitive skills will attain a greater net benefit from education, whereby α_e^c and α_e^N may be considered parameters on preferences, technology and endowments of skills f^c and f^N which generate academic outcomes. e_e represents an idiosyncratic error for each level of education, independent across levels of education and independent of f^c, $f^N$and X.

This model highlights that the socioeconomic gradient in university education could arise from heterogeneity in both the benefits and the costs, financial and non-financial, of university education. Consider specifically the psychic costs, these may be greater for students from disadvantaged backgrounds. If this is true, then the model would predict that disadvantaged students are less likely to attend university than students from more privileged backgrounds. Jacob (2002) suggests that non-cognitive skills are good proxies for these psychic costs, as they help

students to "navigate college life" (Jacob 2002).[1] Other studies proxy psychic costs with measures of cognitive ability (Carneiro et al. 2003, Cunha et al. 2005).

The consequence is that students at university are highly selected by their cognitive and non-cognitive skills (see Kassenboehmer, Leung & Schurer 2018), and this is especially true for students from disadvantaged backgrounds. Thus, controls for a student's pre-study skill endowments are important, as without them the likely result is a biased estimate of the socioeconomic gradient in university participation and performance.

2.2. Socioeconomic gradients in university performance

The empirical evidence on the socioeconomic gradient shows that students from disadvantaged backgrounds perform worse at university. The literature predominantly focuses on the impact of financial constraints on university education. The argument is that to cover living expenses, students from disadvantaged backgrounds are more likely to work outside university which takes away time spent studying (Walpole 2003, 2008, Pascarella, Pierson, Wolniak & Terenzini 2004).

Pascarella et al. (2004) is one of the few studies that focuses on students who are the first in their family to attend university. Using data from eighteen four-year colleges in the United States, the study finds that first-generation students perform worse at university relative to their peers. First-generation students are more likely to engage in outside employment and to study fewer units, while also being less likely to live in college, to participate in extra-curricular activities or to be accustomed to college expectations (for example, the importance of deadlines).[2] Cobb-Clark & Gørgens (2012) find strong socioeconomic inequalities in parental financial support of young adults studying in Australia, yet suggest that the differences in financial support are not the main cause of the socioeconomic gradient in educational outcomes. Stinebrickner & Stinebrickner (2003), using administrative data from a US college which provides full tuition and board subsidies,

[1] Jacob (2002) finds that non-cognitive skills affect the gender gap in higher education attendance, where non-cognitive skills are measured by grades and effort in school, student behaviour and if a student had ever been retained in a grade.
[2] The evidence on the link between outside work and academic performance is mixed. Using data from the 1996 Beginning Postsecondary Students Longitudinal Study, Bozick (2007) finds that students from low socioeconomic backgrounds are more likely to engage in outside employment. He estimates that working more than 20 hours per week is associated with higher incidence of dropping out of college, conditional on sociodemographic characteristics, family obligations, financial aid and state unemployment rates. DeSimone (2008) uses an instrumental variable approach to measure the relationship between employment and academic performance. Using paternal schooling achievement and religion as a factor affecting student labour supply but unrelated to academic performance, the study finds an additional work hour each week is associated with a fall in GPA by 0.011 points. In contrast, Dustmann and van Soest (2007) find that employment does not significantly affect performance of full-time students.

find a socioeconomic gradient in college performance and dropout, despite the absence of financial costs.

Other studies focus on family social support and cultural capital. Cheng, Ickes & Verhofstadt (2012) find that family encouragement is a key predictor of both levels and stability of grade point averages, particularly for female students. Doren & Grodsky (2016) show that inequality across families in parents' skills is likely to explain the socioeconomic gradient in student academic performance. Walpole (2003, p.49) suggests that the achievement gap is likely due to differences in cultural capital, that is – "insider knowledge which is not taught in schools, such as knowledge of high culture," but which is valued within the college community.

Some studies argue that previous academic achievement (a noisy proxy of cognitive ability), used in many countries to regulate university access, may explain the socioeconomic gradient in university access and performance. In Australia, universities select students based on a standardised university admissions test score, the so-called ATAR score (see Section 3.2. for details). Previous research shows a strong socioeconomic gradient in these ATAR scores (Li & Dockery 2015). Yet, students from low socioeconomic status schools do not perform worse at university than students from more privileged schools, holding their past academic achievement constant. Messinis & Sheehan (2015) show that the socioeconomic gradient of academic performance holds only for students with low ATAR scores, while students from disadvantaged backgrounds, but with high ATAR scores, outperform more privileged students.

2.3. Non-cognitive skills and academic performance

Non-cognitive skills are considered important predictors of educational attainment and academic performance (see Almlund et al. 2011 for an overview). Some traits are more important than others. Komarraju, Karau, Schmeck & Avdic (2011) find that personality, measured by the Big Five personality traits, explains 14 percent of the variance in grade point averages. Frequently credited as a super-trait (Roberts, Lejuez, Krueger, Richards & Hill 2014), Conscientiousness is singled out as one of the most important non-cognitive skills in determining study outcomes, both in high school (Noftle & Robins 2007) and at university (Chamorro-Premuzic & Furnham 2003, Kappe & van der Flier 2012, Trapmann, Hell, Hirn & Schuler 2007). Delaney, Harmon & Ryan (2013) find that this trait shapes undergraduate study behaviours, including lecture attendance and additional study hours, behaviours important for exam outcomes. Some argue therefore that Conscientiousness is considerably more powerful in predicting grade point averages than intelligence (Kappe & van der Flier 2012).

Other important traits often mentioned in this literature are internal Locus of Control or Grit. Multon, Brown & Lent (1991) find that internal control beliefs explain approximately 14 percent of the variance in academic performance of school children. They are also important predictors of college grade point averages (Richardson, Abraham & Bond 2012). Duckworth, Peterson, Matthews & Kelly (2007) find that Grit, a measure of an individual's motivation to reach a goal and to surpass obstacles, explains 4 percent of variance in GPA at the University of Pennsylvania, suggesting it is not simply cognitive ability which allows high performing students to succeed.

More importantly, the literature has identified critical interaction effects between cognitive and non-cognitive skills. Carneiro, Crawford & Goodman (2007) show that the impact of social skills on educational attainment is potentiated for students with high levels of cognitive ability. The authors also find that social skills (for example, Extraversion), although uncorrelated with parents' education, are strongly related to a student's family relationships, number of siblings and the interest of the mother in her child's education.

While there has been extensive research on the relationship between non-cognitive skills and academic outcomes, the moderating effects of non-cognitive skills on the relationship between socioeconomic status and university performance has not received much attention in the literature. We contribute to the previous literature by rigorously testing for the moderating effects of cognitive and non-cognitive skills in the relationship between socioeconomic status and academic performance, with a special focus on first-in-their family students.

3. Data

3.1. Linked Survey and Administrative Data

This study exploits unique data from a survey fielded at a leading Australian university, linked to administrative student records. We collected five different personality measures: three of the Big

Five personality inventory[3], Locus of Control[4], and Grit[5]. A full description of the instruments used to collect the personality traits is presented in Table 1. We also collected information on socioeconomic status (parental education levels), and important family determinants of the decision to pursue university education (financial support, encouragement, role models). We also added anchoring vignettes, fictional characters whose personality should be described by survey respondents alongside their own personality, into the survey for Conscientiousness and Openness to Experience (Appendix C). These allow us to control for reporting heterogeneity in the personality self-assessments.

[TABLE 1 ABOUT HERE]

The linked administrative student record data includes information on four semesters of students' grade point averages (2015-2016), university records on parental socioeconomic status, and proxies for pre-university cognitive skills (standardised university admissions test scores, ATAR). Our analysis benefits from the use of multiple indicators of socioeconomic status, including an official university definition based on the postcode of residence at enrolment, whether the student is first-generation student, as well as school type and self-reported measures of family support.

We fielded our survey during the first weeks of the 2015 university academic year. In total, around 1,000 first-year students responded to the survey. With the permission of the participants, the survey data was linked to university administrative data from student records (around 800 students). Academic performance at university was tracked over the subsequent four semesters for those students who remained enrolled at the university. The administrative data also provides

[3] Participants completed 20 self-report items designed to elicit measures of three Five Factor Model (FFM) personality traits: conscientiousness (C), extraversion (E) and openness to experience (O). The Big Five 'trait descriptive adjective' (TDA) marker method upon which our measurement strategy is based was originally developed by Goldberg (1992) and a shorter version (the 'Mini-Markers'), was developed and validated by Saucier (1994). We utilise an adapted subset of the Mini-Markers based on Losoncz (2009) and used in Cobb-Clark & Schurer (2012), Elkins, Kassenboehmer & Schurer (2017), and Kassenboehmer, Leung, and Schurer (2018). The facets C, E, and O are indexed by seven, six, and six TDA items respectively. Participants indicate the degree to which each of the 20 adjectives describes them on a seven-point response scale, ranging from 1 ("Does not describe me at all") to 7 ("Describes me very well").

[4] We collected seven of the original items from the Psychological Coping Resources component of the Mastery Module developed by Pearlin & Schooler (1978). Mastery refers to beliefs about the extent to which life's outcomes are under one's own control. Those with an internal locus of control generally believe that life's outcomes are due to their own efforts, while those with an external locus of control believe that outcomes are mainly due to external factors (Gatz and Karel 1993). The same measures have been used in Cobb-Clark & Schurer (2013), Cobb-Clark, Kassenboehmer, & Schurer (2014), and Elkins et al. (2017).

[5] We measure grit using the self-report Short Grit Scale (Grit-S) developed and validated by Duckworth et al. (2007) and Duckworth & Quinn (2009). Grit-S is designed as a brief measure of trait-level perseverance and passion for long-term goals. The scale consists of eight items, divided evenly between two subscales, 'consistency of interest' and 'perseverance of effort', with half of the items reverse-coded. Participants provide an indication of the degree to which each item applies to them on a five-point scale from 1 ("Very much like me") to 5 ("Not like me at all").

information on the students' secondary schooling background, prior academic performance and socioeconomic status.

The analysis is restricted to bachelor degree students who were in their first year in 2015, who self-selected into the survey and gave permission to link their survey to the administrative data. After dropping 66 students either with missing information or over 30 years of age, we obtained key variables on an estimation sample of 613 students. Only 6 percent of the sample dropped out of university after their first year, suggesting that the data does not suffer from severe attrition bias.

3.2. Summary Statistics

Table 2 shows summary statistics for sociodemographic, socioeconomic and employment characteristics, standardised university admissions test scores, as well as a summary of performance indicators, which are explained in depth below. The average age of our sample of students is just under 19 years, with less than 10 percent of students above the age of 20 years. Almost three quarters of all respondents were female.

Students' socioeconomic status is measured using data on location of residence and family characteristics. We construct two different measures. First, we use the official university definition of socioeconomic status, which is based on the place of students' residence (postcode) upon enrolment. The university administrative data indicates 5.2 percent of students are of low socioeconomic status (referred to as 'Low SES'). Second, we proxy for socioeconomic disadvantage with a measure that indicates whether the student is the first in their family to attend university, a measure which we collected from the student survey. Specifically, a student is counted as a first-generation student if neither parent attended university. This is our primary and preferred measure of socioeconomic status, as parents' education level may act as a proxy for family wealth while also encompassing family attitudes towards learning, parental skills and familiarity with the university system. Of the sample, 27 percent of students are first-generation.

Cognitive skills are assessed by students' university admissions test score, the Australian Tertiary Admission Rank (ATAR), which provides a measure of a student's overall academic achievement at high school (Universities Admissions Centre 2017). The ATAR is the main criteria by which a student is accepted into universities in Australia. An ATAR was available for 72 percent of the students in the sample. The average ATAR score was 89.0 with a standard deviation of 8.2.

Students' academic performance at university was tracked through their first four semesters, from semester 1 in 2015 to semester 2 in 2016. A student's weighted average mark

(WAM) for course work per semester is our primary measure of university performance outcomes. Students achieved an average WAM of 66.4 (out of 100) with a standard deviation of 16.9.

In comparison to nationally representative data (Table 2, column (5)), our surveyed students are more likely to be married (0.04 vs 0.02), diagnosed with anxiety (0.26 vs 0.14), and living at home (0.68 vs 0.51). They have higher ATAR scores than the normalised national average (89 vs 70). They are less likely to be international students (0.11 vs 0.25), of low socioeconomic status (0.05 vs 0.14), less likely to be first in their family to study at university (0.27 vs 0.54), and to have graduated from a public high school (0.35 vs 0.51). These differences are not surprising given the academic selectivity of the university at which the survey was fielded.

[TABLE 2 ABOUT HERE]

3.3. Characteristics of First-Generation Students

First-generation students are significantly more likely, than their non-first-generation peers, to come from a residential area with higher levels of socioeconomic disadvantage. First-generation students are also significantly less likely to receive parental financial support and encouragement to attend university (Table 3).

[TABLE 3 ABOUT HERE]

Figure 1 indicates that the distributions of WAM differ markedly across socioeconomic groups, with student academic performance positively correlated with SES. First-in-family students (shown in the left panel) scored lower WAMs on average, with a mean of 63.6, while non-first-in-family students scored 67.5 on average. The distributions of WAM between these two groups, and between the medium and high and low and high SES groups, are significantly different (Table A1, Online Appendix).

[Figure 1 ABOUT HERE]

Figure 2 shows a positive association between ATAR scores and socioeconomic status, independent of the definition of socioeconomic status. The mean score for first-generation

students was 86.8, while for non-first-generation students it was 89.8. The distributions of ATAR scores are significantly different.[6]

[Figure 2 ABOUT HERE]

Figure 3 shows the distribution of non-cognitive skills by socioeconomic status. The figures suggest a modest yet statistically significant difference in the distributions of non-cognitive skills by socioeconomic status (Table A1, Online Appendix). The distribution of first-generation and low SES students generally lies to the left of that for non-first-generation or higher SES students.

[Figure 3 ABOUT HERE]

4. Estimation results

In our empirical analysis, we seek to answer the following questions: (1) Is there a socioeconomic gradient in cognitive and non-cognitive skills among university students?, (2) Is there a socioeconomic gradient in academic performance at university?, and (3) Do cognitive and non-cognitive skills moderate the socioeconomic gradient in academic performance?

