Quantitative Economics 11 (2020), 231–275 1759-7331/20200231

A dynamic model of personality, schooling, and occupational choice

Petra E. Todd
Department of Economics, University of Pennsylvania

Weilong Zhang
Faculty of Economics, University of Cambridge

This paper develops a dynamic model of schooling and occupational choices that incorporates personality traits, as measured by the “big five” traits. The model is estimated using the HILDA dataset from Australia. Personality traits are found to play an important role in explaining education and occupation choices over the lifecycle. Results show that individuals with a comparative advantage in schooling and white-collar work have, on average, higher cognitive skills and higher personality trait scores. Allowing personality traits to evolve with age and with schooling proves to be important to capturing the heterogeneity in how people respond to educational policies. The estimated model is used to evaluate two education policies: compulsory senior secondary school and a 50% college tuition subsidy. Both policies increase educational attainment and also affect personality traits.

Keywords. Personality traits and education policies, occupational choice, unobserved types, human capital investment, dynamic discrete choice.

JEL classification. D61, I26, J24.

Petra E. Todd: ptodd@econ.upenn.edu
Weilong Zhang: weilong.zhang@econ.cam.ac.uk
Petra Todd is Edmund J. and Louise W. Kahn Term Professor of Economics at the University of Pennsylvania and a member of the U Penn Population Studies Group, NBER, HCEO and IZA. Weilong Zhang is an Assistant Professor at University of Cambridge. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to either DSS or the Melbourne Institute. We have received very helpful comments from the co-editor and two anonymous referees. This paper was also presented at the Washington University Graduate Student Conference, the summer school on socioeconomic inequality at Univ of Chicago, the ZEW summer workshop, the 2017 China meeting of the Econometric Society, SOLE 2018, 2018 Cowles foundation conference at Yale, UCL, Univ of Rochester, Toulouse School of Economics, and at a workshop on the economics of education and expectations at Royal Holloway University of London. We would like to thank Lawrence Blume, Jeffery Smith, Flavio Cunha, James Heckman, Hanming Fang, Christopher Flinn, David Wiczer, Jerome Adda, Andrew Shephard, Holger Sieg, Stephanie Schurer, Ariel Pakes, Richard Blundell and Thierry Magnac for helpful comments. The authors acknowledge research funding from the Pension Research Council/Boettner Center for Pensions and Retirement Research at the Wharton School and the Population Studies Center of the University of Pennsylvania. Opinions and views are solely the responsibility of the authors.

Available at http://qeconomics.org. https://doi.org/10.3982/QE890

1. Introduction

Cognitive skills are known to be important determinants of labor market success, but there is increasing evidence that noncognitive skills also play a salient role (Becker (1964) and Griliches (1977)). For example, using data from the Perry Preschool randomized experiment, Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) found that the ability to plan and to exert self-control significantly affects lifetime earnings and employment. Devising social policies that maximize the potential for human development requires an understanding of how cognitive and noncognitive skills jointly evolve and influence individuals' education and labor market trajectories.

This paper develops and estimates a discrete choice dynamic programming (DCDP) model of schooling, work, and occupational choices that incorporates both cognitive and noncognitive skills, where the latter are measured by the "big five" personality traits. The model allows both cognitive and noncognitive traits to influence educational and labor market outcomes, by affecting pecuniary or nonpecuniary returns from schooling and by affecting the reward from choosing different occupational sectors. Our analysis is inspired in part by the pioneering work of Keane and Wolpin (1997) that estimates a similar model without personality traits.

A key finding of Keane and Wolpin (1997) is that 90% of the total variance in expected lifetime utility is explained by unobserved skill endowments at age 16. Other studies also emphasize the importance of unobserved endowment heterogeneity. For example, Yamaguchi (2012) found that endowment differences prior to labor market entry account for 70% of the log-wage variance in the first year and 35% after 20 years. Sullivan (2010) found that 56% of the variance in discounted expected lifetime utility is explained by initial heterogeneity. Huggett, Ventura, and Yaron (2011) concluded that 61.5% of the variation in lifetime earnings and 64% of the variation in lifetime utility are attributable to initial conditions.

Although accumulated evidence shows that endowment heterogeneity is important in explaining educational and labor market outcomes, its precise components remain unclear. Keane and Wolpin (1997) found that family background accounts for less than 10% of the total variation in lifetime utility and that adding cognitive ability as an initial condition only increases the explained variation to 14%. Prior studies have not considered the potential role of personality traits as a component of endowment heterogeneity, because the datasets typically used in estimation do not include personality trait measurements.

In the psychology literature, personality traits have been shown to be correlated with many aspects of individuals' lives, but study of their effects on economic outcomes is relatively scarce (e.g., Almlund, Duckworth, Heckman, and Kautz (2011)). The five-factor model (the so-called "big-five") is the most widely adopted measurement of personality in psychology (Goldberg (1992), Saucier (1994), Gosling, Rentfrow, and Swann (2003)). The big five traits include openness to experience, conscientiousness, extraversion, agreeableness and neuroticism (OCEAN). Below we describe the meaning of these traits and their determination.

In economics, some studies consider the role of the "big-five" in explaining wage, employment, education, and marriage outcomes. However, none of these studies introduce personality traits within a life-cycle framework in which these outcomes are jointly determined. Also, the few dynamic models in the literature that incorporate noncognitive traits usually represent traits using a single factor rather than multidimensional measures.[1]

The goals of this paper are: (i) to incorporate the "big-five" personality traits within a dynamic life-cycle framework model of education and employment choices (ii) to explore the role of personality traits as determinants of unobserved heterogeneity, and (iii) to use the estimated model to evaluate the distributional effects of two kinds of educational policies. To this end, we develop and estimate a dynamic model of schooling, work, and occupational choices that assumes that individuals ages 15 to 49 make one of four mutually exclusive choices: attending school, staying home, working in a white-collar job, or working in a blue-collar job. After age 49, individuals are assumed to stay in their most recent sector choice until retirement (age 65) (to ease the computational burden of estimating the model). Individual endowments at age 15 consist of personality traits, cognitive ability, and family background characteristics, which include parental schooling, number of siblings, birth order, and whether the person lived with both parents at age 14. To allow for unobserved heterogeneity in a tractable way, we assume each individual is one of four types (denoted I–IV). An individual's type can affect their pecuniary and nonpecuniary rewards from choosing particular schooling or work options.

We incorporate the "big five" personality traits into the model in a parsimonious way as a determinant of the unobserved type probabilities. In the data, personality traits are observed to change with age and to be affected by school attendance. We therefore allow in the model for an individual's unobserved type to potentially vary over time. In the DCDP literature, the standard approach is to assume fixed types (e.g., Keane and Wolpin (1997), Yamaguchi (2012), and Sullivan (2010)). However, recent methodological papers by Hu, Shum, Tan, and Xiao (2015) and Arcidiacono and Miller (2011) consider time-varying types that follow a Markov process similar to our specification.[2]

Our model is estimated using Simulated Method of Moments and using the Household Income and Labour Dynamics in Australia (HILDA) longitudinal data, waves 1 (2001) through 13 (2013). The data have repeated measures of personality traits as well as a one-time measure of cognitive ability. The estimation results show that the unobserved types are malleable, particularly at early ages. At age 15, individuals have on average a 75% probability to change type, but by age 30 their type stabilizes.[3] We perform a test for type stability, which is rejected in our data.

We use the estimated model to evaluate two education policies: making senior secondary school compulsory and providing a 50% cost subsidy to attend college. Both

[1] For example, Borghans, Duckworth, Heckman, and Ter Weel (2008).

[2] To our knowledge, varying-type models have been considered from a theoretical perspective and have not yet been implemented in a DCDP context.

[3] Our results are broadly consistent some psychology studies on personality trait stability. For example, Terracciano, Costa, and McCrae (2006) and Terracciano, McCrae, and Costa (2010) reported that intraindividual stability increases up to age 30 and thereafter stabilizes.

policies provide incentives to enroll in school but they differ in their distributional implications. We find that individuals belonging to types I and IV, who tend to come from more advantaged SES backgrounds, have a comparative advantage in the schooling sector, and receive the most benefit from the college subsidy policy. Their average years of completed education increases by almost 1 year. However, individuals belonging to types II and III, who tend to come from lower SES backgrounds, also experience significant benefits from the tuition subsidy in terms of increased education, earnings, and utility. In contrast, the impacts of compulsory senior secondary school are concentrated among individuals from lower SES backgrounds (types II and III).

To study the relevance of time-varying heterogeneity and personality traits in assessing educational policy impacts, we also simulate policy effects under an assumption that the types are fixed, in the spirit of Keane and Wolpin (1997). In such a model, there is less incentive for disadvantaged groups to pursue education, because they can no longer alter their disadvantaged types. The educational policy impacts on annual earnings and on educational attainment attributable are significantly smaller and the distribution of impacts is more unequal.

This paper proceeds as follows. Section 2 reviews related literature. Section 3 describes the HILDA data and the big five personality measures. Section 4 describes the model. Section 5 discusses identification and Section 6 explains the estimation strategy. Section 7 presents the estimation results and provides evidence on model fit. Section 8 explores the relationship between personality traits, types, and choices, evaluating the importance of personality traits in explaining the ex ante utility, and also performs a test for the type stability. Section 9 reports results from simulating two policy experiments and Section 10 concludes.

2. Related literature

The "big five" personality traits are defined as follows: (1) extraversion: an orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by positive affect and sociability, (2) neuroticism: a chronic level of emotional instability and proneness to psychological distress, (3) openness to experience/intellect: the tendency to be open to new esthetic, cultural, or intellectual experiences, (4) conscientiousness: the tendency to be organized, responsible, and hardworking and (5) agreeableness: the tendency to act in a cooperative, unselfish manner.

Several studies examine the influence of personality traits on wages and on occupational choices. For example, both Nyhus and Pons (2005) and Salgado (1997) found that emotional stability and conscientiousness are strongly correlated with wage and job performance. Cubel, Nuevo-Chiquero, Sanchez-Pages, and Vidal-Fernandez (2016) examined whether big five personality traits affect productivity using data gathered in a laboratory setting where task effort is directly measured. They find that individuals who exhibit high levels of conscientiousness and emotional stability perform better on the task. Fletcher (2013) found a robust relationship between personality traits and wages using sibling samples and family fixed effect estimators. Specifically, conscientiousness, emotional stability, extraversion, and openness to experience were all found to positively affect wages.

There are a few papers that examine the correlation between personality traits and educational attainment. For example, Lundberg (2013) found positive correlations between personality traits (such as conscientiousness, agreeableness, and openness to experience) and college entrance. Dahmann and Anger (2014), Kassenboehmer, Leung, and Schurer (2018) and Schurer (2017) argued that educational experiences in secondary school and at university shape students' personalities.

Our paper is also related to the burgeoning literature examining the non-cognitive skill formation process. Heckman, Stixrud, and Urzua (2006) studied the effect of noncognitive skills on schooling decisions and subsequent labor market outcomes. Cunha and Heckman (2007) found that dynamic complementarity and self-productivity are key features that generate multiplier effects in the production of human capital. Cunha and Heckman (2008) estimated a linear dynamic model to study the formation of cognitive and noncognitive skill as it depends on parental investment. Cunha, Heckman, and Schennach (2010) extended the linear production technology assumption to a nonlinear setting, which allows identification of elasticity of substitution parameters governing the trade-off between early and late investments. Agostinelli and Wiswall (2016) considered ways of further relaxing some of the identification assumptions in Cunha, Heckman, and Schennach (2010). Mullins (2019) developed and estimated a model of children's cognitive and behavioral skill development that he used to evaluate the impact of government antipoverty programs. Heckman and Raut (2016) formulated a dynamic structural model that relates preschool investment choices that affect skill formation with schooling and earning outcomes later in life.

3. Data

Our analysis is based on a sample of individuals from the Household Income and Labour Dynamics in Australia (HILDA) longitudinal data set. HILDA is a representative one in one thousand sample of the Australian population. It is an ongoing annual survey starting from the year 2001 with 19,914 initial individuals from 7682 households (Department of Social Services, Melbourne Institute of Applied Economic and Social Research (2018)). We make use of the following variables: (1) labor market outcomes including occupational information (coded following the ANZSCO system[4]), annual labor earnings, and working hours; (2) family background information including parental education levels, number of siblings and birth order as well as measures of household composition; (3) educational attainment levels; (4) cognitive ability measured in wave 12; and (5) the "big five" personality traits assessment repeatedly collected in waves 5, 9, and 13.

To the best of our knowledge, HILDA has the best quality measures of personality traits among all nationwide data sets. For the majority of respondents, we observe three

[4]In practice, we classify all occupations into two categories: blue-collar job and white-collar job. White-collar jobs includes managers, professionals, technicians, and trades person as well as clerical and administrative workers. Blue-collar jobs include community and personal service workers, sales workers, machinery operators, and drivers as well as laborers. See Table S5 in the Appendix in the Online Supplemental Material (Todd and Zhang (2020)) for details.

measurements of personality traits over an 8-year time window.[5] HILDA's "big five" information is based on 36 personality questions.[6] Respondents were asked to pick a number between 1 to 7 to assess how well each personality adjective describes them. The lowest number 1 denotes a total opposite description and the highest number 7 denotes a perfect description. According to Losoncz (2009), 28 of the 36 items load well onto five components when performing factor analysis. The other 8 items are discarded due to either their low loading value or their ambiguity on several traits. We construct "big five" measures using the average scores of the items belonging to each personality component.[7]

The big five personality traits are available for 4938 males interviewed in wave 5 and for 5048 and 6771 male respondents in waves 9 and 13. We include in our analysis individuals who have at least one measure of personality traits. Cognitive ability is only surveyed once in wave 12.[8] We construct a one-dimensional measure of cognitive ability from three different measurements: (i) Backward Digits Span, (ii) Symbol Digits Modalities, and (iii) a 25-item version of the National Adult Reading Test. Specifically, we rescale the three cognitive ability task scores to be mean 0 and variance 1 and then obtain the average value of these three measurements that is used in our analysis.

3.1 *Additional background variables and sample restrictions*

In addition to the cognitive and noncognitive trait measures, we use the following family background information in our analysis: sibling information (including whether the person has siblings, whether he is the eldest child in the family and how many siblings), an indicator of growing up in an intact family, parental education, and parental working status.[9] We also use state of residence.

Our estimation focuses on males age 15–44. Women are not included to avoid the complication of modeling fertility, which likely impacts schooling and labor supply decisions. Individuals serving in the military are also excluded. Lastly, we drop person-year observations with missing information for key state space variables in our model. The remaining sample has 36,639 total observations from 4215 individuals.

[5] One alternative national-wide data set providing personality traits inventory assessment is German Socio-Economic Panel (GSOEP) study. GSOEP also surveys "big five" three times in years 2005, 2009, and 2013.

[6] The personality questionnaire is shown in the Appendix available in the Online Supplemental Material (Todd and Zhang (2020)).

[7] More specifically, openness to experience is constructed from average scores on six adjective items including imaginative, creative, intellectual, philosophical, deep and complex. Conscientiousness is constructed from average scores on six adjective items including orderly, disorganized, efficient, sloppy, inefficient, and systematic. Extraversion is constructed from average scores on six adjective items including quiet, shy, talkative, extroverted, bashful, and lively. Agreeableness is constructed from average scores on four items including warm, kind, sympathetic, and cooperative. Lastly, emotional stability is constructed from average scores on six items including moody, temperamental, jealous, fretful, envious, and touchy. We verify the internal construct reliability by examining the most widely index—Cronbach's alpha. All traits show relatively high internal consistency ($\alpha > 0.7$) in all waves 5, 9, and 13.

[8] According to Wooden (2013), the response rate is high, approximately 93%.

[9] All the parental questions pertain to the time when the respondent was age 14.

TABLE 1. Sample summary statistics.

Variable	Proportions (%)	Variable	Proportions (%)
State		*Background info when you were* 14	
New South Wales	31.02	***Father Education***	
Victoria	25.46	College	64.52
Queensland	20.42	Not college	35.48
South Australia	8.83	***Mother Education***	
Western Australia	8.83	College	49.05
Tasmania	2.83	Not college	50.95
Northern Territory	0.55	***Father Working***	
Australian Capital Territory	2.08	Employed	95.81
Year (Cohort)		Not employed	4.19
1961–1969	20.01	***Father Occupation***	
1970–1979	28.83	White collar	72.32
1980–1989	27.13	Blue collar	27.68
1990–1998	24.03	***Mother Working***	
Ever had siblings		Employed	63.70
Had siblings	95.71	Not employed	36.30
No siblings	4.29	***Mother Occupation***	
Sibling numbers		Not asked	16.43
Not Asked	4.36	White collar	52.86
1	29.41	Blue collar	30.71
2	32.48	***Family Intactness***	
3	17.01	Both parents	80.3
4	8.79	Father and step	1.3
5 or more	7.95	Mother and step	5.01
Eldest Sibling		Father only	3.03
Not asked	4.29	Mother only	8.76
Oldest	34.9	Other	1.6
Not oldest	60.8		
Total individuals		2934	

Note: This table shows summary statistics for males age 15–44. The family background information (parental education, employment, occupations, and family intactness) refer to the time when individuals were age 14. Data source: HILDA, 2001–2013.

Selected summary statistics of individual's characteristics are reported in Table 1. Our sample is distributed across eight states and territories. Most individuals (>95%) have siblings. Approximately one-third are the eldest child. Table 1 also provides statistics on parental education and occupations at the time the individual was age 14. Almost two-thirds of fathers have a college degree whereas only half of mothers have a college degree. Most fathers were employed (>95%), but only about two-thirds of the sample had working mothers (64%). The majority of fathers' jobs were in white-collar occupations (72%), in comparison to 53% of mothers' jobs. The majority of individuals (80%) report residing with both parents at the age of 14.

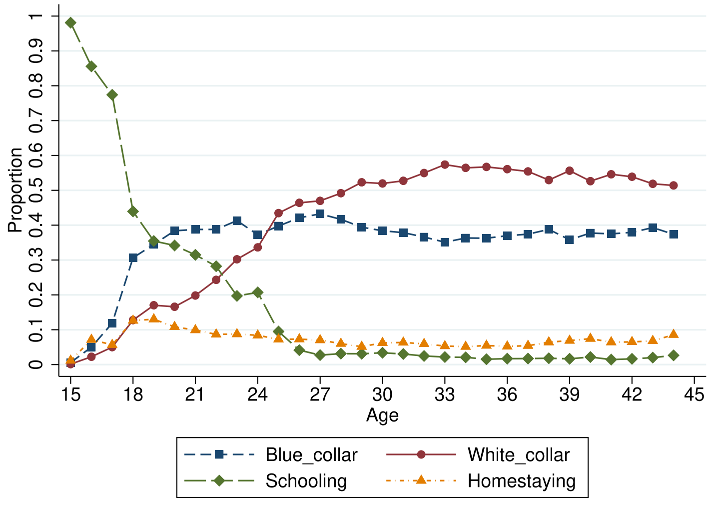

FIGURE 1. Proportion working and attending school by age (% of the same age cohort). *Note*: This figure shows the proportion in each of four choices (schooling, blue collar, white collar, home staying) for men ages 15–44. Data source: HILDA data, 2001–2013.