4.1. Socioeconomic gradient in skills

We start our analysis by assessing whether pre-university skills differ by student socioeconomic background. We estimate a linear regression model in which a measure of skills (S_i) is regressed on socioeconomic status (SES_i) and a vector of control variables (X_i):

$$S_i = \alpha_0 + \alpha_1 SES_i + \beta X_i' + \varepsilon_i \qquad (1)$$

Our parameter of interest is α_1, which isolates the socioeconomic gradient in non-cognitive skill, expressed in standard deviation (SD) difference. The controls variables are gender, a full set of age fixed effects, a dummy variable for international students and a dummy variable for whether the student has ever been diagnosed with anxiety.

Table 4 presents estimation results of interest obtained from Eq. (1) using either non-cognitive or cognitive skills (ATAR score) as outcome measures. In column (1), each row reports

[6] The sample ATAR scores are generally higher than the Australian average score of 70. This is primarily due to the selection of high-calibre students to the university, using minimum ATAR requirements for most courses.

the coefficient on first-generation status from a separate regression with each skill as the dependent variable. In columns (2)-(3) we report the coefficient on low and high socioeconomic status (relative to medium SES).[7]

[TABLE 4 ABOUT HERE]

We find no statistically significant differences in non-cognitive skills between first-generation students and students who are not for all skills except for the Locus of Control. There is suggestive evidence that first-in-family students have a more external Locus of Control than their non-first-in-family counterparts (-0.157 SD, $p<0.10$). We obtain similar findings using nationally representative data on young Australians sourced from the Household, Income and Labour Dynamics Survey (HILDA), in which we have the same non-cognitive skill measures available (except for Grit) and a comparable measure of first-generation status. We use a general summary measure of cognitive skills available in HILDA (Table B1, Online Appendix). We find a very similar socioeconomic gradient in cognitive ability of about one third of a standard deviation in both our university sample and the nationally representative survey data and somewhat similar gradients in non-cognitive skills.[8]

When using the official university (administrative) definition of socioeconomic disadvantage, we find only one significant difference in non-cognitive skills. Relative to students from midrange SES, students from higher SES report higher levels of Extraversion by 0.2 SD ($p<0.10$), while students from lower SES do not differ in their Extraversion scores. Similarly, we find no consistent evidence for differences in non-cognitive skills by other available measures of SES.[9]

However, disadvantaged students unambiguously score lower on the university entry examinations, independent of the SES definition. First-generation students have ATAR scores that are around a third of a standard deviation lower ($p<0.01$). Students from more privileged backgrounds (relative to midrange SES) according to the official university definition score close

[7] Full results are available from the authors on request.

[8] We use unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey (Melbourne Institute 2017). 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 author and should not be attributed to either DSS or the Melbourne Institute.

[9] See Table A2 in the Online Appendix.

to 0.4 SD higher ($p<0.01$), while low SES students score about 0.3 SD lower (not statistically significant) relative to their midrange SES peers.

Together, these results demonstrate that first-in-family students are not so different in terms of their non-cognitive skills but they are more disadvantaged in terms of pre-university academic achievement. These findings imply that socioeconomically disadvantaged students either have lower cognitive innate ability, or that the ATAR score – despite being a strong predictor of university readiness – is a noisy measure of cognitive ability. That is, the ATAR score appears to reflect elements of economic privilege, including for example, being better prepared at taking tests or having access to educational resources.

These findings are also consistent with the model of selection into university presented above (Section 2.1). In that model, the net benefit to tertiary education is increasing in cognitive ability which can be a proxy for psychic costs. Hence, the estimated relationship between socioeconomic status and ATAR suggests that the net benefit of tertiary education may be smaller for socioeconomically disadvantaged students.

4.2. Socioeconomic gradient in university performance

We model academic performance using a random effects (RE) specification, exploiting the time variation in test scores, and controlling for a set of observable characteristics as follows:

$$WAM_{it} = \gamma_0 + \gamma_1 SES_i + \gamma_2 X_i' + \gamma_3 CS_i + \mu_i + \varepsilon_{i,t} \qquad (3)$$

$$\text{for } i = 1, \dots, n \text{ and } t = 1, \dots, 4$$

WAM_{it} is the standardised weighted average mark for individual *i* in semester *t* and γ_1, the coefficient on the SES indicator, is our parameter of interest. Demographic characteristics (X) and pre-study cognitive skills (CS) (ATAR score) are included as controls to capture past inputs and cognitive ability. Random shocks, $\varepsilon_{i,t}$, and time-invariant, individual-specific heterogeneity, μ_i, capture unobserved random and student-specific variation in WAM, respectively.

To test whether non-cognitive skills mediate the relationship between socioeconomic status and academic performance, we add our measures of NCS, as shown in Equation (4):

$$WAM_{it} = \gamma_0 + \gamma_1 SES_i + \gamma_2 X_i' + \gamma_3 CS_i + \gamma_4 NCS_i + \mu_i + \varepsilon_{i,t} \qquad (4)$$

Although Equations (3) and (4) allow for individual-specific unobserved heterogeneity, the RE specification assumes that this unobserved heterogeneity is independent of the other regressors in the model, including socioeconomic status. This does not allow for unobservable, cumulative

inputs into the youth production function of academic achievement. Following Todd & Wolpin (2003), one solution is a value-added model, in which we condition the analysis on lagged measures of the outcome variable.[10] We use lagged achievement at university ($WAM_{i,t-1}$) as a proxy for both innate ability and previous inputs in the study process, over and above the influence of pre-university, measured cognitive (ATAR) and non-cognitive skills:[11]

$$WAM_{it} = \gamma_0 + \gamma_1 SES_i + \gamma_2 X_i' + \lambda_1 CS_i + \lambda_2 NCS_i + \tau WAM_{i,t-1} + \mu_i + \varepsilon_{i,t} \quad (5)$$

The main coefficient of interest is γ_1. It identifies the association between WAM and socioeconomic status conditional on controls and past weighted average marks.

Table 5 shows the estimation results obtained from Eqs. (3) to (5). Going across the table, each subsequent column gradually adds sets of control variables. In a model without control variables, first-generation students achieve WAMs that are more than a quarter of a standard deviation (SD) lower than that of their peers ($p<0.01$). This penalty is equivalent to a gap of approximately five marks, or half of the number of marks lying between a credit and distinction average. As control variables are gradually added, the size of the estimated coefficient on the indicator for first-generation status shrinks. In a model with full controls (column (4)), the socioeconomic gradient in WAM is equivalent to less than 0.2 SD ($p<0.10$).[12]

[TABLE 5 ABOUT HERE]

The coefficient on the first-generation student indicator drops most significantly in absolute value when controlling for pre-university cognitive skills (ATAR), from 0.25 SD ($p<0.01$) to 0.16 SD (not significant). This is consistent with the findings and discussion in Section 4.1 above. The ATAR score appears to be a noisy measure of cognitive skills, proxying also for, or at least highly correlated with, socioeconomic disadvantage. When using alternative definitions of low socioeconomic status, including school type, the university's official socioeconomic status indicator, or whether the student lives with their parents, we obtain similar findings. There is no

[10] Todd & Wolpin (2003) and Fiorini & Keane (2014) use the so-called value-added model in the context of skill development of children. Both studies model the production function of cognitive and non-cognitive skills in children, explicitly modelling a child's development as dependent on the historical accumulation of family inputs, schooling inputs and innate ability. Kassenboehmer et al. (2018) and Elkins et al. (2017) use the value-added model in the context of youth non-cognitive skill development.

[11] The key assumption of the model is that the effect of the non-cognitive and cognitive skills decline at the rates λ_1 and λ_2 over time respectively.

[12] The full set of estimates for Equation (4) are shown in Table A3 in the Online Appendix. In the same table we show the results from a series of sensitivity checks on these estimates. The results are extremely similar, and our conclusions remain unchanged when adding further controls such as having children, being married, birth order or hand used for writing.

statistically significant association between socioeconomic status and WAM when defining disadvantage through geographic inequality (postcode) or school type.[13]

Controlling for non-cognitive skills has little impact on the first-in-family penalty (column (4)). Pre-university non-cognitive skills are however significantly associated with academic performance. For instance, a one standard-deviation increase in Conscientiousness corresponds to a 0.14 SD increase in WAM ($p<0.01$), while a one standard deviation increase in Extraversion is associated with a fall of 0.11 SD in WAM ($p<0.01$). Although not as large as the effect size of the ATAR score, these associations are sizable. After controlling for reporting heterogeneity in personality assessment using the anchoring vignettes, we obtain similar estimated coefficients (Table C4, Online Appendix).

Thus far, we have considered the impact of socioeconomic status and skills on performance data averaged across four semesters. We now explore whether first-generation students catch up over time. In Table 6, we report the estimated coefficients on the first-generation indicator for each semester (with a full set of control variables, as in column (4), Table 5). In semester 1, the first-in-family performance penalty is 0.20 SD ($p<0.01$). From semester 2 onward, the penalty shrinks significantly. By semester 4, the penalty is neither statistically nor economically significant.

[TABLE 6 ABOUT HERE]

Our findings suggest that performance in semester 1 strongly determines subsequent performance. This implies intertemporal dependence that, if left unaccounted for, may bias our estimate on the first-generation status. We therefore re-estimate our benchmark model using the value-added specification (Eq. (5)). These results are shown in the final column of Table 5. When controlling for the intertemporal dependence in academic performance, the first-in-family academic penalty disappears. The coefficient estimate is statistically and economically insignificant with a magnitude of less than 0.02 SD. The coefficient on the lagged WAM is statistically significant and large in magnitude (0.6 SD, $p<0.01$). That is, consistent with the semester by semester results from Table 6, the value-added model suggests that first generation students perform poorly only in the first semester but catch up over time. This result is in line with the hypothesis that familiarity with and an understanding of university expectations shapes student achievement.[14]

[13] See Table A4 in the Online Appendix.

[14] We also considered the relationship between the variability in academic performance and first-generation status. Using the standard deviation and range of the weighted average mark over the four semesters as our measure of

Separating the analysis by gender, shows that the first-generation penalty is larger for young men. This also holds true for the results across semesters – the penalty is larger for young men but becomes smaller and less significant over the four semesters.[15] Furthermore, we explored whether first-in-family students choose easier subjects than their peers. In their first and second semesters, first-generation students are more likely to take introductory units, which are generally easier than advanced subjects, by 4 and 6 percentage points respectively.[16] This may imply that these students are more unsure of their long-term study plans relative to non-first generation students. However, there is no significant heterogeneity in subject choice in the second year by first-generation status. Therefore, it does not appear that our main findings are biased by variation in the subject choice of first-generation students.

4.3. Skills as moderator of the socioeconomic gradient in academic performance

Finally, we test whether the socioeconomic gradient is moderated by cognitive and non-cognitive skills. To do so, we extend Eq. (4), including interactions between socioeconomic status and each of the cognitive or non-cognitive skill measures S_i in turn ($S_i \times SES_i$):

$$WAM_{it} = \gamma_0 + \gamma_1 SES_i + \gamma_2 X_i + \gamma_3 CS_i + \gamma_4 NCS_i + \gamma_5 S_i \times SES_i + \mu_i + \varepsilon_{i,t}, \quad (6)$$

γ_5 is our parameter of interest. If γ_5 is zero, we conclude that skills do not moderate the socioeconomic gradient in university performance. To allow for non-linearities in this relationship, we include higher polynomials of the skill measure (using the Akaike information criteria to select the optimal number of polynomials).[17] We find significant interaction effects for the ATAR score, Conscientiousness, Extraversion and the Locus of Control. According to the AIC, we model the interaction effects for the ATAR score and Extraversion with a quadratic polynomial in the skill, and for Conscientiousness and the Locus of Control with a cubic polynomial.

Figure 4 summarises the estimated interaction effect between first-generation status and the ATAR score (Panel (a)) or Conscientiousness (Panel (b)) and their 95% confidence intervals.[18] In each figure the vertical axis shows the marginal effect of being a first-generation student on the

variability, we do not find a significant association between first-in-family and variability in academic performance. First-generation students do not appear to experience greater variability in their academic performance than students from more privileged backgrounds. Results available from the authors on request.

[15] Results available from the authors on request.

[16] Results available from the authors on request.

[17] We use the AIC for interactions of up to a fifth order polynomial in the skill measure of interest with the first in family characteristic.

[18] Full estimation results are presented in Table A5 and comparable figures for each skill measure in Figure A1 in the Online Appendix.

weighted average mark (WAM), expressed in standard deviations, while the horizontal axis presents the full distribution of the skill measure, also expressed in standard deviations.

Figure 4(a) demonstrates that first-generation students with cognitive skills (ATAR) between the middle and lower end of the skill distribution experience a first-in-their family WAM penalty of up to -0.6 SD. In contrast, first-generation students with very high ATAR scores (1 or 2 SD above the mean) score up to 1.5 SD above the WAM mean, offsetting the first-in-family performance penalty. These differences are statistically significant at the 5% level.

[FIGURE 4 ABOUT HERE]

Conscientiousness also offsets the performance penalty of first-generation students (Figure 4(b)), although to a lesser extent in comparison to cognitive skills. Students at the upper end of the Conscientiousness distribution have a WAM premium of about 0.3 SD above the mean. In contrast, for low to medium values on the Conscientiousness score, there is a statistically significant penalty of about -0.5 to -0.3 SD. Surprisingly, there is no statistically significant penalty for first-generation students with extremely low Conscientiousness scores. One explanation for this counter-intuitive result may be that very low-performing first-generation students rank themselves too highly on the Conscientiousness scale, while mid to low-performing students rank themselves too negatively relative to others, a phenomenon reported in West, Kraft, Finn, Martin, Duckworth & Gabrieli et al. (2016).[19]

Controlling for reporting heterogeneity that may systematically vary with socioeconomic background with our anchoring vignette experiment, we find less evidence for a compensatory effect of Conscientiousness. The anchoring vignettes allow us to correct for individual heterogeneity in response style, and this adjustment predominantly affects the most extreme reports of Conscientiousness (Figure C4, Online Appendix C). Using the adjusted scores, we continue to find that first-generation students in the low to middle range of the Conscientiousness distribution experience a significant WAM penalty (-0.5 to -0.25 SD). However, the impact of

[19] West et al (2016) find evidence for this hypothesis analysing data on students who entered a Boston-based charter school through a lottery. The authors find that students who enter highly-selective charter schools tend to adjust their Conscientiousness and Grit scores downward because they adopt a new, higher standard of what they consider as high level.