3.2 *Education and occupation choices over the life cycle*

In the HILDA survey, individuals report school enrollment and employment information annually.[10] The employment information includes employment status, hours worked, total annual earnings, and occupational codes. Figure 1 shows the choice distribution of schooling, staying at home, blue-collar jobs, and white-collar jobs by age. At age 15, almost everyone is enrolled in school but after age 19, the fraction drops sharply to around 35%. The majority of secondary school graduates choose to work immediately rather than to continue their tertiary education. The school enrollment rate decreases from 20% at age 23 to around 3% at age 27.

We define an individual to be "working" if he reports to be working positive hours and he is not enrolled in school.[11] An individual is defined to be "staying home" if neither working nor in school.[12] The blue-collar participation rate increases to around 30% at age 18 and then stabilizes at around 40%. The increase in the white-collar participation rate seen between ages 22 to 25 suggests that a college degree is a prerequisite for many white-collar occupations. The white-collar participation rate continues to increase after age 26, as some workers switch from blue-collar to white-collar jobs. The percentage staying home increasing shortly after graduation from secondary school graduation and then declines to roughly 5%.

[10]A rough classification of the tertiary education certificates includes 1. Certificates I–IV; 2. Diploma, Advanced Diploma, Associate Degree; 3. Bachelor degree and honors; 4. Graduate Certificate and Graduate Diploma; 5. Master degree; 6. Doctoral degree.

[11]A small fraction of individuals report working and attending school simultaneously. When it happens, we record an individual as in school if his age is less than 25 and as at work if his age is greater or equal than 25.

[12]We do not distinguish between being unemployed and out of labor force, as the decision to be unemployed is always considered voluntary under our model.

Table 2. Average personality traits by education level.

Education Level	Open. to Exp.	Conscientiousness	Agreeableness	Extraversion	Emotional Stab.
Sec. school or lower	4.243	4.824	5.070	4.447	5.026
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
College dropouts	4.357	4.883	5.122	4.350	5.012
	(0.022)	(0.023)	(0.019)	(0.023)	(0.022)
College graduates	4.577	5.028	5.160	4.293	5.113
	(0.012)	(0.012)	(0.010)	(0.013)	(0.012)

Note: This table shows the mean and standard error (in parentheses) of the "big-five" personality traits. The sample is HILDA, waves 5, 9, and 13. Each trait has a value ranging from 1 to 7.

Data on personality traits were gathered in 2005, 2009, and 2013. Table 2 reports the average personality trait scores for three different education levels: senior secondary school or lower, college dropouts, and college graduates. College graduates have higher average scores on emotional stability, openness to experience, conscientiousness, and agreeableness. However, they tend to be less extraverted.

Table 3 reports the difference in personality traits between white-collar and blue-collar workers. White-collar workers have higher average scores on emotional stability, openness to experience, conscientious, and agreeableness. The greatest differences in scores by occupation sector are seen in conscientiousness and openness to experience scores.

3.3 *Stability of personality traits*

The stability of personality traits is an important issue discussed both in the psychology and economics literature. Some studies find that personality traits are stable for adults (Terracciano, Costa, and McCrae (2006), Terracciano, McCrae, and Costa (2010)) and Cobb-Clark and Schurer (2012). Other studies find that personality traits change with age, particularly during younger ages (Almlund et al. (2011), Cunha and Heckman (2007), Cunha, Heckman, and Schennach (2010)). In this section, we examine whether personality traits vary by age in the HILDA data. Figure 2 shows the average score on the "big five" over the life cycle using the 2013 data wave. Compared with the other traits, conscientiousness is the only personality trait that exhibits a significant increase with age. Changes in openness to experience, agreeableness and emotional stability are

Table 3. Average personality traits by occupation sector.

Occupation	Open. to Exp.	Conscientiousness	Agreeableness	Extraversion	Emotional Stab.
White collar	4.452	5.035	5.159	4.381	5.091
	(0.010)	(0.010)	(0.009)	(0.011)	(0.010)
Blue collar	4.153	4.887	5.088	4.382	5.015
	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)

Note: This table shows the mean and standard error (in parentheses) of the "big five" personality traits by occupation group (white collar and blue collar). The sample is HILDA, waves 5, 9, and 13. The personality trait values range from 1 to 7.

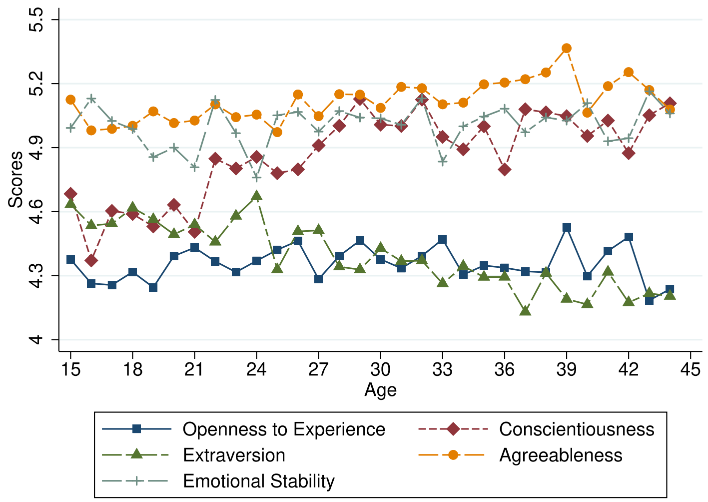

FIGURE 2. The "big five" personality trait scores by age. *Notes*: This figure shows changes in the "big five" personality traits with age. The measures are based on males ages 15–44 who report their personality traits in wave 13, HILDA.

moderate and not statistically distinguishable from 0. Extraversion decreases with age until age 35 and then stays stable. Our findings are consistent with patterns described in Elkins, Kassenboehmer, and Schurer (2017). They find that conscientiousness increases with age by 0.38 standard deviations, but the changes in other personality traits are moderate and do not exceed 0.15 standard deviations.

3.4 *Personality traits, schooling, and occupation sector*

We investigate in Table 4 how working and schooling correlate with changes in personality traits. After standardizing the scores to each have mean 0 and variance 1, we estimate fixed effects models of personality traits regressed on education, age, age squared, age interacted with education and indicators for whether the individual is in a white- or blue-collar occupation. Each column of Table 4 reports the estimated coefficients for a different trait. The coefficients associated with age and education are statistically significantly different from zero for three out of the five traits (openness, conscientiousness, and emotional stability). The estimated coefficients imply that the marginal effects of education on personality traits are greatest at younger ages and then diminish with age.[13] Occupational sector does not have a systematic relationship with personality traits.

Table 5 shows the relationship between log earnings, personality traits, and cognitive ability. The specification is analogous to a Mincer log earnings regression (esti-

[13] For example, the marginal effect of education on openness to experience is 0.076 ($0.150 - 0.493 * 15/100$) at age 15 then decreases to 0.002 ($0.150 - 0.493 * 30/100$) at age 30. The marginal effect of education on emotional stability decreases from -0.056 ($-0.147 + 0.609 * 15/100$) at age 15 to 0.005 ($-0.147 + 0.609 * 25/100$) at age 25. The results are consistent with the findings reported in the literature that personality traits are less malleable at older ages (see, e.g., Cunha and Heckman (2008, 2009)).

TABLE 4. The effect of education/occupation changes on personality traits, fixed effect models.

	(1) Opn	(2) Cos	(3) Agr	(4) Ext	(5) Stb
Education (β_1)	0.150	−0.0489	0.0943	−0.0553	−0.147
	(0.0587)	(0.0592)	(0.0619)	(0.0509)	(0.0642)
$Age * Edu/100$ (β_2)	−0.493	0.218	−0.211	0.189	0.609
	(0.228)	(0.230)	(0.241)	(0.198)	(0.250)
Age (β_3)	0.046	0.112	0.0482	0.0193	−0.00726
	(0.0322)	(0.0325)	(0.0340)	(0.0279)	(0.0352)
$Age^2/100$ (β_4)	−0.0694	−0.152	−0.0639	−0.0687	0.0226
	(0.0566)	(0.0572)	(0.0598)	(0.0491)	(0.0619)
White collar (β_5)	−0.103	0.03	−0.0897	$-3.69e^{-4}$	−0.068
	(0.0622)	(0.0628)	(0.0657)	(0.0680)	(0.0540)
Blue collar (β_6)	−0.0267	0.00946	−0.0592	0.0122	0.0100
	(0.0632)	(0.0638)	(0.0667)	(0.0549)	(0.0691)
R^2	0.828	0.823	0.762	0.866	0.800
Observations	2800	2800	2800	2800	2800

Note: The sample includes males whose personality traits are measured at least once between ages 15 and 30. Standard errors in parentheses.

mated for individuals with positive earnings). The first column presents estimates where the included variables are education, potential experience, and potential experience-squared.[14] The so-called "rate of return" to education is around 11%.[15] The second column adds personality traits to the specification. All of the traits have associated coefficients that are statistically different from zero. The most important trait associated with higher earnings is conscientiousness. Three of the traits (openness, emotional stability, and agreeableness) have, ceteris paribus, negative effects on earnings. The regression also includes cognitive ability (standardized to have mean zero and variance 1). Ceteris paribus, a one standard deviation increase in cognitive ability increases earnings by 7–8%, an effect comparable in magnitude to the effect of conscientiousness. The third column adds to the specification a set of family background variables as additional control variables (described in the table notes). The estimated coefficients on all the variables change little when the family background variables are added, although the overall R-squared increases.

4. The model

We develop a discrete choice dynamic programming (DCDP) model of decision-making with regard to education, employment, and occupation sector between ages 15 to 49. At each age, individuals maximize their remaining discounted lifetime utility. The terminal age is 65 but to facilitate computation, we assume that individuals make choices until age 49 and then stay in their age 49 sector choice from ages 50–65. The choice set in each year consists of four mutually exclusive options $m \in M$: working in either a blue- or

[14] Potential experience is defined as age-years of education-6.

[15] It is relatively high, in part because the sample is restricted to individuals ages 17–44. The estimated rate of return is lower when older age individuals are included.

TABLE 5. How personality traits and cognitive ability relate to log wages.

	Log Earnings		
	(1)	(2)	(3)
Education	0.116	0.110	0.133
	(0.003)	(0.003)	(0.003)
Potential experience	0.149	0.148	0.172
	(0.004)	(0.004)	(0.005)
Potential experience squared/100	−0.396	−0.396	−0.377
	(0.013)	(0.013)	(0.014)
Openness		−0.054	−0.056
	...	(0.006)	(0.006)
Conscientiousness		0.078	0.083
	...	(0.006)	(0.006)
Emotional stability		−0.015	−0.011
	...	(0.006)	(0.006)
Agreeableness		−0.026	−0.030
	...	(0.007)	(0.007)
Extraversion		0.031	0.028
	...	(0.006)	(0.006)
Cognitive		0.077	0.073
	...	(0.009)	(0.009)
Family characteristics	No	No	Yes
Observations	16,408	16,408	16,408
R Square	0.303	0.318	0.351

Note: The sample includes males whose personality traits are measured at least once between ages 15 and 44. The family background variables includes family intactness, father's occupation, parental education, sibling numbers, birth order and cohort. Standard errors are reported in parentheses.

white-collar occupation, attending school, or staying home. Let $d_m(a) = 1$ if alternative m is chosen at age a, $d_m(a) = 0$ otherwise.

Individual endowments at age 15 consist of personality traits, cognitive ability, and family background characteristics. These include parental schooling, the number of siblings, birth order, and whether the person lived with both parents at age 14. To allow for unobservable heterogeneity in a tractable way, we assume each individual is one of four types $k(a) = \{1, 2, 3, 4\}$. An individual's type can affect their pecuniary and nonpecuniary reward from choosing particular alternatives. As noted in the Introduction, one important aspect of our model that deviates from most of the literature (e.g., Keane and Wolpin (1997)) is that it allows types to evolve in a way that may depend on age and personality traits.[16]

We use $\Theta(a)$ to represent personality traits and $k(a)$ to denote the unobserved type at age a, assumed to be known by the individual but not by the econometrician. $s_o(a)$ represents all other observed state variables. At age 15, the initial type $k(15)$ is determined by the initial endowment $s_o(15)$. Then given the initial type $k(15)$ and observed state variables $s_o(15)$, the agent chooses the alternative $d_m(a)$ that gives the highest con-

[16] Arcidiacono and Miller (2011) and Hu et al. (2015) discussed conditions needed to identify models with time-varying types. They are further described below.

tinuation value. The state variables, $s_o(16)$, are updated according to the choice $d_m(15)$, and then the new type $k(16)$ is drawn with type probabilities depending on $s_o(16)$ and the previous period type $k(15)$.

4.1 *Laws of motion for* $s_o(a)$ *and* $k(a)$

The time-varying part of $s_o(a)$ consists of four components, $s_o(a) = (g(a), x_1(a), x_2(a), \Theta(a))$. $g(a)$ represents accumulated education while $x_1(a)$ and $x_2(a)$ represent accumulated blue-collar and white-collar experience at age a. We first specify the law of motion for states $g(a)$, $x_1(a)$, $x_2(a)$ and then discuss the transition probability functions governing the personality traits $\Theta(a)$ and types $k(a)$.

Years of schooling and occupation-specific experience evolve in a deterministic way. The updating of $g(a)$, $x_1(a)$ and $x_2(a)$ proceeds as follows:

$$\begin{aligned} g(a) &: g(a+1) = g(a) + d_3(a), \\ x_m(a) &: x_m(a+1) = x_m(a) + d_m(a), \quad m = \{1, 2\}. \end{aligned} \tag{1}$$

We assume that the true nth personality trait $\theta_n \in \Theta$, $\{n = 1, 2, 3, 4, 5\}$ is measured with error and denote the measurement error shock as $\zeta_n(a)$. In Section 3.3 (Table 4), we reported estimates for a fixed effect model that suggested that personality traits change with schooling. We therefore adopt the following specification for the evolution of each trait:

$$\begin{aligned} &\theta_n^M(a) = \theta_n(a) + \zeta_n(a), \\ &\theta_n(a) = \theta_n(15) + \gamma_{0n} + \gamma_{1n} g(a) + \gamma_{2n}(a-15) g(a) + \gamma_{3n}(a-15) + \gamma_{4n}(a-15)^2, \end{aligned} \tag{2}$$

where $\theta_n^M(a)$ is the measure of the nth personality trait at age a and $\theta_n(a)$ is the true trait without measurement error. γ_{3n} and γ_{4n} capture age effects. The term $\gamma_{1n} + \gamma_{2n}(a-15)$ captures the potential effect of schooling on personality traits, which may again vary by age.

As previously noted, we specify a Markov process for the evolution of the discrete types. After the initial period, the type $k(a)$ can stay the same with probability $1 - p(a)$ or change with probability $p(a)$.[17] Conditional on changing, we use notation $q_k(a)$ to represent the probability of becoming type $k \in \{1, 2, 3, 4\}$. Let $L(a)$ denote the Markov type transition matrix between period a to period a+1, which has the following form:

$$L(a) = \big(1 - p(a)\big) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + p(a) \begin{bmatrix} q_{k=1}(a) & q_{k=1}(a) & q_{k=1}(a) & q_{k=1}(a) \\ q_{k=2}(a) & q_{k=2}(a) & q_{k=2}(a) & q_{k=2}(a) \\ q_{k=3}(a) & q_{k=3}(a) & q_{k=3}(a) & q_{k=3}(a) \\ q_{k=4}(a) & q_{k=4}(a) & q_{k=4}(a) & q_{k=4}(a) \end{bmatrix}, \tag{3}$$

where

$$p(a) = \frac{1}{1 + \exp\big(\gamma_7 + \gamma_8(a-15) + \gamma_9(a-15)^2\big)}, \tag{4}$$

[17] We assume the probability of changing types $p(a)$ depends on age but does not vary by type k.

$$q_k(a) = \frac{\bar{v}_k^a(\Theta, d_z)}{\sum_{k=1}^{K=4} \bar{v}_k^a(\Theta, d_z)}, \tag{5}$$

$$\begin{aligned} \log v_k^a(\Theta, d_z) &= \log \bar{v}_k^a(\Theta, d_z) + \eta_k(a) \\ &= \gamma_{3k} + \sum_{n=1}^{N=5} \gamma_{4kn}\theta_n(a) + \sum_{z=1}^{Z} \gamma_{5zk} d_z + \eta_k(a). \end{aligned} \tag{6}$$

At age 15, the initial types are directly drawn from the distribution $q_k(15)$. In subsequent ages, types are updated following the Markov transition matrix $L(a)$. When $p(a)$ is close to 0, then $L(a)$ corresponds to an identity matrix $I_{4\times 4}$ and the types, k, are essentially fixed. When $p(a) = 1$, types do not persist from the previous time period. We estimate $p(a)$, allowing for the possibility that types become more or less persistent with age. The probability of each type $q_k(a)$ follows a multinomial logit form (equation (5)). Equation 6 shows how the type probability may depend on personality traits $\theta_n(a)$ and background characteristics d_z.[18]

4.2 *Rewards associated with each alternative*

An individual can choose to work in either a blue-collar or a white-collar occupation. The reward to a particular sector includes the wage compensation $w_m(a)$ and any non-pecuniary reward $r_m(a)$. $\epsilon_m(a)$ is the preference shock when choosing mth alternative. $m = 1$ denotes the blue collar and $m = 2$ the white-collar alternative. The utility function at age a is

$$u_m(a) = w_m(a) + r_m(a) + \epsilon_m(a), \quad m = \{1, 2\}. \tag{7}$$

As in Keane and Wolpin (1997), the wage is specified as a human capital pricing equation. It is given by the product of the price per unit of human capital p_m and the amount of human capital $e_m(a)$ embodied in the individual, that is, $w_m(a) = p_m e_m(a)$. Human capital is determined by work experience $x_m(a)$, schooling years $g(a)$ as well as cognitive ability c:[19]

$$\begin{aligned} e_m(a) = \exp\Bigg(& e_m^k + \sum_{i=1}^{I} \beta_{m0i} d_i + \beta_{m1} g(a) + \big(\beta_{m2} + \beta_{m3} I\{x_m(a) \le 2\}\big) x_m(a) \\ & + \beta_{m4} x_{3-m}(a) + \beta_{m5} x_m^2(a) + \beta_{m6} x_m(a) g(a) + \beta_{m7} c + \xi_m(a)\Bigg), \\ & m = \{1, 2\}, \end{aligned} \tag{8}$$

[18] The family background information includes sibling numbers, birth order, and parental education level.

[19] Cognitive ability was only measured in one wave of the data, so we do not attempt to model its dynamics over time.

which yields a log-wage equation the form:

$$\begin{aligned}\log w_m(a) = \log p_m + e_m^k + \sum_{i=1}^{I} \beta_{m0i} d_i + \beta_{m1} g(a) + \bigl(\beta_{m2} + \beta_{m3} I\bigl\{x_m(a) \leq 2\bigr\}\bigr) x_m(a) \\ + \beta_{m4} x_{3-m}(a) + \beta_{m5} x_m^2(a) + \beta_{m6} x_m(a) g(a) + \beta_{m7} c + \xi_m(a), \\ m = \{1, 2\}.\end{aligned} \tag{9}$$

In (9), d_i, $i \in \{state \times cohort\}$ is a fixed effect for being a member of particular age cohort and residing in a particular state. e_m^k is the type-specific component of reward, which represents the advantage or disadvantage of type k when choosing alternative m. $g(a)$ represents the years of schooling, and $x_m(a)$ denotes work experience in sector m. The component $\beta_{m3} I\{x_m(a) \leq 2\} x_m(a)$ captures a potential differential in returns to experience when the agent is new in an occupation (has 2 years or less experience). $x_{3-m}(a)$ represents working experience in the other sector $3 - m$. Therefore, the coefficient β_{m4} captures the return to other sector work experience. The component $\beta_{m6} x_m(a) g(a)$ captures the interaction term between work experience $x_m(a)$ and education $g(a)$, which is included to allow returns to experience to differ with education. The component $\beta_{m7} c$ captures the return to cognitive abilities. $\xi_m(a)$ is a skill technology shock, which is assumed to be i.i.d. normal distribution. The second term in equation (7), $r_m(a)$, represents nonpecuniary aspects of choosing a certain occupation (such as working hours flexibility) expressed in monetary equivalent units. For the purpose of identification, we normalize the nonpecuniary utility from white-collar job $r_1(a)$ equal to 0. We allow the nonpecuniary utility from the blue-collar sector $r_2(a)$ to vary with education level:

$$\begin{aligned} r_1(a) &= 0, \\ r_2(a) &= \beta_8 + \beta_9 I\bigl[g(a) \leq 12\bigr]. \end{aligned} \tag{10}$$

If a person chooses to attend school, the per-period utility consists of two parts: a nonpecuniary component, which may reflect any physical and mental costs when attending school, and a pecuniary component, such as tuition costs and fees.