Conscientiousness on performance does not differ between first-generation and non-first-generation students above the mean and at the upper end of the Conscientiousness distribution.[20]

5. Conclusion

In this study, we analyse novel linked survey and administrative data which we collected on an incoming cohort of students at a leading Australian university (2015). The aim of the study was to better understand the facilitators and constraints that first-in-their-family students face at university. Little empirical evidence exists on this policy-relevant population.

Our findings are multi-fold. Most importantly, first-in-their-family students experience no inequalities in pre-university non-cognitive skills but arrive at university with lower pre-university cognitive skills, which we proxied with standardised university admissions test scores. This finding is consistent with supplementary evidence obtained from the analysis of sample sourced from a nationally representative survey, the so-called Household, Income, and Labour Dynamics in Australia (HILDA) survey. Second, first-in-their-family students have lower grade-point averages at the beginning of their studies, but they tend to catch up over time. Non-cognitive skills (Conscientiousness, Extraversion) predict academic performance almost as strongly as cognitive skills. Third, both cognitive skills and Conscientiousness compensate for the academic penalties produced by social origin. For instance, low levels of Conscientiousness exacerbate performance penalties, while high levels of conscientiousness over-compensate for the performance penalties by social origin. Yet, the over-compensating effects at the high-end of the Conscientiousness spectrum disappear once controlling for extreme response styles using vignettes.

Our findings contribute to an international literature that considers the role of both cognitive and non-cognitive skills instrumental in facilitating social mobility (Heckman and Mosso 2014; Heckman 2000) and success in life (Almlund et al. 2011). Our findings accentuate the importance of non-cognitive skills, and, in particular, Conscientiousness, in influencing academic outcomes for disadvantaged students. Conscientiousness has often been discussed in the literature as a super-trait because of its health benefits and its association with job and academic performance (Roberts et al. 2014). It is positive to see that first-in-their-family students have no

[20] We also find that first-generation students with average or slightly above average levels of Extraversion (up to one standard deviation above the mean of the Extraversion distribution) experience a statistically significant penalty in their academic performance of around -0.25 SD. Those with high or low levels of Extraversion do not experience a first-in-family penalty. In contrast, students that are at the extremes of the Locus of Control distribution, that is, those that have a very internal, or very external, locus of control experience a statistically significant first-in-family penalty of -0.5 SD or more in the WAM.

Conscientiousness disadvantage upon entry into university and that this skill helps students to achieve high grade point averages. This finding holds even when correcting for extreme response styles in self-assessed personality measures through our anchoring vignettes.

Our study also contributes to an emerging literature that questions the reliability of self-assessed non-cognitive skill measures (see Almlund et al. 2011, West et al. 2016). We build on previous studies which developed and applied so-called anchoring vignettes to be able to compare education outcomes based on self-assessed measures across cultures (He, Buchholz & Klieme 2017) and in the context of personality assessment (Mõttus, Allik, Realo, Rossier, Zecca & Ah-Kion et al. 2012, Bolt, Lu & Kim 2014, Primi, Zanon, Santos, De Fruyt & John 2016). To the best of our knowledge, our study is one of the first to show that individual-specific, extreme response styles in personality assessment tasks may lead to over-estimates of the benefits of Conscientiousness in the context of inequalities in educational achievement. Certainly, more research is needed to better understand the breadth of response styles in a given population. Specifically, in our study we assume that response heterogeneity is fixed across personality assessments of your own and fictional others' profiles. Other types are possible and should be explored.

Our study has limitations that require mention. First, our survey includes only pre-study information, and lacks inputs or changes in circumstances throughout four semesters of data. Our analysis therefore does not allow for a dynamic modelling approach. Kassenboehmer, Leung and Schurer (2018) suggest that some non-cognitive skills, although not Conscientiousness, are endogenous to the experiences students have at university. Pascarella et al. (2004) find that student experiences have heterogeneous effects on non-cognitive skill development. Therefore, the study recommends that models of academic achievement should reflect the fact that disadvantaged students accumulate more cultural capital during university life than more privileged students. Indirectly, our models have done so by allowing for heterogeneous achievement dynamics across students.

Another limitation of our study sample comes with self-selection of students into the survey. The survey was advertised widely across campus through posters and fliers. We also sent a series of emails to incoming students in a faculty of arts and social sciences. Thus, our cohort insights may not be externally valid. This could partially explain why our student cohort characteristics differ from the national average. Students also had to give their permission to link survey and administrative data. Selection and external validity concerns are of course problems that plague every university cohort study (e.g. Stinebrickner and Stinebrickner 2003). In our case, 98

percent of survey respondents agreed to have their records linked, which makes it unlikely to have caused additional selection bias. It is furthermore reassuring that the distribution of pre-university non-cognitive skills for this sample of students is representative of a broader sample of university students using nationally representative data.

References

Almlund, M, Duckworth, AL, Heckman, JJ, & Kautz, T 2011, 'Personality Psychology and Economics'. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Vol. 4, Volume 4, Chapter 1, pp. 1-181, Elsevier B.V.

Attanasio, O, Blundell, R, Conti, G, & Mason, G 2020, 'Inequality in Socio-Emotional Skills: A Cross-cohort Comparison', *Journal of Public Economics*, vol. 96, no. 4, pp. 898-912.

Blanden, J, Gregg, P, & Machin, S 2005, 'Intergenerational Mobility in Europe and North America', London: Centre for Economic Performance, London School of Economics.

Blanden, J, Gregg, P, & Macmillan, L 2007, 'Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education', *Economic Journal, vol.* 117, pp. C43–C60.

Bolt DM, Lu Y, & Kim J-S 2014, 'Measurement and control of response styles using anchoring vignettes: A model-based approach'. *Psychological Methods*, vol 19, no 4, pp. 528.

Borghans, L., Duckworth, A., Heckman, J. & Weel, B. 2008, 'The Economics and Psychology of Personal Traits', *The Journal of Human Resources*. 43. 10.1353/jhr.2008.0017.

Bowles, S, Gintis, H, & Osborne, M 2001, 'Incentive-Enhancing Preferences: Personality, Behavior, and Earnings'. *American Economic Review*, vol. 91, no. 2, pp. 155-158.

Bozick, R 2007, 'Making it through the First Year of College: The Role of Students' Economic Resources, Employment and Living Arrangements', *Sociology of Education*, vol. 80, no. 3, pp. 261-284.

Breen, R., & Müller, W. 2020, 'Education and Intergenerational Social Mobility in Europe and the United States'. Stanford, California: Stanford University Press.

Bütikofer, A, & Peri, G 2020, 'How Cognitive Ability and Personality Traits Affect Geographic Mobility', *Journal of Labor Economics* (forthcoming).

Carneiro, P, Crawford, C, & Goodman, A 2007, 'The Impact of Early Cognitive and Non-Cognitive Skills on Late Outcomes', London School of Economics Centre for the Economics of Education, ISBN 978-0-85328-188-7

Carneiro, P, Hansen, KT, & Heckman, JJ 2003, '2001 Lawrence R. Klein Lecture: Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice', *International Economic Review*, vol. 44, no. 2, pp. 361-422.

Chamorro-Premuzic, T, & Furnham, A 2003, 'Personality predicts academic performance: Evidence from two longitudinal university samples', *Journal of Research in Personality* vol. 37, no. 4, pp. 319-338.

Cheng, W, Ickes, W, & Verhofstadt, L 2012, 'How is family support related to students' GPA scores? A longitudinal study', *Higher Education*, vol. 64, no. 3, pp. 399-420.

Chesters, J, & Watson, L 2013, 'Understanding the Persistence of Inequality in Higher Education: Evidence from Australia', *Journal of Education Policy* 28(2), 198–215.

Cobb-Clark, D, & Gørgens, T 2012, 'Parents' Economic Support of Young-Adult Children: Do Socioeconomic Circumstances Matter?', *Journal of Population Economics*, vol. 27, pp. 447–471

Cobb-Clark, D., & Schurer, S. (2012). The stability of the Big-Five personality traits. Economics Letters 115(1), 11-15.

Cobb-Clark, D, & Schurer, S 2013, 'Two economists' musings on the stability of locus of control', *Economic Journal* 123(570), F358-F400.

Cobb-Clark, D, Kassenboehmer, S, & Schurer, S 2014, 'Healthy habits: What explains the connection between diet, exercise, and locus of control?' *Journal of Economic Behavior & Organization* vol. 98, pp. 1-28.

Conley, D. T. 2003, Understanding university success. Eugene, OR: Center for Educational Policy Research, University of Oregon.

Cunha, F, Heckman, JJ, & Navarro, S 2005, 'The 2004 Hicks Lecture: Separating Uncertainty from Heterogeneity in Life Cycle Earnings', *Oxford Economic Papers*, vol. 57, no. 2, pp. 191-261.

Delaney, L, Harmon, C, & Ryan, M 2013, 'The role of noncognitive traits in undergraduate study behaviours', *Economics of Education Review*, vol. 32, pp. 181-195.

Department of Education and Training 2016, *Driving Innovation, Fairness and Excellence in Australian Higher Education*, viewed 12 May 2017, https://docs.education.gov.au/system/files/doc/other/he_reform_paper_driving_innovation_fairness_and_excellence_3_may_2016.pdf.

Department of Education and Training 2017, *Completion Rates of Higher Education Students – Cohort Analysis, 2005-2014*, viewed 12 May 2017, https://docs.education.gov.au/system/files/doc/other/cohort_analysis_2005-2014_0.pdf.

DeSimone, JS 2008, 'The impact of employment during school on college student academic performance', *NBER Working Paper 14006*, viewed 24 October 2017, http://www.nber.org/papers/w14006.

Doren, C, & Grodsky, E 2016, 'What Skills Can Buy: Transmission of Advantage through Cognitive and Noncognitive Skills', *Sociology of Education*, vol. 89, no. 4, pp. 321-342.

Duckworth, AL, Peterson, C, Matthews, MD & Kelly, DR 2007, 'Grit: Perseverance and Passion for Long-Term Goals', *Journal of Personality and Social Psychology*, vol. 92, no. 6, pp. 1087-1101.

Duckworth AL, & Quinn PD 2009, 'Development and Validation of the Short Grit Scale (GRIT-S)', *Journal of Personality Assessment*, vol. 91, no. 2, pp. 166-74.

Dustmann, C, & van Soest, A 2007, 'Part-time work, school success and school leaving', *Empirical Economics*, vol. 32, pp. 277-299.

Elkins, RK, Kassenboehmer, SC, & Schurer, S 2017, 'The stability of personality traits in adolescence and young adulthood', *Journal of Economic Psychology*, vol. 60, pp. 37-52.

Elkins, R, & Schurer, S 2020, 'Exploring the role of parental engagement in noncognitive skill development over the lifecourse', *Journal of Population Economics,* vol. 33, pp. 957-1004.

Fiorini, M, & Keane, MP 2014, 'How the Allocation of Children's Time Affects Cognitive and Noncognitive Development', *Journal of Labour Economics*, vol. 32, no. 4, pp. 787-836.

Gatz, M., & Karel, M. J. 1993, 'Individual change in perceived control over 20 years', *International Journal of Behavioral Development*, 16(2), 305–322.

Gensowski, M, Goertz, M, & Schurer, S 2020, 'Inequality in Personality over the Life Cycle', IZA Discussion Paper Nr 13378. IZA Bonn. http://ftp.iza.org/dp13378.pdf

Goldberg, LR 1992, 'The Development of Markers for the Big-Five Factor Structure', *Psychological Assessment*, vol. 4, no. 1, pp. 26-41.

Haveman, R., & Smeeding, T. 2006, 'The Role of Higher Education in Social Mobility', The Future of children / Center for the Future of Children, the David and Lucile Packard Foundation. 16. 125-50. 10.1353/foc.2006.0015.

He, J., Buchholz, J., & Klieme, E 2017, 'Effects of anchoring vignettes on comparability and predictive validity of student self-reports in 64 cultures', *Journal of Cross-Cultural Psychology*, vol. 48, pp. 319–334.

Heckman, JJ 2000. 'Policies to foster human capital', *Research in Economics*, vol. 54, pp. 3–56

Heckman, JJ, Stixrud, J, & Urzua, S 2006, 'The Effects of Cognitive and Noncognitive Abilities on Labour Market Outcomes and Social Behaviour', *Journal of Labour Economics*, vol. 24, no. 3, pp. 411-482.

Heckman, J.J., & Mosso, S. 2014, 'The Economics of Human Development and Social Mobility', *Annual Review of Economics,* vol. 6, no. 1, pp. 689-733.

Jacob, BA 2002, 'Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education', *Economics of Education Review*, vol. 21, no. 6, pp. 589-598.

Jerrim, J, & Vignoles, A 2015, 'University access for disadvantaged children: A comparison across countries', *Higher Education*, 70. 10.1007/s10734-015-9878-6.

Jongbloed, B, & Vossensteyn, H 2016, 'University funding and student funding: international comparisons', *Oxford Review of Economic Policy*, vol. 32, no. 4, pp. 576–595

Kappe, R. & van der Flier, H. 2012, 'Predicting academic success in higher education: what's more important than being smart?', E*uropean Journal of Psychology of Education*, vol. 27, no. 4, pp. 605-619.

Kassenboehmer, S, Leung, F, & Schurer, S 2018, 'University education and non-cognitive skill development', *Oxford Economic Papers*, vol. 70, no. 2, pp. 538-562.

King, G, Murray, CJL, Salomon, JA, & Tandon, A 2005. 'Enhancing the validity and cross-cultural comparability of measurement in survey research' *American. Political Science Review*, vol. 98, no., 1:191-207.

King G, & Wand J 2007, 'Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes', *Political Analysis*, vol 15, no 1, pp. 46-66.

Komarraju, M, Karau, SJ, Schmeck, RR & Avdic, A 2011, 'The Big Five personality traits, learning styles, and academic achievement', *Personality and Individual Differences*, vol. 51, no. 4, pp. 472-477.