The utility associated with school attendance at age a is

$$\begin{aligned} u_3(a) = e_3^k + \sum_{r=1}^{R} \alpha_r d_r + \alpha_c c + \alpha_0 I(age < 18) - \alpha_1 I(college) \\ - \alpha_2 I(graduate) + \epsilon_3(a). \end{aligned} \tag{11}$$

d_r is a cohort-specific effect, and c is the effect of cognitive ability on the education choice. The term $\alpha_0 I(age < 18)$ captures the extra utility of attending school when the agent is under the age 18. α_1 and α_2 are per period schooling costs of attending college and attending graduate school. Lastly, e_3^k is the type-specific reward from attending school.

The reward from staying home, $u_4(a)$, consists of the type-specific component e_4^k, an age effect and an age squared effect, α_3 and α_4, and a preference shock $\epsilon_4(a)$, that is,

$$u_4(a) = e_4^k + \alpha_3 \cdot age + \alpha_4 \cdot age^2 + \epsilon_4(a). \tag{12}$$

Personality traits do not directly appear in the choice-specific utilities. Instead, they affect the choices indirectly through their influence on an individual's type probability.[20] Different types have different type-specific components e_m^k for each choice m.

4.3 *Information structure*

In our model, individual heterogeneity comes from two sources: ex ante endowments $\{k(15), \Theta(15), c, Z, state, cohort\}$ at age 15 and ex post realized shocks $(\epsilon_m(a), \xi_m(a), \zeta_n(a), \eta_k(a))$.[21] In terms of timing, we assume that the shocks governing the evolution of personality and of types are realized first. After that, individuals observe preference shocks and choose their preferred sector. Then wage shocks are realized.

Let $S^v(s) \subseteq S$ denote the set of visited states and $S^f(s) \subseteq S$ as the set of feasible states that can reached from s. Given the timing assumptions, we define $\iota(s)$ as the information set in state s by specifying all components known in the state, where

$$\iota(s) = \begin{cases} \epsilon_m(a); \zeta_n(a); \xi_m(a); \eta_k(a): & \text{for all } s(a) \in S^v(s), \\ \epsilon_m(a+1): & \text{for } s'(a+1) \in S^f(s), \\ k(15), \Theta(15), c, Z, state, cohort; \Omega: & \text{and for all } s. \end{cases}$$

An individual in state s knows all state variable laws of motion, $\Pr(s(a+1)|s(a), d_m(a))$. He uses the distribution of wage shocks $F_m(\xi(s))$, idiosyncratic preference shocks $F_m(\epsilon(s))$, traits transition shocks $F_n(\zeta(s))$, and type transition shocks $F_k(\eta(s))$ to form an expectation over future states. For computational simplicity, $\xi_m(a)$ and $\zeta_n(a)$ are assumed to be uncorrelated and normally distributed, whereas $\epsilon_m(a)$ and $\eta_k(a)$ are assumed to be type I extreme value distributed. Conditional on the unobserved types, the other shocks are assumed to be i.i.d. over time.

5. Identification

The general procedure for incorporating multinomial types into longitudinal models dates back to Heckman (1981) and Heckman and Singer (1984). The method was first used in the context of DCDP models with fixed types in Keane and Wolpin (1997). Identification of a discrete choice model with serially correlated, unobserved types is discussed in Hu et al. (2015). They discuss two key assumptions required for identification: (1) the choice at age a $d_m(a)$ is independent of variables from last period

[20]Each of the five traits can take values 1 through 7. The structure we assume avoids the need to include a five-dimensional personality trait vector in the time-varying state space. Only the initial personality traits are included in the state space. The traits evolve with age and with attendance at school according to equation (2) and are assumed to be measured with error.

[21]Three state variables are constant for every individual at age 15. $g(15) = 0$, $x_1(15) = 0$, $x_2(15) = 0$.

$a-1$ after conditioning on the state variables at age a: $\Pr(d_m(a)|s_{-k}(a), k(a), s_{-k}(a-1), k(a-1), d_m(a-1)) = \Pr(d_m(a)|s_{-k}(a), k(a))$, where $s_{-k}(a)$ represents the set of state variables excluding type $k(a)$, and (2) the type $k(a)$ is independent of last period choices $d_m(a-1)$ conditional on other current period and last period state variables: $\Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1), d_m(a-1)) = \Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1))$. Following Hu et al. (2015), the conditional probability of observing a given choice, state space and type can be factored into three terms:

$$\begin{aligned}
&\Pr\bigl(d_m(a), s_{-k}(a), k(a)|d_m(a-1), s_{-k}(a-1), k(a-1)\bigr) \\
&\quad = \underbrace{\Pr\bigl(d_m(a)|s_{-k}(a), k(a)\bigr)}_{\text{CCP}} \underbrace{\Pr\bigl(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1)\bigr)}_{\text{Law of motion for type } k} \\
&\qquad \times \underbrace{\Pr\bigl(s_{-k}(a)|d_m(a-1), s_k(a-1), k(a-1)\bigr)}_{\text{Law of motion for } s_{-k}}.
\end{aligned}$$

Hu et al. (2015) showed that the right three terms can be identified with observations from at least three time periods $\{d_m(a), s_{-k}(a), d_m(a-1), s_{-k}(a-1), d_m(a-2)\}$. Appendix A discusses how to apply their result in our context.[22]

The utility values associated with the schooling choice and with the home choice as well as the nonpecuniary values of choosing a white- or blue-collar job are not directly observed. In the last time period, the set-up of the choice problem is analogous that of a multinomial logit model given the types. The choices we observe allow us to infer relative but not absolute utilities, so identification requires normalizing one of the utility values. We normalize the nonpecuniary value of the white-collar sector choice to be zero, so that utility in that sector corresponds to the wage. Lastly, the difference in conditional choice probabilities by type identifies the type-specific components e_m^k of the flow utility functions.[23] Wages are directly observed, but for selected subgroups that choose each sector. Using one period of data (e.g., the last period which is a static problem), a control function method (e.g., Heckman and Honore (1990)) could be used to identify the parameters of the wage equations.

Personality traits are observed in multiple time periods, so it is possible to directly identify the transition process from the data (equation (2)). The final parameter that we need to identify is the discount rate, which is identified through functional form assumptions that allow separation of the current period utility from future expected utility.

It would, in principal, be possible to construct a maximum likelihood estimator from the estimated conditional choice probabilities for each household. Model parameters are then identified if the first-order conditions are linearly independent. As previously described, we instead use a moment-based estimator, because of the problem of having some missing state variables. When using moment-based estimators, typically it is not possible to explicitly demonstrate the identification of all of the model parameters. The

[22] Hu et al. (2015) also made a stationary assumption on the conditional probability, but their results can be generalized to our case where it is is age-dependent.

[23] Identification of these kinds of models is discussed in Horowitz (1981).

hope is that by including enough sample statistics the model parameters will be identified and precisely estimated. As described below, we choose moments that capture data variation similar to that captured by the MLE first-order conditions. For example, the moments include (i) the proportions choosing different sectors by age, (iii) average wages by age, (ii) average personality traits by education, age, and occupation, and (iii) sector transition rates. The moments that we use are described in detail below.

Whether or not the model is "well identified" using a particular vector of sample moments is often determined after estimation has been attempted. Different sets of moments can yield different point estimates and associated standard errors in small samples, but it is seldom possible to determine an "optimal" vector of moments in a reasonably complex estimation problem. A specific parameter is said to be precisely estimated if the ratio of its point estimate to its estimated standard error is large in absolute value. In our case, it is almost never the case that this ratio of the parameter to its standard error is close to zero.

6. Estimation strategy

6.1 *Solving the dynamic programming problem*

At the beginning of age a, an individual has the state vector $s(a)$, determined by his choices up to age a. As previously described, the evolving state variables include the accumulated sector-specific experience $x_i(a)$, $i = 1, 2$, the completed schooling $g(a)$, personality traits $\Theta(a)$, and the unobserved type $k(a)$.[24] Let $d_m(t) = 1$ denote that alternative m is chosen at age t. The value function at age a is the maximum over all possible sequences of future choices given the current state space:

$$V\bigl(s(a), a, \Omega\bigr) = \max_{\{d_m(t)\}} E\left[\sum_{t=a}^{A} \delta^{\tau-a} \sum_{m=1}^{4} u_m(t) d_m(t) \middle| s(a)\right],$$

where Ω denotes a set of parameter values. The summation over t denotes the ages and the summation over m denotes the different sector choices. The problem can be written in Bellman equation form.