Leigh, A. 2007, 'Intergenerational Mobility in Australia', *The B.E. Journal of Economic Analysis & Policy*, vol. 7(2),no.6.

Li, IW, & Dockery, AM 2015, 'Does School Socio-economic Status Influence University Outcomes?', *Australian Journal of Labour Economics*, vol. 18, no. 1, pp. 75-94.

Lindqvist, E. & Vestman, R 2011, 'The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment', *American Economic Journal: Applied Economics*, vol. 3, pp. 101-128.

Losoncz, I. 2009, 'Personality Traits in HILDA', *Australian Social Policy* No. 8, 169.

Lundberg, S 2013, 'The College Type: Personality and Educational Inequality', *Journal of Labour Economics*, vol. 31, no. 3, pp. 421-441.

Melbourne Institute, *HILDA Survey*, viewed 29 October 2017, http://melbourneinstitute.unimelb.edu.au/hilda.

Mendolia, S. & Siminski, P. 2016, 'New Estimates of Intergenerational Mobility in Australia', *Economic Record*, 92: 361-373. doi:10.1111/1475-4932.12274

Messinis, G, & Sheehan, P 2015, 'The academic performance of first year students at Victoria University by Entry Score and SES, 2009-2013', *Victoria Institute of Strategic Economic Studies*, Melbourne, https://www.vu.edu.au/sites/default/files/cses/pdfs/the-academic-performance-of-first-year-students-at-VU-by-entry-score-and-SES-2009-2013.pdf.

Mõttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., Amoussou-Yéyé, D., Bäckström, M., Barkauskiene, R., Barry, O., Bhowon, U., Björklund, F., Bochaver, A., Bochaver, K., de Bruin, G., Cabrera, H.F., Chen, S.X., Church, A.T., Cissé, D.D., Dahourou, D., Feng, X., Guan, Y., Hwang, H.S., Idris, F., Katigbak, M.S., Kuppens, P., Kwiatkowska, A., Laurinavicius, A., Mastor, K.A., Matsumoto, D., Riemann, R., Schug, J., Simpson, B., Tseung-Wong, C.N., Johnson, W 2012, 'The effect of response style on self-reported conscientiousness across 20 countries', *Personality and Social Psychology Bulletin*, vol. 38, pp. 1423–1436.

Multon, KD, Brown, SD, & Lent, RW 1991, 'Relation of Self-Efficacy Beliefs to Academic Outcomes: A Meta-Analytic Investigation', *Journal of Counseling Psychology*, vol. 38, no.1, pp. 30-38.

Noftle, EE, & Robins, RW 2007, 'Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores', *Journal of Personality and Social Psychology*, vol. 93, no. 1, pp. 116-130.

OECD (2018), Equity in Education: Breaking Down Barriers to Social Mobility, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264073234-en.

Page, LC, & Scott-Clayton, J 2016, 'Improving college access in the United States: Barriers and policy responses', *Economics of Education Review*, vol.51, pp. 4-22.

Pascarella, ET, Pierson, CT, Wolniak, GC & Terenzini, PT 2004, 'First-Generation College Students: Additional Evidence on College Experiences and Outcomes', *The Journal of Higher Education*, vol. 75, no. 3, pp. 249-284.

Pearlin, L, & Schooler, C 1978, 'The structure of coping', *Journal of Health and Social Behavior*, vol. 19, pp. 2-21.

Primi, R, Zanon, C, Santos, D, De Fruyt, F, & John, OP 2016, 'Anchoring vignettes: can they make adolescent self-reports of social-emotional skills more reliable, discriminant, and criterion-valid?', *European Journal of Psychological Assessment*, vol. 32, pp. 39–51.

Richardson, M, Abraham, C, & Bond, R 2012, 'Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis', *Psychological Bulletin*, vol. 138, no. 2, pp. 353-387.

Roberts, BW, Lejuez, C, Krueger, R.F, Richards, JM & Hill, PL 2014, 'What is conscientiousness and how can it be assessed?', *Developmental Psychology,* vol. 50, no. 5, pp. 1315-1330.

Saucier, G. 1994, Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. Journal of Personality Assessment, 63(3), 506–516.

Stinebrickner, R, & Stinebrickner, TR 2003, 'Understanding Educational Outcomes of Students from Low-Income Families: Evidence from a Liberal Arts College with a Full Tuition Subsidy Program', *Journal of Human Resources*, vol. 38, no. 3, pp. 591-617.

Todd, PE, & Wolpin, KI 2003, 'On the specification and estimation of the production function for cognitive achievement', *The Economic Journal*, vol. 113, no. 485, pp.F3-F33.

Trapmann, S, Hell, B, Hirn, J-OW, & Schuler, H 2007, 'Meta-analysis of the relationship between the Big Five and academic success at university', *Zeitschrift fur Psychologie, vol.* 215, no. 2, pp. 132-151.

Walpole, M 2003, 'Socioeconomic Status and College: How SES Affects College Experiences and Outcomes', *The Review of Higher Education*, vol. 27, no. 1, pp. 45-73.

Walpole, M 2008, 'Emerging from the Pipeline: African American Students, Socioeconomic Status, and College Experiences and Outcomes', *Research in Higher Education*, vol. 49, no. 3, pp. 237-255.

West, MR, Kraft, MA, Finn, AS, Martin, RE, Duckworth, AL, Gabrieli, CFO & Gabrieli, JDE 2016, 'Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling', *Educational Evaluation and Policy Analysis,* vol. 38, no. 1, pp. 148-170.

Universities Admissions Centre 2017, *Australian Tertiary Admission Rank (ATAR)*, viewed 5 May 2017, http://www.uac.edu.au/atar/

Universities Australia 2016*, University Participation and Quality*, viewed 21 May 2017, https://www.universitiesaustralia.edu.au/uni-participation-quality

Tables and Figures

Table 1. Personality Instruments

Big Five Personality Traits

A7 How well do the following words describe you? For each word, cross one box to indicate how well that word describes you. There are no right or wrong answers.

Does not describe me at all						Describes me very well
1	2	3	4	5	6	7
Tick X one box for each word						

A7a Talkative
A7c Orderly
A7e Deep
A7h Systematic
A7j Philosophical
A7k Bashful
A7m Inefficient
A7o Creative
A7p Quiet
A7r Sloppy
A7t Intellectual
A7u Extroverted
A7w Disorganized
A7y Complex
A7z Shy
A7ab Efficient
A7ad Imaginative
A7aj Lively

Grit

A9 Here are a number of statements that may or may not apply to you. For the most accurate score, when responding, think of how you compare to most people -- not just the people you know well, but most people in the world. There are no right or wrong answers, so just answer honestly!
Please answer with the following categories

Very much like me	Mostly like me	Somewhat like me	Not much like me	Not like me at all
1	2	3	4	5
Tick X one box for each statement				

A9a New ideas and projects sometimes distract me from previous ones.
A9b Setbacks don't discourage me.
A9c I have been obsessed with a certain idea or project for a short time but later lost interest.
A9d I am a hard worker.
A9e I often set a goal but later choose to pursue a different one.
A9f I have difficulty maintaining my focus on projects that take more than a few months to complete.
A9g I finish whatever I begin.
A9h I am diligent.

Locus of control

A8 Please indicate, by crossing one box on each line, how much you agree or disagree with each of the following statements. The more you agree, the higher the number of the box you should cross. The more you disagree, the lower the number of the box you should cross.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**

A8a I have little control over the things that happen to me
A8b There is really no way I can solve some of the problems I have
A8c There is little I can do to change many of the important things in my life
A8d I often feel helpless in dealing with the problems of life
A8e Sometimes I feel that I'm being pushed around in life
A8f What happens to me in the future mostly depends on me
A8g I can do just about anything I really set my mind to

Table 2. Summary statistics

	Mean	Std Dev	Min	Max	National***
	(1)	(2)	(3)	(4)	(5)
Age in Semester 1, 2015	18.75	1.93	17	29	20.3[b]
Female	0.72	0.45	0	1	0.54[a]
Married	0.04	0.20	0	1	0.02[b]
International student	0.11	0.32	0	1	0.25[a]
Ever diagnosed with anxiety	0.26	0.44	0	1	0.14[b]
ATAR Score*	89.01	8.19	65.57	99.95	70 [a]
First in family	0.27	0.44	0	1	0.58[b]
Lives at Home	0.68	0.47	0	1	0.51[b]
Parent Encouragement	0.91	0.29	0	1	NA
Receives Family Financial Support	0.68	0.47	0	1	NA
Monthly Financial Support ($), if received	694.57	1208.46	15	15750	NA
Low Socioeconomic Status**	0.05	0.22	0	1	0.14[b]
Medium Socioeconomic Status	0.33	0.47	0	1	0.33[b]
High Socioeconomic Status	0.50	0.50	0	1	0.53[b]
Public High School Student**	0.35	0.48	0	1	0.51 [b]
Catholic High School Student	0.14	0.35	0	1	0.25
Private High School Student	0.22	0.41	0	1	0.23
Number of Courses taken per semester	3.86	0.72	1	7	NA
Weighted Average Mark (WAM)	66.42	16.94	0		NA
Drops out after first year	0.06	0.23	0	1	0.16[a]
Number of Students	613				NA
Observations	2179				NA

Notes:
*ATAR Score is available for 442 students. Students without an ATAR or comparable entry score are mostly international or mature aged students. We control for these missing ATAR scores in the regression analysis.
**All domestic students have a socioeconomic indicator, whilst the 68 that do not are international students. For this reason, the proportion of students in the low, medium and high socioeconomic status groups do not sum to 1. The same applies for the type of high school attended – this is unavailable for international students.
*** Data constructed from Australian Government reports, Department of Education, Skills and Employment ([a]) and Household, Income, Labour Dynamics in Australia survey (HILDA) ([b]).

Table 3: Socioeconomic indicators for first- and non first-generation students

	First-Generation Students (1)	Non First-Generation Students (2)	t-test for difference in means: (1)-(2) (3)
	Mean	Mean	p-value
Source: Administrative data			
Low SES	0.094	0.034	0.002
Medium SES	0.412	0.284	0.002
High SES	0.353	0.582	0.000
International student	0.141	0.099	0.140
Source: Survey data			
Parental encouragement	0.853	0.916	0.020
Lives at home	0.653	0.682	0.497
Receives financial support	0.576	0.711	0.001
Financial support ($), if received	574.95	697.80	0.361

Table 4: Socioeconomic gradient in cognitive and non-cognitive skills

Socioeconomic indicator:	First-generation		Low SES	High SES
	Student survey[a]	HILDA[b]		
Dependent variable:				
Non-cognitive skills				
Conscientiousness	0.043	0.055	0.084	0.040
	(0.093)	(0.054)	(0.224)	(0.093)
Openness	-0.098	-0.187***	-0.246	0.035
	(0.094)	(0.053)	(0.214)	(0.091)
Extraversion	-0.119	-0.035	-0.012	0.199*
	(0.090)	(0.054)	(0.188)	(0.093)
Internality	-0.157*	0.019	0.123	0.051
	(0.094)	(0.059)	(0.189)	(0.097)
Grit	-0.065	NA	-0.003	0.0002
	(0.088)		(0.197)	(0.091)
Cognitive skills				
ATAR/Ability tests	-0.370***	-0.303***	-0.326	0.388***
	(0.101)	(0.064)	(0.236)	(0.096)

Notes: The estimated equation is (1). Robust standard errors are shown in parentheses. *** p<.001, ** p<0.05, *p<0.1. Each row shows the coefficient estimate on the socioeconomic indicator variable from a separate regression with the dependent variable as listed in the left-most column. Each of the dependent variables are standardised to have a mean 0 and variance of 1. Covariates included but not shown here are a gender dummy, a full series of age (in years) fixed effects, a dummy for international students and a dummy variable for whether the student has ever been diagnosed with anxiety. In the regressions with a non-cognitive skill as the dependent variable, we include the ATAR score as an additional covariate.

[a] Student survey: Number of observations is 613 for each of all non-cognitive skills, and 442 for cognitive skills (ATAR).

[b] Household Income and Labour Dynamics in Australia (HILDA) survey: Number of observations: 1,611 for Conscientiousness, Openness to experience and Extraversion; 1,847 for Internality and 1,018 for cognitive ability, which is a summary measure of backward digit span, symbol coding and word knowledge. Skill data in HILDA is available in Waves 3, 4, 5, 7, 9, 11, 12, 15, and 18.

Table 5: First-generation status and Academic Achievement

	Benchmark model: Student performance (WAM) across all four semesters				Value added model
	(1)	(2)	(3)	(4)	(5)
First-generation	-0.268***	-0.250***	-0.161*	-0.168*	-0.015
	(0.095)	(0.092)	(0.090)	(0.089)	(0.047)
Female		0.087	0.152*	0.185**	0.101*
		(0.090)	(0.085)	(0.083)	(0.055)
Anxiety		-0.275***	-0.244**	-0.237**	-0.036
		(0.101)	(0.097)	(0.098)	(0.062)
International student		-0.065	-0.072	-0.089	-0.093
		(0.116)	(0.143)	(0.142)	(0.086)
ATAR score (std)			0.353***	0.328***	0.147***
			(0.045)	(0.044)	(0.027)
Conscientiousness (std)				0.138***	0.081***
				(0.040)	(0.025)
Openness (std)				0.048	0.018
				(0.038)	(0.026)
Extraversion (std)				-0.112***	-0.055***
				(0.037)	(0.021)
Grit (std)				-0.002	0.016
				(0.033)	(0.023)
Internality (std)				0.046	0.014
				(0.038)	(0.024)
Lagged WAM score (in *t-1*)					0.599***
					(0.045)
Constant	-0.005	0.054	-0.082	-0.019	-0.189*
	(0.044)	(0.156)	(0.170)	(0.101)	(0.104)
Observations	2179	2179	2179	2179	1529
No. of students	621	621	621	621	577

Notes: Cluster robust standard errors (clustered at the individual) are in parentheses. *** p<0.001, ** p<0.05, * p<0.1.
The dependent variable is the standardized WAM. In columns (1) to (4) the estimated equation is (4). In column (5) the estimated equation is (5). Additional controls not shown are a series of age (in years) fixed effects and a dummy equal to one if a student is missing an ATAR score.

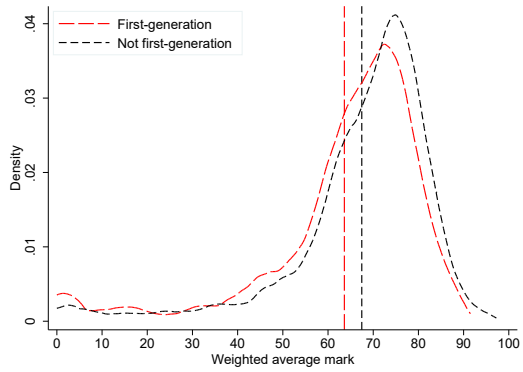
Table 6: First-generation status and Academic Achievement, by semester

	Semester 1	Semester 2	Semester 3	Semester 4
	(1)	(2)	(3)	(4)
First-generation	-0.206**	-0.131	-0.152	-0.041
	(0.098)	(0.098)	(0.099)	(0.094)
Observations	613	566	524	476

Notes: Robust standard errors are in parentheses. *** p<0.001, ** p<0.05, * p<0.1. The dependent variable is the standardized WAM for the given semester. Control variables included but not shown are those also included in the estimated models in column (4) of Table 5.

Figure 1. Weighted Average Marks at university by socioeconomic status indicators

(a) by first-generation



(b) by SES

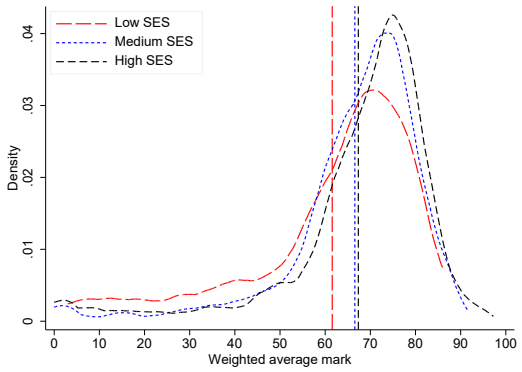
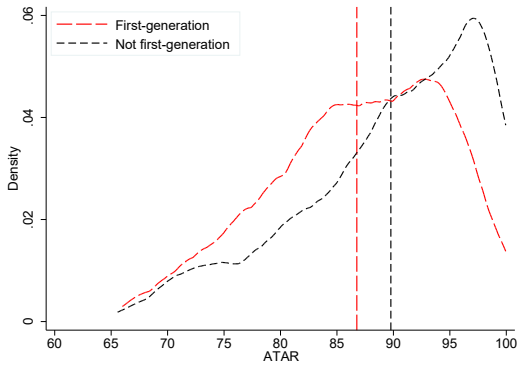


Figure 2. Cognitive skills (ATAR score) by socioeconomic status

(a) by first-generation



(b) by SES

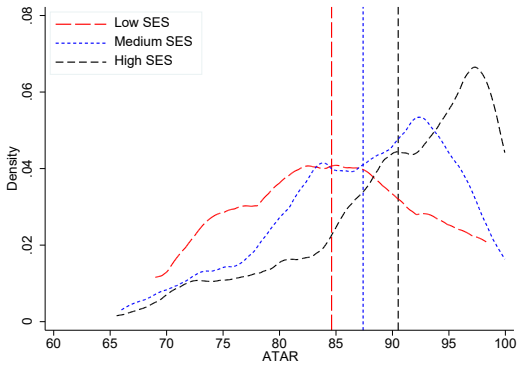
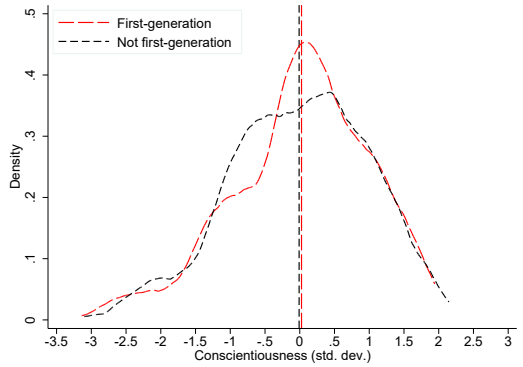


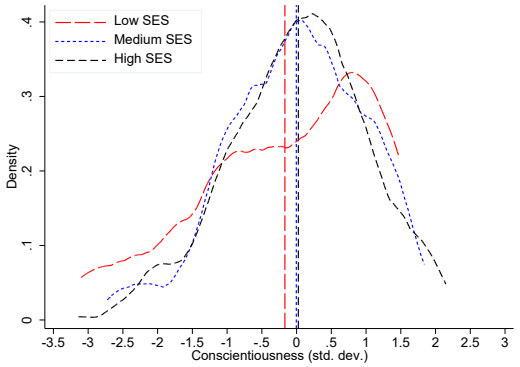
Figure 3. Non-cognitive skills by socioeconomic status

(a) by first-generation

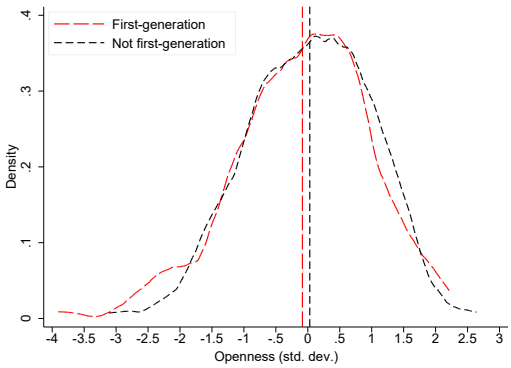
(b) by SES

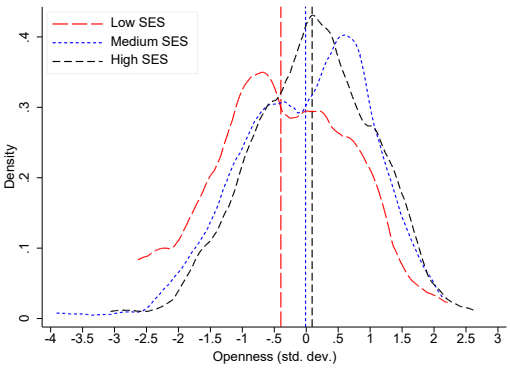
(i) Conscientiousness



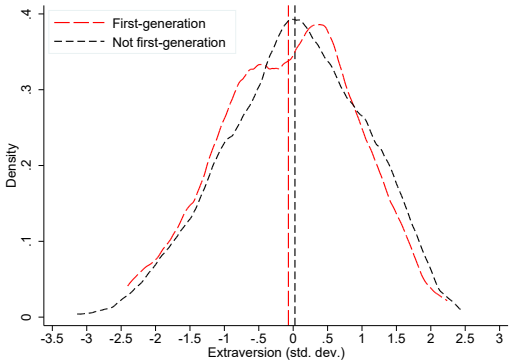


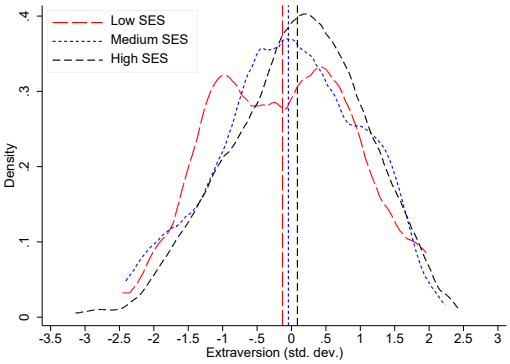
(ii) Openness



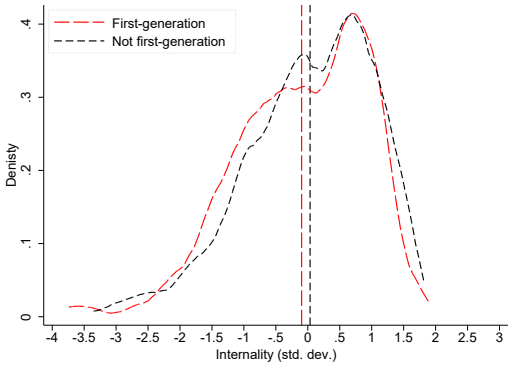


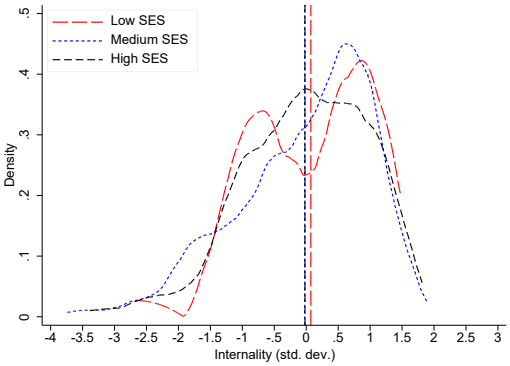
(iii) Extraversion



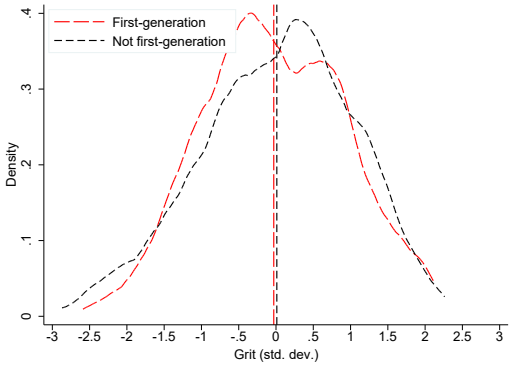


(iv) Internality





(v) Grit



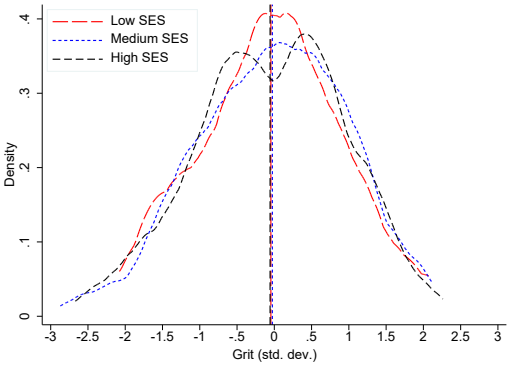
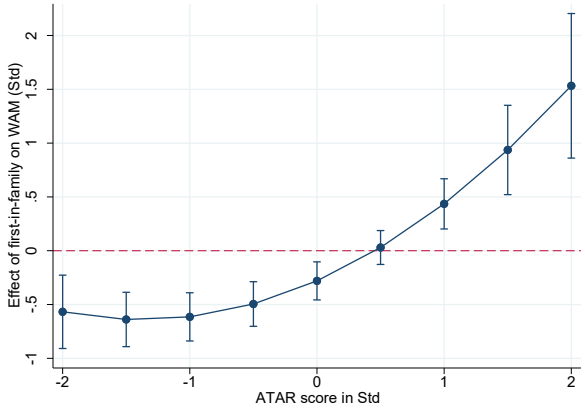
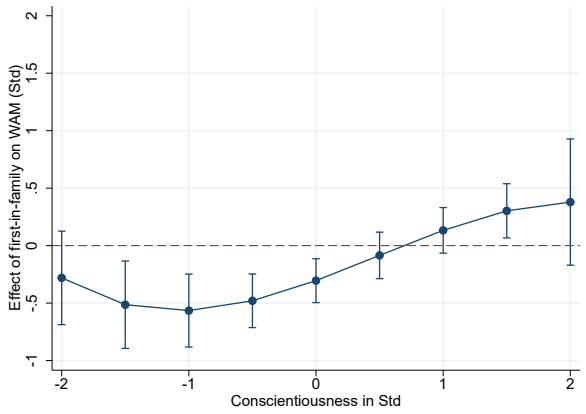


Figure 4. First-generation status and academic achievement across the skill distribution

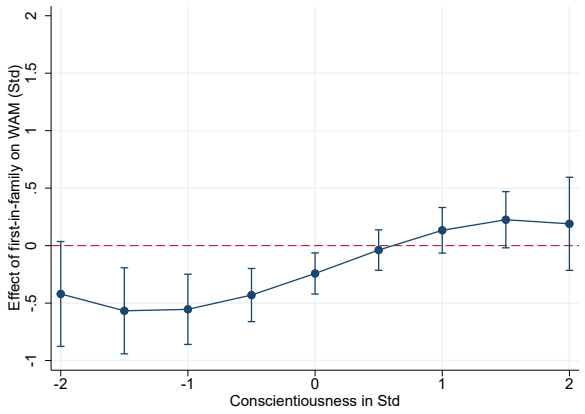
(a) ATAR



(b) Conscientiousness



(c) Conscientiousness vignette adjusted



Notes: Spikes indicate 95% confidence intervals. Figures show the marginal effect associated with being first-generation relative to not at different points across the distribution of each skill. The estimated specification is (6). Estimated coefficients are shown in Table A5 in the Online Appendix. Figures for all skills are shown in Figure A1 in the Online Appendix.

Supplementary Material: Online Appendix

Appendix A

Table A1. Tests for the equality of skill distributions by socioeconomic status

Variable	First-generation and Not first-generation	Low SES and Medium SES	Medium SES and High SES	Low SES and High SES
	p-values			
ATAR	0.000	0.000	0.000	0.000
Conscientiousness	0.005	0.056	0.386	0.012
Openness	0.028	0.000	0.021	0.000
Extraversion	0.012	0.025	0.001	0.001
Internality	0.002	0.088	0.002	0.051
Grit	0.003	0.407	0.140	0.258
Weighted Average Mark	0.000	0.105	0.002	0.001

Notes: p-values shown are from the Kolmogorov Smirnov test for the equality of distributions. The corresponding distributions are shown in Figures 1, 2 and 3 in the main text.