The alternative specific value function is

$$V_m\bigl(s(a), a, \Omega\bigr) = \tilde{u}_m\bigl(s(a), a\bigr) + \delta E\bigl[V\bigl(s(a+1), a+1, \Omega\bigr) | s(a), d_m(a)\bigr]$$

for $a < A$, and

$$V_m\bigl(s(A), A, \Omega\bigr) = \tilde{u}_m\bigl(S(A), A\bigr)$$

in the last time period. As previously noted, to facilitate computation, we impose an assumption on the model that the sector is chosen after preference shocks are realized but before the wage shock is realized. We denote $\tilde{u}_m(s(a), a)$ as $u_m(s(a), a)$ after integrating over the wage shock distribution (i.e., $\tilde{u}_m(s(a), a) = \int_{\xi_{m(a)}} u_m(s(a), a) \times$

[24] The personality traits at the initial age may not directly be observable, so in some cases we infer them using the approach described in Appendix B.

$f(w(\xi_m(a)))\,d\xi_m(a))$. Wages in the white- and blue-collar sectors are assumed to be normally distributed and uncorrelated (conditional on type). The expectation in the Bellman equation is taken over future wage and preference shocks and over the unobserved type transition process.[25] The value function is the max over the alternative specific value functions:

$$V\big(s(a), a, \Omega\big) = \max_{m \in M} V_m\big(s(a), a, \Omega\big).$$

Recall that the preference shocks enter additively into $u_m(s(a), a)$ and, for computational simplicity, are assumed to follow an i.i.d. type I extreme value distribution with a location parameter 0 and a common scale σ_c.

Let $\tilde{V}_m(s(a), a, \Omega)$ denote the choice-specific value function excluding the contemporaneous sector-specific preference shock $\epsilon_m(a)$,

$$V_m\big(s(a), a, \Omega\big) = \tilde{V}_m\big(s(a), a, \Omega\big) + \epsilon_m(a).$$

Because of the preference shock distributional assumptions, we have

$$\Pr\big(d_m(a) = 1 | s(a), \Omega\big) = \frac{\exp\big(\tilde{V}_m\big(s(a), a, \Omega\big)/\sigma_c\big)}{\sum_{j=1}^{4} \exp\big(\tilde{V}_j\big(s(a), a\big)/\sigma_c\big)}.$$

As shown by Rust (1987), the expected value function can be written as

$$\begin{aligned} E\big[V\big(s(a+1), a+1, \Omega\big) | s(a), d_m(a)\big] &= E_{\epsilon_m(a)} \max_{d_m(a)} \sum_{m=1}^{4} d_m(a)\big\{\tilde{V}_m\big(s(a), a, \Omega\big) + \epsilon_m(a)\big\} \\ &= \sigma_c \log\left(\sum_{m=1}^{4} \exp\big(\tilde{V}_m\big(s(a), a, \Omega\big)/\sigma_c\big)\right) + \sigma_c \gamma, \end{aligned}$$

where γ is the Euler's constant and σ_c is the scale parameter of the preference shock.[26]

The dynamic programming problem uses backward recursion for each set of parameter values under consideration. That is, in the last period A, when there is no future expected value function and using the previous equation, one obtains $E[V(s(A), A)|s(A-1), d_m(A-1), A-1]$ for each possible point in the state space. Plugging in $E[V(s(A), A)|s(A-1), d_m(A-1), A-1]$ into $\tilde{V}_j(s(A-1), A-1)$, one can then use the same expression to obtain $E[V(s(A-1), A-1)|s(A-2), d_m(A-2), (A-2)]$, and so on, back until the first time period. After solving the dynamic programming problem,

[25] Even though the realized wage shocks do not affect the contemporaneous utility associated with different sectors, the expected value functions will depend on the variance of the wage shocks.

[26] This closed-form representation of the value function is a big advantage in estimation because, without it, numerical integration over the structural errors is required to get the expected value function. It also generates an analytic one-to-one mapping between the choice probability and utility level of each choice. This tractable i.i.d. generalized extreme value (GEV) distributions assumption is also adopted in other recent DCDP papers such as Chan (2013) and Kennan and Walker (2011).

one obtains the expected future value functions for all possible state points. It is then possible to use the model to simulate choices and to implement a simulated method of moments optimization algorithm to estimate the parameters.

6.2 *Simulated method of moments estimation*

Our model parameters are estimated by simulated method of moments. We use an unconditional simulation approach starting from age 15, because occupation-specific experience stocks, which are part of the model's state space, are not directly observed and, therefore, need to be simulated from initial conditions.

The simulation process is as follows: For each individual i, given a set of parameters Ω:

1. Solve backward for choice-specific value function $V_m(s(a), \Omega)$ and choice probability $\Pr(d_m(a)|s(a), \Omega)$ following the procedure described previously.

2. Impute initial personality traits $\theta_n(15)$ from the observed personality traits (at up to three ages) following the procedure described Appendix B. Initial unobserved types $k(15)$ are drawn from equation (5).

3. Starting from $s(15) = g(15) = 0, x_i(15) = 0, k(15), \theta_n(15)$, simulate sequential shocks $\{\epsilon_m(a), \zeta_n(a), \xi_m(a), \eta_k(a)\}$ and compute the following outcomes: (1) agents' lifetime choices $d_m(a)$; (2) wage realizations $w_m(a)$ when $m = \{1, 2\}$, $a = \{18, \dots, 58\}$; and (3) personality traits $\theta_n(a), n = \{1, 2, \dots, 5\}$.

The simulation process is repeated for all $i = 1, 2, \dots, N$ individuals, given their initial state variables.

We then compute R moments using both the N simulated samples and the observed data, and then calculate the weighted difference between those R simulated moments $\tilde{M}_{N,R}(\Omega)$ and the data moments M_R, using the following objective criterion:

$$\hat{\Omega}_{N,R,W} = \arg\min_{\Omega}\big(\big(M_R - \tilde{M}_{N,R}(\Omega)\big)' W_R\big(M_R - \tilde{M}_{N,R}(\Omega)\big)\big), \tag{13}$$

where M_R denotes the data moments, and $\tilde{M}_{N,R}(\Omega)$ represents the simulated moment evaluated at the parameter set Ω based on N repeated simulations.[27]

We use the variance information of each data moment to form the weighting matrix, W_R. Del Boca, Flinn, and Wiswall (2014) showed the consistency for this type of estimator for large sample sizes, $\operatorname{plim}_{N\to\infty} \tilde{M}_{N,R}(\Omega_0) = M_R(\Omega_0)$.[28] In total, we match 314 moments to estimate 118 parameters. The following moments are used in estimation:

1. Sequential life-time choices (120 moments)

[27] This unconditional simulation algorithm is often used to estimate dynamic discrete choice models when some state variables are unobserved(e.g., Keane and Wolpin (2001), Keane and Sauer (2010)). The consistency and other asymptotic properties of this estimator based on unconditional simulation are discussed in Gourieroux and Monfort (1996, Section 2.2.2).

[28] Compared with directly calculating the optimal weighting matrix, this method simplifies computation significantly. Altonji and Segal (1996) discussed that gains from using an optimal weighting matrix may be limited.

- The fraction of individuals in the blue-collar occupation sector by age (15–44).
- The fraction of individuals in the white-collar occupation sector by age (15–44).
- The fraction of individuals in school by age (15–44).
- The fraction of individuals at home by age (15–44).

2. Earning profiles (108 moments)[29]

- Average log earnings of blue-collar workers by age (18–44).
- Average log earnings of white-collar workers by age (18–44).
- The standard deviation of log earnings of blue-collar jobs by age (18–44).
- The standard deviation of log earnings of white-collar jobs by age (18–44).

3. The transition matrix of the four sector choices from the current period to the next period (16 moments)

4. The mean value of personality traits by age, education level, and occupation sector (50 moments)

- Mean values of "big five" (openness to experience, conscientiousness, extraversion, agreeableness, emotional stability) by 5-year age groups.[30]

- Mean values of "big five" (openness to experience, conscientiousness, extraversion, agreeableness, emotional stability) by education group (educational years ≤ 12, educational years > 12).[31]

- Mean values of "big five" (openness to experience, conscientiousness, extraversion, agreeableness, emotional stability) by blue-collar workers and white-collar sector.

5. Moments that equate the distribution of initial personal traits for different age groups (20 moments):

- The difference (mean and std deviation) of "big five" (openness to experience, conscientiousness, extraversion, agreeableness, emotional stability) between the young age group (15–24) and the middle age groups (25–34, 35–44).

As previously noted, for older cohorts we only observe personality traits at later ages and have to impute the initial values. The last 20 moments are included in estimation, so that the imputed initial distributions for older cohorts closely match the observed distributions for younger cohorts.[32]

[29] We do not attempt to fit earning between ages 15–17, because there are few observations have earning information at these ages.

[30] The 4-year age groups are 15–19, 20–24, 25–29, 30–34, 35–39, 40–44.

[31] The moments conditioning on educational are constructed based on individuals who are beyond age 30 and have for the most part finished their education.

[32] This is a type of stationary assumption on the distribution of initial personality traits. Other initial conditions of older and younger cohort may still differ, though, because of differences in family background. Thus, the type distribution will not necessarily be the same across age cohorts. The model also allows for cohort effects on the sector-specific rewards.

TABLE 6. Estimates of the reward function parameters.