Table A2. Relationship between skills and alternative SES measures

	Regional	School type (reference is public)		Does ***not*** receive financial support from parents	Does ***not*** live with parents	Did not receive parental encouragement
		Catholic	Private/Indep			
Dependent variable:						
Non-cognitive skills						
Conscientious	-0.0001	0.055	0.029	-0.097	-0.194*	-0.185
	(0.185)	(0.122)	(0.108)	(0.094)	(0.111)	(0.140)
Openness	-0.042	0.007	0.087	-0.099	-0.031	-0.019
	(0.160)	(0.127)	(0.114)	(0.094)	(0.107)	(0.159)
Extraversion	-0.083	-0.039	-0.006	-0.013	-0.002	-0.189
	(0.148)	(0.127)	(0.111)	(0.091)	(0.102)	(0.134)
Internality	0.226	0.010	0.007	-0.079	-0.023	-0.177
	(0.139)	(0.127)	(0.114)	(0.093)	(0.102)	(0.140)
Grit	-0.009	0.060	0.160	-0.004	-0.026	0.007
	(0.142)	(0.120)	(0.113)	(0.092)	(0.102)	(0.138)
Cognitive skills						
ATAR	-0.354	-0.078	0.020	-0.285***	-0.070	-0.015
	(0.223)	(0.139)	(0.119)	(0.109)	(0.130)	(0.180)

Notes: The estimated equation is (1). Robust standard errors are shown in parentheses. *** p<.001, ** p<0.05, *p<0.1.
Each row shows the coefficient estimate on the socioeconomic indicator variable from a separate regression with the dependent variable as listed in the left-most column. Each of the dependent variables are standardised to have a mean 0 and variance of 1. Covariates included but not shown here are a gender dummy, a full series of age (in years) fixed effects, a dummy for international students and a dummy variable for whether the student has ever been diagnosed with anxiety. The number of observations is 613 for each of the non-cognitive skills, and there are 442 observations in the regression for cognitive skills, that is, for the ATAR score.

Table A3. First-generation status and Academic Achievement – full estimates of preferred model and robustness to additional controls

		Additional covariate added				
	Preferred model[a]	Semester fixed effects	Has children (yes/no)	Married (yes/no)	Birth order indicator variables	Left-handedness (yes/no)
	(1)	(2)	(3)	(4)	(5)	(6)
First-generation	-0.168*	-0.169*	-0.179**	-0.182**	-0.176**	-0.172*
	(0.089)	(0.089)	(0.088)	(0.089)	(0.089)	(0.089)
Female	0.185**	0.183**	0.185**	0.189**	0.203**	0.200**
	(0.083)	(0.083)	(0.083)	(0.083)	(0.084)	(0.084)
Age – 18 years	-0.074	-0.073	-0.075	-0.070	-0.063	-0.064
	(0.089)	(0.089)	(0.089)	(0.090)	(0.092)	(0.092)
Age – 19 years	0.092	0.091	0.089	0.092	0.119	0.117
	(0.107)	(0.108)	(0.108)	(0.108)	(0.110)	(0.110)
Age – 20 years	-0.109	-0.107	-0.117	-0.119	-0.088	-0.086
	(0.222)	(0.223)	(0.226)	(0.225)	(0.231)	(0.231)
Age – 21 years	-0.192	-0.191	-0.079	-0.076	-0.087	-0.084
	(0.251)	(0.252)	(0.232)	(0.231)	(0.244)	(0.244)
Age – 22 years	-0.532*	-0.534*	-0.538*	-0.565*	-0.548	-0.565*
	(0.321)	(0.320)	(0.321)	(0.319)	(0.335)	(0.339)
Age – 23 years	0.383	0.376	0.373	0.356	0.382	0.395
	(0.415)	(0.417)	(0.413)	(0.399)	(0.411)	(0.413)
Age – 24 years	-0.069	-0.068	-0.076	-0.097	-0.092	-0.076
	(0.436)	(0.437)	(0.437)	(0.435)	(0.444)	(0.444)
Age – 25 years	0.042	0.036	-0.187	-0.207	-0.190	-0.176
	(0.304)	(0.304)	(0.352)	(0.366)	(0.359)	(0.359)
Age – 26 years	-0.837	-0.829	-0.845	-0.858	-0.835	-0.818
	(0.595)	(0.595)	(0.597)	(0.606)	(0.597)	(0.598)
Age – 27 years	0.205	0.187	0.186	0.141	0.155	0.169
	(0.307)	(0.306)	(0.302)	(0.300)	(0.294)	(0.294)
Age – 28 years	-1.254	-1.252	-2.078*	-2.024*	-2.011*	-1.999*
	(1.443)	(1.458)	(1.069)	(1.103)	(1.087)	(1.089)
Age – 29 years	-1.208	-1.209	-1.202	-1.267	-1.205	-1.194
	(0.946)	(0.953)	(0.951)	(0.906)	(0.927)	(0.930)
Anxiety	-0.237**	-0.238**	-0.241**	-0.244**	-0.251**	-0.250**
	(0.098)	(0.098)	(0.098)	(0.098)	(0.099)	(0.098)
International student	-0.089	-0.087	-0.100	-0.097	-0.114	-0.117
	(0.142)	(0.143)	(0.142)	(0.142)	(0.148)	(0.148)
ATAR score (std)	0.328***	0.327***	0.324***	0.323***	0.337***	0.343***
	(0.044)	(0.044)	(0.045)	(0.045)	(0.045)	(0.046)
Missing ATAR score	-0.055	-0.057	-0.039	-0.047	-0.042	-0.046
(dummy)	(0.126)	(0.126)	(0.125)	(0.124)	(0.125)	(0.124)
Conscientiousness	0.138***	0.139***	0.146***	0.146***	0.145***	0.144***
(std)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
Openness (std)	0.048	0.048	0.044	0.046	0.042	0.040
	(0.038)	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)
Extraversion (std)	-0.112***	-0.112***	-0.109***	-0.110***	-0.113***	-0.112***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Grit (std)	-0.002	-0.002	-0.005	-0.007	-0.011	-0.011
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Internality (std)	0.046	0.046	0.038	0.038	0.042	0.041
	(0.038)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)
Semester 2		-0.055*	-0.053*	-0.053*	-0.056*	-0.056*
		(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Semester 3		-0.082**	-0.087**	-0.087**	-0.089**	-0.089**

		(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Semester 4		-0.125***	-0.126***	-0.126***	-0.128***	-0.128***
		(0.042)	(0.042)	(0.042)	(0.043)	(0.043)
Has children			1.607	1.529	1.560	1.561
			(1.062)	(1.101)	(1.082)	(1.083)
Married				0.148	0.175	0.173
				(0.207)	(0.212)	(0.212)
First-born (with siblings)[b]					-0.090	-0.094
					(0.108)	(0.108)
Last-born					0.008	0.002
					(0.113)	(0.113)
Middle child					-0.163	-0.170
					(0.132)	(0.132)
Right-handed for writing						-0.124
						(0.103)
Constant	-0.019	0.043	0.044	0.038	0.075	0.194
	(0.101)	(0.101)	(0.101)	(0.101)	(0.133)	(0.167)
Observations	2179	2179	2163	2144	2144	2144
No. of students	621	621	617	617	612	612

Notes: The dependent variable is the standardized WAM. Cluster robust standard errors are in parentheses. *** p<0.001, ** p<0.05, * p<0.1. [a] In column (1), we reproduce the results from Table 5, column (4) in the main text. In columns (2) to (5), we add additional control variables as shown. [b]The omitted category or reference group is singleton child students.

Table A4. Alternative socioeconomic status indicators and Academic Achievement

	Preferred model^	SES	School type	Lives with parents	Regional	Parental encourage-ement	Parental financial support	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First-generation	-0.168*							
	(0.089)							
Low SES		-0.127						
		(0.192)						
High SES		-0.075						
		(0.082)						
School type: Catholic			-0.012					
			(0.108)					
School type: Private/indep			-0.014					
			(0.094)					
Does not live with parent(s)				-0.291***				
				(0.098)				
Regional					-0.229			
					(0.177)			
Did not receive parental encouragement						0.125		
						(0.123)		
Does not receive financial support from parents							-0.001	-0.019
							(0.084)	(0.085)
Financial support if received ($'0000s)								-0.284
								(0.307)
Female	0.185**	0.187**	0.182**	0.161**	0.177**	0.178**	0.183**	0.176**
	(0.083)	(0.083)	(0.083)	(0.082)	(0.083)	(0.083)	(0.083)	(0.084)
Anxiety	-0.237**	-0.214**	-0.226**	-0.195**	-0.225**	-0.228**	-0.227**	-0.220**
	(0.098)	(0.098)	(0.097)	(0.096)	(0.098)	(0.097)	(0.097)	(0.098)
International student	-0.089	-0.157	-0.110	0.031	-0.131	-0.109	-0.108	-0.076
	(0.142)	(0.151)	(0.169)	(0.153)	(0.142)	(0.144)	(0.143)	(0.146)
ATAR score (std)	0.328***	0.347***	0.340***	0.337***	0.335***	0.341***	0.340***	0.341***
	(0.044)	(0.046)	(0.044)	(0.044)	(0.045)	(0.044)	(0.044)	(0.044)
Conscientiousness (std)	0.138***	0.134***	0.134***	0.122***	0.134***	0.136***	0.134***	0.136***
	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)	(0.040)	(0.040)	(0.040)
Openness (std)	0.048	0.050	0.052	0.054	0.051	0.051	0.052	0.051
	(0.038)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Extraversion (std)	-0.112***	-0.106***	-0.109***	-0.107***	-0.110***	-0.107***	-0.109***	-0.109***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Grit (std)	-0.002	-0.001	-0.000	-0.000	-0.002	-0.002	-0.001	-0.001
	(0.033)	(0.033)	(0.033)	(0.032)	(0.033)	(0.033)	(0.033)	(0.033)
Internality (std)	0.046	0.052	0.051	0.051	0.055	0.053	0.051	0.054
	(0.038)	(0.039)	(0.039)	(0.039)	(0.038)	(0.039)	(0.039)	(0.039)
Constant	-0.019	-0.019	-0.053	-0.012	-0.051	-0.065	-0.059	-0.044
	(0.101)	(0.101)	(0.116)	(0.099)	(0.101)	(0.101)	(0.100)	(0.101)
Observations	2179	2179	2179	2179	2179	2179	2179	2179
No. of students	621	621	621	621	621	621	621	621

Notes: The dependent variable is the standardized WAM. Cluster robust standard errors are in parentheses. *** p<0.001, ** p<0.05, * p<0.1. ^In column (1), we reproduce the results from Table 5, column (4) in the main text. In columns (2) to (8), we use alternative indicators of socioeconomic status as shown. Additional controls not shown are a series of age (in years) fixed effects and a dummy equal to one if a student is missing an ATAR score.

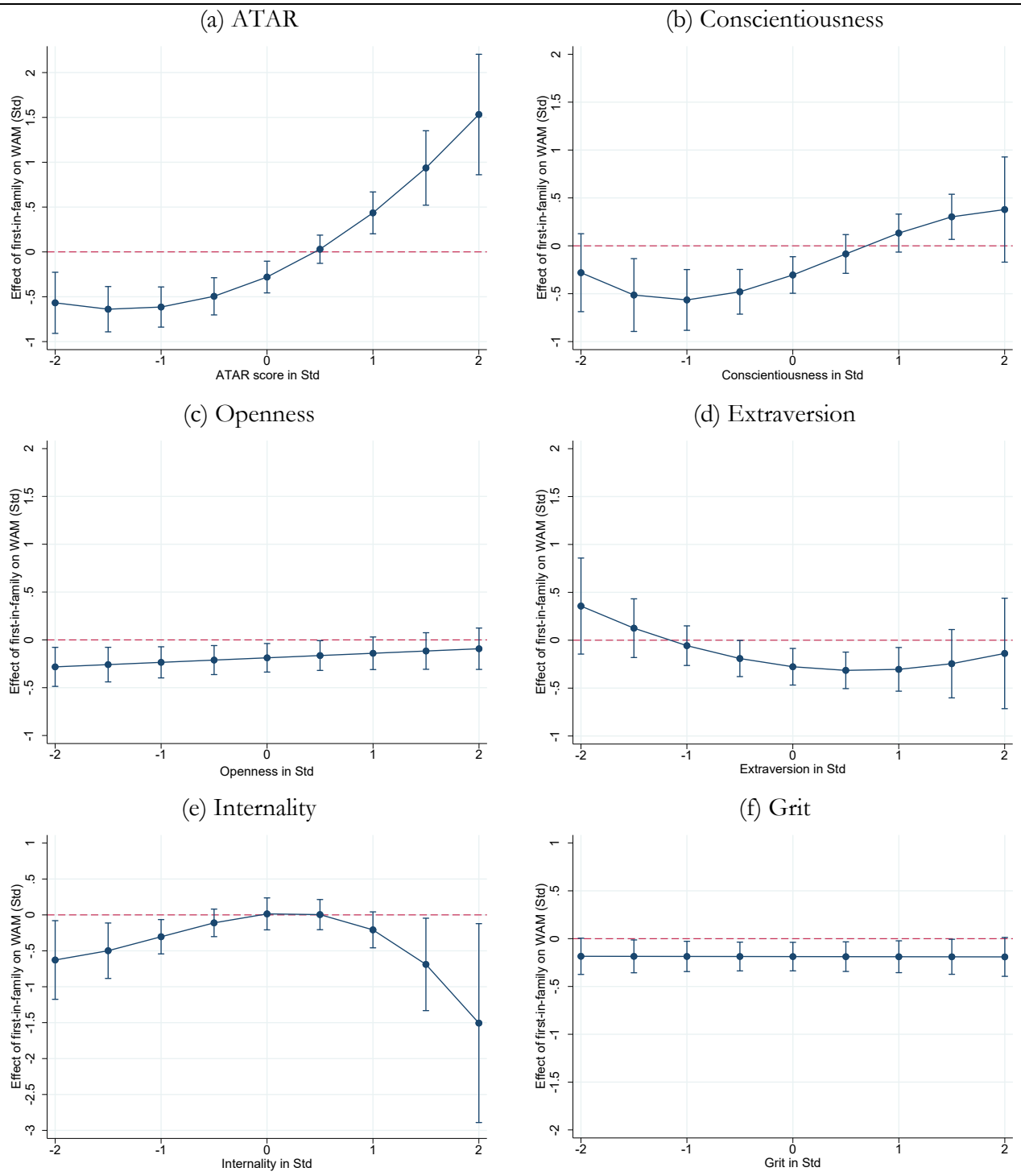
Table A5. Interaction effects between first-generation and skill measure

	(1) ATAR	(2) Consc	(3) Open	(4) Extra	(5) LoC	(6) Grit
First-generation	-0.1669*	-0.1669*	-0.1669*	-0.1669*	-0.1669*	-0.1669*
	(0.088)	(0.088)	(0.088)	(0.088)	(0.088)	(0.088)
Skill	0.3268***	0.1373***	0.0474	-0.1120***	0.0449	-0.0015
	(0.044)	(0.039)	(0.038)	(0.036)	(0.038)	(0.033)
BIC	5178.1474	5178.1474	5178.1474	5178.1474	5178.1474	5178.1474
AIC	5030.2953	5030.2953	5030.2953	5030.2953	5030.2953	5030.2953
	(1)	(2)	(3)	(4)	(5)	(6)
First-generation	-0.1651*	-0.1659*	-0.1701*	-0.1695*	-0.1666*	-0.1648*
	(0.086)	(0.087)	(0.088)	(0.088)	(0.088)	(0.088)
Skill	0.3235***	0.1177***	0.0610	-0.1007**	0.0435	-0.0217
	(0.052)	(0.043)	(0.041)	(0.040)	(0.039)	(0.035)
First-gen × skill	0.0133	0.0721	-0.0453	-0.0450	0.0050	0.0829
	(0.097)	(0.088)	(0.084)	(0.089)	(0.100)	(0.089)
BIC	5185.8163	5185.0221	5185.4863	5185.5363	5185.8300	5184.8354
AIC	5032.2776	5031.4833	5031.9475	5031.9975	5032.2912	5031.2967
	(1)	(2)	(3)	(4)	(5)	(6)
First-generation	-0.1315	-0.2735**	-0.1934**	-0.1884*	-0.1296	-0.1050
	(0.105)	(0.109)	(0.098)	(0.110)	(0.102)	(0.111)
Skill	0.4361***	0.1318***	0.0617	-0.0917**	0.0357	-0.0160
	(0.049)	(0.043)	(0.041)	(0.038)	(0.043)	(0.037)
First-gen × skill	0.0889	0.1332	-0.0261	-0.0320	-0.0336	0.0724
	(0.106)	(0.084)	(0.094)	(0.089)	(0.113)	(0.091)
Skill × skill	0.1678***	0.0517	0.0121	0.0670**	-0.0111	0.0220
	(0.042)	(0.034)	(0.028)	(0.027)	(0.034)	(0.028)
First-gen × skill × skill	0.0228	0.1052*	0.0208	0.0295	-0.0392	-0.0638
	(0.071)	(0.054)	(0.047)	(0.072)	(0.076)	(0.078)
BIC	5181.5187	5188.4389	5200.0772	5194.5112	5200.0134	5199.3039
AIC	5016.6066	5023.5269	5035.1652	5029.5992	5035.1014	5034.3919
	(1)	(2)	(3)	(4)	(5)	(6)
First-generation	-0.1119	-0.2528**	-0.1731	-0.2022*	-0.0193	-0.0829
	(0.114)	(0.110)	(0.105)	(0.110)	(0.126)	(0.110)
Skill	0.3640***	0.2295***	0.0291	0.0235	0.1246*	0.0032
	(0.090)	(0.074)	(0.062)	(0.072)	(0.070)	(0.064)

First-gen × skill	0.1705	0.1809	0.0390	-0.1962	0.0126	0.2506*
	(0.165)	(0.157)	(0.129)	(0.141)	(0.124)	(0.150)
Skill × skill	0.2208***	0.0186	0.0196	0.0461	-0.0813*	0.0180
	(0.058)	(0.038)	(0.031)	(0.030)	(0.042)	(0.030)
First-gen × skill × skill	-0.0388	0.0698	-0.0054	0.0600	-0.1890	-0.0888
	(0.133)	(0.070)	(0.071)	(0.071)	(0.140)	(0.079)
Skill × skill × skill	0.0398	-0.0412	0.0130	-0.0451**	-0.0431*	-0.0066
	(0.036)	(0.026)	(0.017)	(0.021)	(0.024)	(0.019)
First-gen × s × s × s	-0.0470	-0.0202	-0.0238	0.0673	-0.0461	-0.0810*
	(0.073)	(0.046)	(0.028)	(0.046)	(0.047)	(0.049)
BIC	5195.9358	5198.4345	5214.8480	5206.1362	5203.4109	5211.2511
AIC	5019.6506	5022.1493	5038.5627	5029.8509	5027.1256	5034.9658
	(1)	(2)	(3)	(4)	(5)	(6)
First-generation	-0.1093	-0.2851**	-0.2100*	-0.2750**	-0.0227	-0.0657
	(0.115)	(0.124)	(0.113)	(0.137)	(0.127)	(0.129)
Skill	0.3613***	0.1988***	0.0096	-0.0234	0.1811**	0.0370
	(0.138)	(0.074)	(0.064)	(0.075)	(0.087)	(0.072)
First-gen × skill	0.0744	0.2277	0.0935	-0.1377	0.0848	0.2318
	(0.265)	(0.170)	(0.162)	(0.148)	(0.194)	(0.167)
Skill × skill	0.2189**	-0.0428	-0.0379	-0.0718	-0.0417	0.0973
	(0.088)	(0.086)	(0.062)	(0.069)	(0.066)	(0.075)
First-gen × skill × skill	-0.0645	0.1557	0.0746	0.2412	-0.1875	-0.1273
	(0.154)	(0.154)	(0.106)	(0.194)	(0.153)	(0.175)
Skill × skill × skill	0.0415	-0.0231	0.0244	-0.0196	-0.0795**	-0.0252
	(0.083)	(0.031)	(0.020)	(0.023)	(0.040)	(0.024)
First-gen × s × s × s	0.0380	-0.0490	-0.0500	0.0346	-0.0913	-0.0704
	(0.184)	(0.062)	(0.051)	(0.049)	(0.124)	(0.063)
Skill × s × s × s	0.0008	0.0161	0.0118	0.0275*	-0.0153	-0.0194
	(0.032)	(0.020)	(0.009)	(0.015)	(0.018)	(0.015)
First-gen × s × s × s × s	0.0327	-0.0225	-0.0171	-0.0433	-0.0088	0.0089
	(0.065)	(0.033)	(0.016)	(0.039)	(0.037)	(0.037)
BIC	5211.0230	5213.1114	5229.2457	5218.0601	5216.7633	5225.4398
AIC	5023.3645	5025.4529	5041.5872	5030.4016	5029.1048	5037.7813
Observations	2179	2179	2179	2179	2179	2179

Notes: The dependent variable is the standardized WAM. Cluster robust standard errors are in parentheses. *** p<0.001, ** p<0.05, * p<0.1. The estimated specification in the uppermost panel is (4). In subsequent panels, the specification is (6) – each panel shows the results as an additional interaction between the particular skill (indicated in the column headings) and the first-generation indicator variable.

Figure A1 – First-generation across the skill distribution – all skills



Notes: Spikes indicate 95% confidence intervals. Figures show the marginal effect associated with being first-generation relative to not at different points across the distribution of each skill. The estimated specification is (6). Estimated coefficients are shown in Table A5 in the Online Appendix. The degree of the polynomial in the given skill and interacted with the first generation indicator variable is chosen according to the AIC.

Appendix B

We compare our estimates against data sourced from the Household, Income and Labour Dynamics in Australia (HILDA) survey, a nationally representative survey on Australians of all ages. Since 2001 (wave 1), it collects information on economic and personal wellbeing in households across Australia (Melbourne Institute 2017). We restricted the HILDA sample to individuals who are currently enrolled in a Bachelor degree across all states in Australia, using information from waves 3, 4, 5, 7, 9, 11, 12, 15, and 18, the waves when noncognitive and cognitive skills were collected.

Table B1. Relationship between personality traits and SES:
HILDA Waves 3-18, Ages 16-19

	(1) Extra	(2) Agree	(3) Consc	(4) Emote	(5) Open	(6) LOC
Pooled male & female sample						
First-generation	-0.466***	-0.460***	-0.428***	-0.462***	-0.601***	-0.0361
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.04)
Observations	2749	2749	2749	2749	2749	2945
Male sample						
First-generation	-0.202	-0.213	-0.204	-0.220	-0.378*	-0.0753
	(0.23)	(0.24)	(0.23)	(0.24)	(0.23)	(0.06)
Observations	1308	1308	1308	1308	1308	1401
Female sample						
First-generation	-0.707***	-0.691***	-0.636***	-0.688***	-0.805***	-0.001
	(0.19)	(0.20)	(0.19)	(0.19)	(0.19)	(0.06)
Observations	1441	1441	1441	1441	1441	1544

Notes. Cluster robust standard errors in parentheses. * $p < .10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix C. Vignettes

Participants completed 13 self-report items designed to elicit measures of two Five Factor Model (FFM) personality traits: conscientiousness (C) and openness to experience (O). The Big Five 'trait descriptive adjective' (TDA) marker method upon which our measurement strategy is based was originally developed by Goldberg (1992) and a shorter version (the 'Mini-Markers'), was developed and validated by Saucier (1994). We utilise an adapted subset of the Mini-Markers based on Losoncz (2009) and Cobb-Clark & Schurer (2012), in which C and O are indexed by seven and six TDA items respectively. Participants indicate the degree to which each of the 13 adjectives describes them on a seven-point response scale, ranging from 1 ("Does not describe me at all") to 7 ("Describes me very well"). The items, in order of presentation, are: Orderly (C), Philosophical (O), Systematic (C), Inefficient* (C), Creative (O), Sloppy* (C), Intellectual (O), Disorganised* (C), Complex (O), Imaginative (O), Efficient (C), Careless* (C), and Deep (O); asterisks identify items that require reverse-coding and letters in brackets indicate the trait for which the item loads most strongly.

Participants are then asked to respond to a set of anchoring vignettes, a strategy designed to allow for the correction of bias induced by heterogeneity in the interpretation of response categories (termed response category differential item function; DIF). The survey instrument is shown below. The task asks the participant to read and assess three vignettes, randomly selected from the eight possible vignettes (listed below). Each consists of a brief sketch depicting a hypothetical third person whose gender is randomised for each scenario. The traits of the hypothetical characters are designed to align with various levels of the traits loading onto C and O. The participant is to assess each character using the same items and response scale as they did to rate themselves. In Table C1 below we provide the classification of each vignette as high/low on C and O.

The vignettes enable the construction of a common response scale across participants and estimation of the measurement error resulting from DIF. The multidimensional item response theory method proposed by King et al. (2004, 2007) and Bolt et al. (2014) is utilised to incorporate the anchoring vignettes as indicators of an individual's response style.

In Table C2 we present the results of our individual fixed effects regression used to extract the adjusted self-assessed conscientiousness and openness to experience score as used in Figure 4, Panel C. To extract the vignette adjusted C and O, we use the responses to the 3 vignettes for each non-cognitive skill C and O and the self-assessed personality scores for C and O, giving up to 8 reports per individual. We use an individual fixed effects regression flexibly controlling for the vignettes the respondent faced in the survey and the vignette and respondent gender. The adjusted C and O scores for each survey respondent are then constructed using the residuals from this regression.

The distributions of the adjusted self-assessed C and O scores are shown relative to the unadjusted scores in Figure C1 for all students and for the sub-sample of first-in-family students. The distribution of C and O scores for first-generation and non-first-generation students is shown in Figure C2. Tables C3 and C4 present our robustness checks of our main results using the adjusted C and O scores.

Survey instruments

Below (C and C1) are the vignette survey instruments presented to respondents. Immediately following, we include the full text of the vignettes, describing the eight hypothetical people.

C Your perceptions about others

Below you will find descriptions of the behaviour of three people. Please rate each person's personality similarly to how you have rated yourself in Part A.

[Note: Survey participants are presented with three randomly chosen sketches from the eight sketches listed below]

C1 How well do the following words describe [Name]. For each word, cross one box to indicate how well that word describes [Name]. There are no right or wrong answers.

Does not describe [Name] at all						Describes [Name] very well
1	2	3	4	5	6	7
Tick X one box for each word (a number for each name and adjective)						

Adjectives	**Name 1**	**Name 2**	**Name 3**
C1a Orderly			
C1b Philosophical			
C1c Systematic			
C1d Inefficient			
C1e Creative			
C1f Sloppy			
C1g Intellectual			
C1h Disorganized			
C1i Complex			
C1j Imaginative			
C1k Efficient			
C1m Deep			

Hypothetical person sketches

1. [Mary] runs a company she founded on her own, raises three children and takes care of her household meticulously. In addition, she is active in sports and in community life. Despite her wide range of activities, she has time for her parents and to go hiking with friends. She likes reading and discussing philosophy and experimenting with new foods.
2. Already as a child [Anette] wanted to become a doctor. At school she was a moderate student lacking depth and creativity and her teachers did not believe she would be admitted to university. She did not succeed the first time, but [Anette] did not give up, she worked as an orderly at a hospital for a year, took private lessons and at second attempt she was admitted to university. Presently [Anette] is a registered doctor and the manager of a small practice.
3. [Nancy] discontinued her studies and she hasn't been able to find a steady job for 10 years. She lives with her parents, who have difficulty coping financially. Due to being

overweight [Nancy] has tried many diets unsuccessfully, she now has heart problems and doctors have advised her to be physically active. In spite of this [Nancy] seldom leaves the house and most of the day she watches TV.

4. Generally [Allan's] friends trust him and enjoy his company because of his ability to think deep and see things from different perspectives. Sometimes, however, they have been really annoyed by him. For example, [Allan] does not always return the things he has borrowed on time. Sometimes he completely forgets about his promises.
5. Five years ago [Tom] finished his medical studies at the university and started working as a surgeon in a local hospital. His colleagues consider him a very good surgeon and lately he was appointed department head in the hospital. In case of problems [Tom] is very dependable. According to [Tom's] wife and her friends, who work as artists and graphic designers, he lacks creativity and rarely tries out new experiences
6. Since childhood [Bruno] has wanted to achieve a lot in his life and he has worked a lot for it. Despite extreme poverty at his parental home [Bruno] managed to get a good education. Continuous self-education and long hours at work have made him a very valued specialist and he has received ever better job offers. [Bruno] enjoys reading different newspapers to broaden his views.
7. [Jeanette] is a very creative young girl. She loves reading and writing, and taking her own time to develop her thoughts. She has been a member of a writer's club for many years, and has written several short stories. [Jeanette] is good in school, but she often daydreams during class, arrives late, and has difficulty meeting deadlines.
8. [Gerry] used to be a handsome man and competitive tennis player in his early 20s. Now in his late 30s he watches a lot of TV and enjoys a drink with his friends, although he doesn't like meeting new people. He works as a key account manager of a large wealth management firm. [Gerry] is reliable in his day-to-day job duties, but does not take the initiative to improve his performance or learn new things.