	1. *White Collar*		2. *Blue Collar*		3. *Schooling*		
	Parameters	S.D.	Parameters	S.D.		Parameters	S.D.
Schooling	0.0447	0.0016	0.0449	0.0026	College cost (per year)	3.9199	0.5982
White-collar experience	0.0385	0.0015	0.0174	0.0013	Graduate cost (per year)	3.7580	1.0751
Blue-collar experience	0.0105	0.0012	0.0347	0.0029	Additional utility before age 19	3.5868	0.1984
"Own" experience squared/100	−0.0387	0.0038	−0.0194	0.0029	Cognitive ability	0.2074	0.0471
"Own" experience × edu	0.0084	0.0029	0.0428	0.0062	**Constant:**		
"Own" experience ≤ 2	0.1804	0.0281	0.2870	0.0474	Type I	7.8651	0.1883
Cognitive ability	0.1028	0.0110	0.3018	0.0232	Type II	4.1195	0.2462
Standard error	0.4808	0.0212	0.4334	0.0309	Type III	1.1857	0.3083
Constant:					Type IV	7.0185	0.0842
Type I	10.3560	0.0187	9.6629	0.0260	**Cohort (Omitted cat:** 60–69)		
Type II	10.2560	0.0177	10.0050	0.0374	70–79	0.3137	0.1057
Type III	9.7467	0.0346	9.2222	0.0339	80–89	2.7839	0.1772
Type IV	9.7990	0.0264	9.0537	0.0452	90–99	3.1411	0.3954
State (Omitted cat: NSW)							
VIC	−0.0708	0.0091	−0.0893	0.0120	***4. Home-staying***		
QLD	−0.1353	0.0171	−1.0000	0.0773	Age	0.0228	0.0022
SA	−0.2336	0.1482	0.5000	0.0827	Age squared/100	0.0300	0.0045
WA	0.0412	0.0386	0.0429	0.0182	**Constant:**		
TAS	−0.2749	0.0804	−0.0778	0.0851	Type I	4.4752	0.0949
NT	−0.0453	0.0266	−0.2186	0.0983	Type II	4.1480	0.0635
ACT	0.2970	0.0959	−0.0330	2.7978	Type III	3.5069	0.0613
Cohort (Omitted cat: 60–69)					Type IV	2.6618	0.1903
70–79	0.1257	0.0110	0.1192	0.0294			
80–89	−0.1086	0.0117	0.2936	0.0151			
90–99	−0.0924	0.0172	−0.4408	0.1189			
Nonpecuniary Values					**Other Primitive Parameters**		
Constant	–		3.2188	0.0920	Std of preference shock	0.8995	0.0322
College premium	–		−2.8277	0.3796	Discount factor	0.8960	0.0100

Note: Data source: HILDA, 2001–2013. The estimates are based on 2934 males whose personality traits are measured at least one time between ages 15–44. The unit for the nonpecuniary, school, and home-staying columns is 10,000AU$.

7. ESTIMATES

7.1 *Parameter values*

Tables 6–8 show the model parameter estimates along with standard errors. Table 6 shows the parameters corresponding to the per-period reward for each of the alternatives (white-collar job, blue-collar job, schooling, and home staying). An additional year of schooling increases both white-collar and blue-collar wage offers by 4.5%. The reward for the first 2 years' work experience (exp ≤ 2) is relatively high. One year of white-collar experience increases white-collar wage offers by 18.0%, and 1 year of blue-collar experience increases blue-collar wages by 28.7%. White-collar experience has a significant return in the blue-collar sector and blue-collar experience is also rewarded in the white-collar sector. The non-pecuniary terms capture the psychic difference between working in a white-collar or a blue-collar job. We normalize the nonpecuniary utility from a white-collar job to 0. The nonpecuniary blue-collar job premium is AU$32,188 for individuals who are not college graduates but only AU$3911 (= 32,188 − 28,277) for college graduates.

For the schooling option, we estimate a utility of AU$35,868 per year if an individual stays in school until age 18; this relatively high utility is needed for the model to be able to capture the drop-off in schooling after high school graduation. We find an annual cost

Table 7. Estimated MNL type probability model coefficients.

Types	I (Baseline)	II	III	IV
Constant term	–	−1.100	−0.310	−0.480
		(0.097)	(0.083)	(0.054)
Openness to experience	–	−0.800	−0.700	0.202
		(0.137)	(0.076)	(0.112)
Conscientiousness	–	−0.490	−0.280	−0.527
		(0.088)	(0.056)	(0.073)
Extraversion	–	0.206	0.114	−0.470
		(0.083)	(0.041)	(0.058)
Agreeableness	–	−0.490	−0.500	−0.480
		(0.095)	(0.071)	(0.056)
Emotional stability	–	−0.014	−0.100	−0.019
		(0.001)	(0.021)	(0.114)
Parental education (Omitted cat: no college)				
One college	−0.500	0.030	−0.086	0.017
	(0.180)	(0.216)	(6.847)	(0.161)
Two colleges	0.138	0.169	−0.115	0.005
	(0.072)	(0.126)	(0.193)	(0.259)
Family Intactness dummy (Omitted cat: intact family)				
Living with at most	0.085	0.065	0.140	−0.035
one parent at age 14	(0.260)	(0.508)	(0.363)	(0.364)
Type persistence		Time shift term	*Age* − 15	$\frac{(Age-15)^2}{100}$
Values		0.320	0.230	1.191
		(0.025)	(0.012)	(0.110)

Note: The estimates are based on 2934 males whose personality traits are measured at least one time between 15–44 over the years 2001–2013. The standard deviations of the estimates are in parentheses.

Table 8. Estimated coefficients for the equation describing the evolution of personality traits.

Traits	*Edu*	$Edu * (Age - 15)/100$	$Age - 15$	$(Age - 15)^2/100$
Openness to experience	0.0796	−0.4025	0.0108	0.0034
	(0.0201)	(0.0691)	(0.0044)	(0.0500)
Conscientiousness	0.0616	−0.2980	0.0486	−0.0736
	(0.0215)	(0.0801)	(0.0089)	(0.0217)
Extraversion	0.0348	−0.1342	−0.0176	0.0115
	(0.0159)	(0.1308)	(0.0041)	(0.0699)
Agreeableness	0.0785	−0.6331	0.0333	−0.0263
	(0.0195)	(0.1067)	(0.0090)	(0.0530)
Emotional stability	0.0716	−0.1544	−0.0247	0.0793
	(0.0353)	(0.2438)	(0.0332)	(0.0741)

Note: The estimates are based on 2934 males whose personality traits are measured at least one time between 15–44. The standard deviations of the estimates are in parenthesis.

of college education of AU$39,199 and a yearly cost of graduate school of AU$37,580.[33] This cost would include both tuition and living expenditures as well as potential psychic costs. Also, higher cognitive ability increases the return from school attendance. There are significant birth cohort effects both on wage offers and on utility from schooling.

With regard to the home staying option, the flow utility is specified as quadratic in age and the age terms are statistically significantly different from zero. Lastly, we estimate a discount rate parameter, β, equal to 0.896 and standard deviation of the preference shock σ_c equal to 0.90.

There is considerable variation in the estimated rewards across occupations for the four types of individuals. For the two working options, types I and II have comparative advantages. Type I receives the highest reward in the white-collar occupation and type II the highest reward in the blue-collar occupation. With regard to the schooling alternative, type I gets the highest reward from attending school, followed by types IV, II, and III. The benefit for type I (AU$78,651) is slightly higher than that of type IV (AU$70,185), but much higher than for types II (AU$41,195) and III (AU$11,857). For the staying home option, the rewards associated with types I-IV are AU$44,752, AU$41,480, AU$35,069, and AU$26,618.

Table 7 shows the parameter estimates from the estimation of the type probability functions (assumed to be multinomial logistic as shown in equation (5)). The type probabilities depend on age 15 personality traits as well as family background (parental education and whether the individual grew up with both parents). A high openness to experience score implies a high probability of being type I or IV but a low probability of being type II or type III. A person with high conscientiousness is more likely to be type I or III and less likely to be types II or IV. High agreeableness leads a higher likelihood of being type I. The last two rows of Table 7 show the malleability of types over time and how types become more persistent with age. Based on our estimates, the probability of changing type is around 0.75 at age 15 then diminishes to almost 0 around age 30. In other words, the types become relatively fixed by the time an individual reaches age 30.

Table 8 shows the estimates of the process governing personality trait changes, which is assumed to potentially depend on education and age. An additional year of education at age 15 increases personality trait scores. It increases the level of openness to experience by 0.08 (std. dev. units), conscientiousness by 0.03, extraversion by 0.08, agreeableness by 0.08, and emotional stability by 0.07.[34] The negative estimated coefficient on the interaction term between education and age (γ_{2n}) implies that the effect of

[33] We compare our estimated costs with the real cost data collected in Australia. For example, a 2014 HSBS report lists a per year cost for undergraduate study as AU$42,093, which includes AU$24,081 for fees and AU$18,012 for living costs. Source: http://www.about.hsbc.com.au/news-and-media/australia-the-most-expensive-country-for-education-hsbc-report. Another official website for Australia gives annual tuition fees for Bachelor's degree, Master's degree, and Doctoral degree in the range of AU$15,000–AU$33,000, AU$20,000–AU$37,000 and AU$14,000 to AU$37,000, respectively. Source: http://www.studyinaustralia.gov.au/global/australian-education/education-costs/education-costs-in-australia.

[34] By comparison, Kassenboehmer, Leung, and Schurer (2018) find that university education increases scores on extraversion and agreeableness for students from disadvantaged backgrounds. Our sample includes individuals with both senior secondary and university education, whereas their sample focuses on individuals with university education.