Hypothetical person sketches with reverse gender

1. [Mark] runs a company he founded on his own, raises three children and takes care of his household meticulously. In addition, he is active in sports and in community life. Despite his wide range of activities, he has time for his parents and to go hiking with friends. He likes reading and discussing philosophy and experimenting with new foods.
2. Already as a child [Adam] wanted to become a doctor. At school he was a moderate student lacking depth and creativity and his teachers did not believe he would be admitted to university. He did not succeed the first time, but [Adam] did not give up, he worked as an orderly at a hospital for a year, took private lessons and at second attempt he was admitted to university. Presently [Adam] is a registered doctor and the manager of a small practice.
3. [Nick] discontinued his studies and he hasn't been able to find a steady job for 10 years. He lives with his parents, who have difficulty coping financially. Due to being overweight [Nick] has tried many diets unsuccessfully, he now has heart problems and doctors have advised him to be physically active. In spite of this [Nick] seldom leaves the house and most of the day he watches TV.
4. Generally [Amy's] friends trust her and enjoy her company because of her ability to think deep and see things from different perspectives. Sometimes, however, they have been really annoyed by her. For example, [Amy] does not always return the things she has borrowed on time. Sometimes she completely forgets about her promises.

5. Five years ago [Tina] finished her medical studies at the university and started working as a surgeon in a local hospital. Her colleagues consider her a very good surgeon and lately she was appointed department head in the hospital. In case of problems [Tina] is very dependable. According to [Tina's] husband and his friends, who work as artists and graphic designers, she lacks creativity and rarely tries out new experiences
6. Since childhood [Beth] has wanted to achieve a lot in her life and she has worked a lot for it. Despite extreme poverty at her parental home [Beth] managed to get a good education. Continuous self-education and long hours at work have made her a very valued specialist and she has received ever better job offers. [Beth] enjoys reading different newspapers to broaden her views.
7. [Jim] is a very creative young boy. He loves reading and writing, and taking his own time to develop his thoughts. He has been a member of a writer's club for many years, and has written several short stories. [Jim] is good in school, but he often daydreams during class, arrives late, and has difficulty meeting deadlines.
8. [Gwyneth] used to be a beautiful woman and competitive tennis player in her early 20s. Now in her late 30s she watches a lot of TV and enjoys a drink with her friends, although she doesn't like meeting new people. She works as a key account manager of a large wealth management firm. [Gwyneth] is reliable in her day-to-day job duties, but does not take the initiative to improve her performance or learn new things.

Table C1. Classification of personality traits in 16 vignettes

Name	Conscientiousness	Openness to experience
Mary, Mark	High	High
Annette, Adam	High	Low
Nancy, Nick	Low*	Low
Allan, Amy	Low	High
Tom, Tina	High	Low
Bruno, Beth	High	High
Jeannette, Jim	Low	High
Gerry, Gwyneth	Low*	Low

Table describes the orientation of the description of the fictive personality. * indicates some ambiguity in the description of the vignette.

Table C2. Individual fixed effects model to purge response heterogeneity from self-assessment

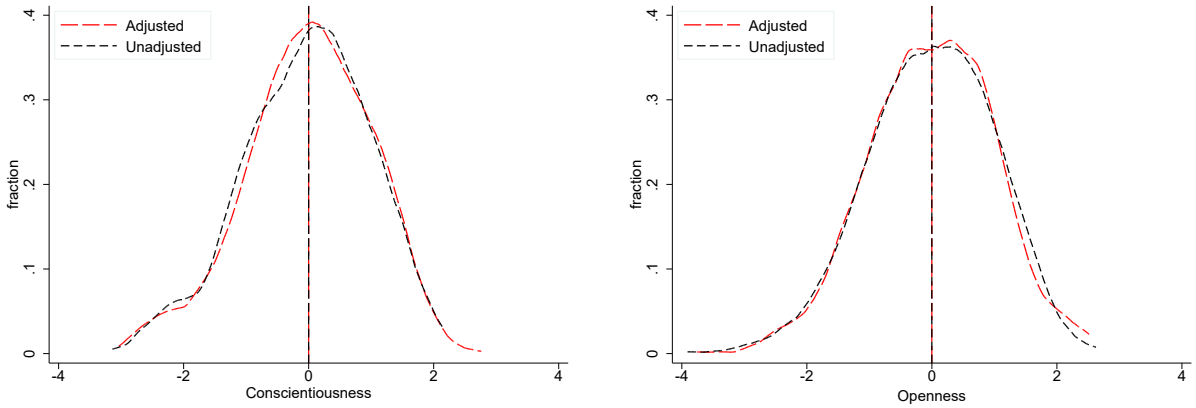
	Personality trait score
Conscientiousness	0.0005
	(0.025)
Self-assessed	-0.899***
	(0.084)
Annette	-0.733***
	(0.108)
Nancy	-1.993***
	(0.100)
Allan	-0.965***
	(0.116)
Tom	-0.715***
	(0.121)
Bruno	-0.291**
	(0.130)
Jeanette	-0.728***
	(0.095)
Gerry	-1.426***
	(0.119)
Mark	-0.033
	(0.127)
Adam	-0.852***
	(0.126)
Nick	-2.029***
	(0.118)
Amy	-0.878***
	(0.099)
Tina	-0.618***
	(0.101)
Beth	-0.252**
	(0.113)
Jim	-0.874***
	(0.129)
Gwyneth	-1.261***
	(0.096)
Female vignette	-0.098
	(0.065)
Constant	0.982***
	(0.102)
N	4961
Individuals	663
Number of obs per individual	
Minimum	2
Average	7.5
Maximum	8
Fraction of variance due to individual fixed effect	0.170
Explained within variation	0.227
Explained between variation	0.142

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Openness is the reference trait, and Nancy (with the highest O and C scores) is the reference vignette.

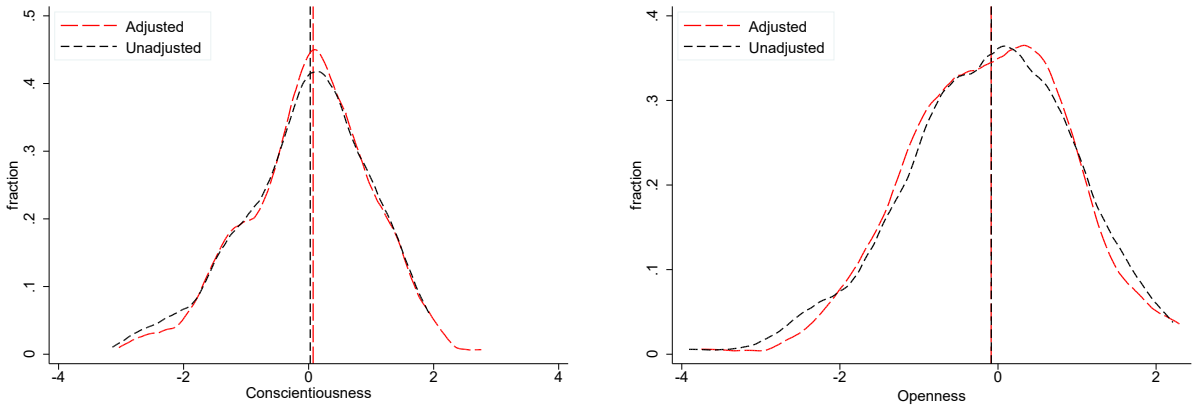
Figure C1 Distribution of personality trait, original and adjusted

(a) Conscientiousness (b) Openness to experience

(i)All students



(ii) First-generation students

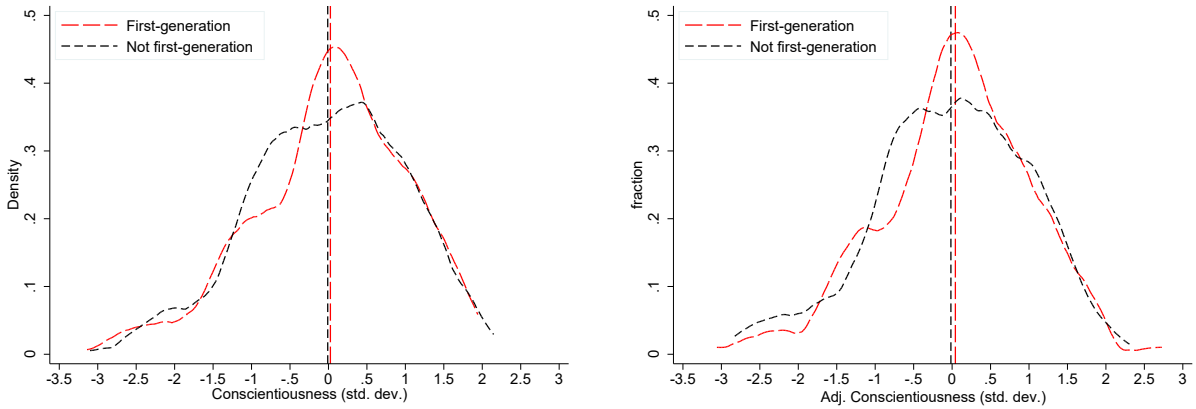


Notes: p-values from the Kolmogorov Smirnov test for the equality of distributions show no statistically significant difference between the adjusted and unadjusted distributions for all students or for the sub-sample of first-generation students. Results available from the authors on request.

Figure C2. Non-cognitive skills by socioeconomic status, vignette-adjusted

(a) Original (b) Vignette adjusted

(i) Conscientiousness



(ii) Openness

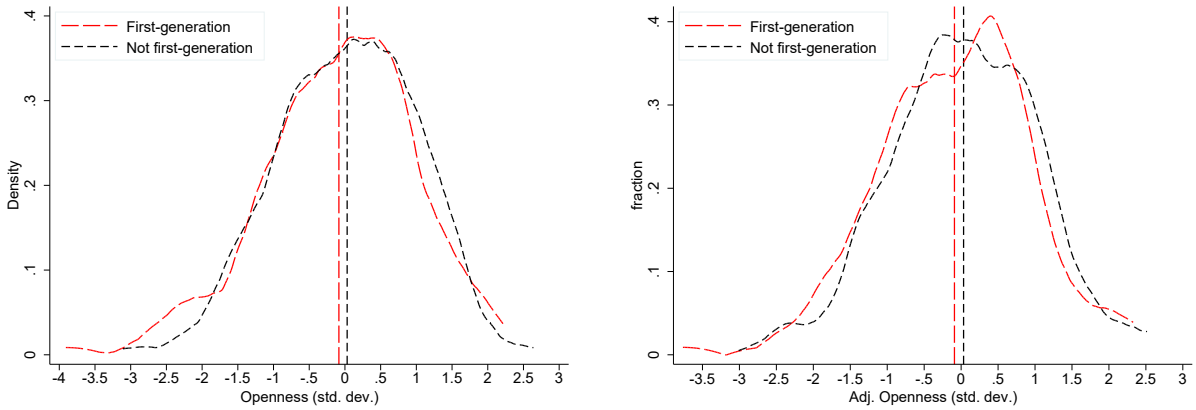


Table C3. Pre-university skills, vignette adjusted

	(1)	(2)	(3)	(4)
	Conscientiousness (Std)		Openness to experience (Std)	
	Original	Adjusted	Original	Adjusted
First-generation	0.043	0.080	-0.098	-0.089
	(0.093)	(0.092)	(0.094)	(0.092)
Female	-0.002	-0.018	-0.123	-0.000
	(0.094)	(0.094)	(0.091)	(0.092)
Ever diagnosed with anxiety	-0.177*	-0.227**	0.233**	0.262**
	(0.103)	(0.103)	(0.100)	(0.103)
International student	-0.024	0.069	-0.056	-0.018
	(0.162)	(0.150)	(0.164)	(0.162)
ATAR score (std)	0.095**	0.094*	0.167***	0.183***
	(0.048)	(0.049)	(0.049)	(0.051)
Constant	-0.062	0.096	-0.135	-0.161
	(0.206)	(0.200)	(0.197)	(0.194)
Observations	613	613	613	613

Notes: Robust standard errors in parentheses. *** p<.001, ** p<0.05, *p<0.1. The estimated equation is (1). Each column shows the coefficient estimates from a separate regression with the dependent variable as listed at the top of the column. Each of the dependent variables are standardised to have a mean 0 and variance of 1. Covariates included but not shown here are a full series of age (in years) fixed effects, and a dummy variable for whether the student has a missing ATAR score.

Table C4. First-generation status and Academic Achievement – robustness to adjusted Conscientiousness and Openness to Experience

	(1) Original	(2) Adjusted
First in family	-0.168*	-0.174*
	(0.089)	(0.089)
Cognitive skills (ATAR, std)	0.328***	0.320***
	(0.044)	(0.045)
Non-cognitive skills		
Conscientiousness (Std)	0.138***	0.161***
	(0.040)	(0.037)
Openness to experience (Std)	0.048	0.074**
	(0.038)	(0.037)
Extraversion (Std)	-0.112***	-0.110***
	(0.037)	(0.037)
Grit (Std)	-0.002	-0.006
	(0.033)	(0.033)
Internal locus of control (Std)	0.046	0.046
	(0.038)	(0.038)
Female	0.185**	0.179**
	(0.083)	(0.082)
Ever diagnosed with anxiety	-0.237**	-0.234**
	(0.098)	(0.098)
International student	-0.089	-0.103
	(0.142)	(0.141)
Constant	-0.107	-0.122
	(0.170)	(0.169)
Observations	2179	2179

Notes: Cluster robust standard errors (clustered at the individual) are in parentheses. *** p<0.001, ** p<0.05, * p<0.1. The dependent variable is the standardized WAM. In column (1) the estimated equation is (4) – this reproduces the results in column (4) of Table 5 in the main text. In column (2) the estimated equation is (4) but the vignette adjusted conscientiousness and openness to experience measures are used. Additional controls not shown are a series of age (in years) fixed effects and a dummy equal to one if a student is missing an ATAR score.