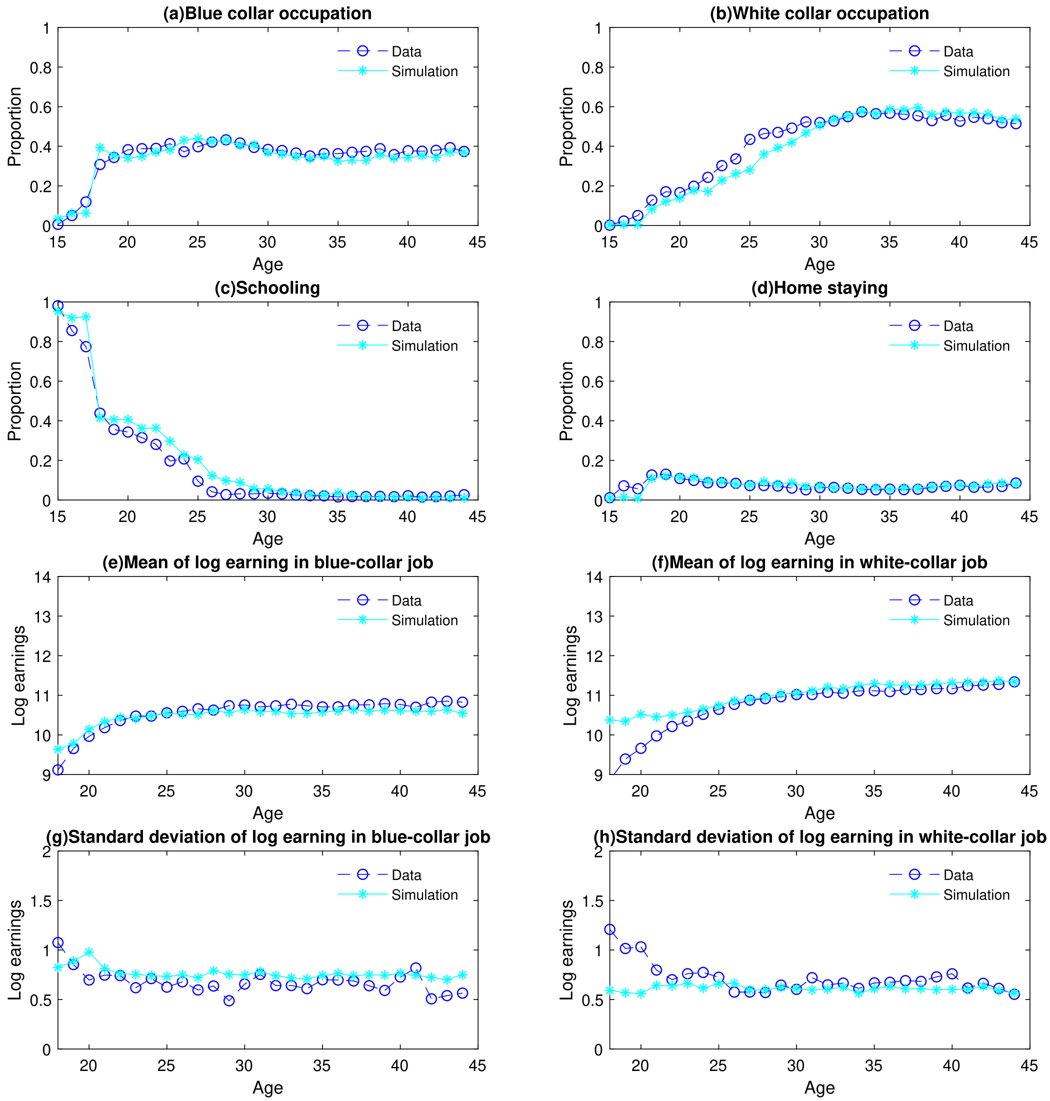

FIGURE 3. The comparison of life-cycle choices and earnings between the data and model simulations.

education diminishes with age. For example, the effect of education on conscientiousness is negligible by age 55. The age effects on conscientiousness, extraversion, and emotional stability are significantly larger than those on the other two traits. Another pattern is that extraversion decreases with age.

7.2 *Model fit*

Figure 3 compares model simulations with the data. The figure panels show the proportion choosing different sectors and the log wage of white-collar and blue-collar occupations at different ages. As seen in Figure 3, the model captures salient features of data:

(1) The fraction in blue-collar occupations exhibits an upward jump at age 18 and then declines gradually. (2) The fraction in white-collar occupations choices grows smoothly from nearly 0 at age 18, reaches its peak in the mid-30s, and then diminishes somewhat. (3) Except for a small hump shape in the early 20s, the fraction that stays home is relatively flat with a slight increase at older ages. (4) The fraction in school rapidly drops at age 18 and continues to fall until reaching a stable level around age 25. (5) The concavity of the earnings profile is also captured in our simulated sample, both for white-collar and blue-collar occupations. However, simulated log earnings for the white collar sector are too high at younger ages (recall that very few people work in that sector at young ages).

To examine whether personality traits distributions are stable across different age cohorts, we impute the initial personality trait distributions for three age groups (15–24, 25–34, 35–44) in Figure 4.[35] Figure 4 shows the personality traits distributions at the age of measurement in the left-hand figure panels and the imputed personality trait distributions at the initial age (age 15) in the right-hand side panels. We conduct a Kolmogorov–Smirnov test to test for equality of the distributions across the age cohorts. The distributions differ when measured at different ages, but we can not reject the hypothesis that people from different age cohorts share identical initial trait distributions.

8. Model simulation results

We next use the estimated model to simulate individuals' choices. First, we explore the link between personality traits, types, and choices. Second, we examine the relative importance of personality traits in explaining ex ante heterogeneity compared with other initial endowments. Third, we implement a test of the hypothesis that the unobserved types are stable over time (which is commonly assumed in the literature).

8.1 *Understanding the link between personality traits, types, and choices*

Table 9 examines the type distributions within the different sectors. White-collar workers tend to be types I or IV, whereas blue-collar workers tend to be types II and III. Also, individuals attending school are more likely to be types I and IV, possibly because longer periods of schooling are usually required to be a white-collar worker. Home-stayers are predominantly type III.

Figure 5, a radar chart, provides a graphical depiction of the average personality trait and cognition levels among types. Each equiangular spokes ("radii") represents one dimension of personality traits. Each star-like hectagon denotes the values of the "big five" along with the cognitive score for each type. It is clear that type I has the highest values of all five traits and also for cognition, because its hectagon totally covers the other three types' hectagon. It seems that high cognitive ability and high values of personality traits tend to be clustered in type I individuals. These individuals are also those that tend to acquire more schooling and to work in the white-collar sector.

[35]The algorithm for calculating the initial personality traits from observed personality traits at a given age can be found in Appendix B.

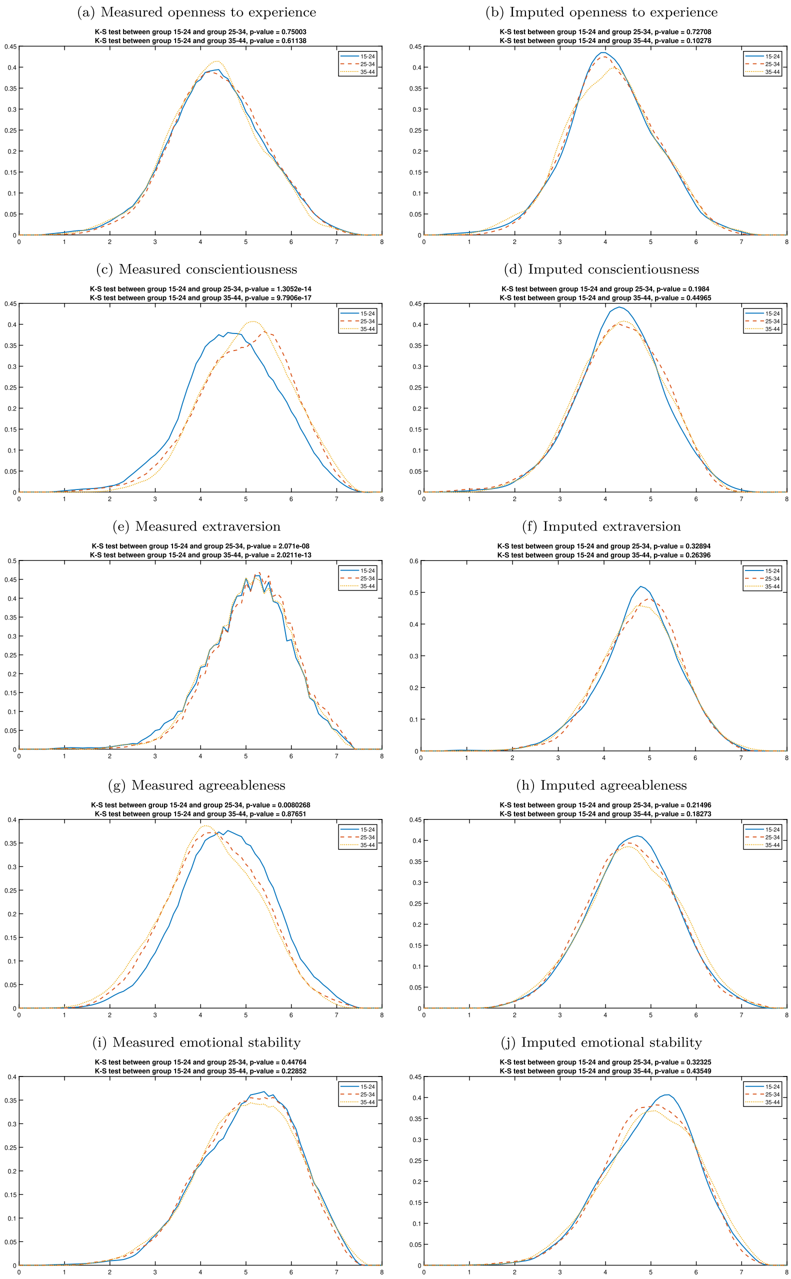

FIGURE 4. Comparison of measured and imputed personality trait initial distributions. *Note*: These figures compare the "big five" initial distribution for individuals for which it is observed (left panels) with the imputed distributions (right panels). We test equality of the distributions using a Kolmogorov–Smirnov test. The test p-values are reported on the top of each figure. Data source: HILDA 2005, 2009 and 2013, Males age 15–44 whose personality traits are measured at least once.

TABLE 9. Simulated type percentages for different sector choices.

Occupation	Type I	Type II	Type III	Type IV
White collar	44.08%	14.46%	11.80%	29.67%
Blue collar	10.66%	28.45%	47.37%	13.53%
Schooling	36.30%	12.70%	13.63%	37.37%
Home staying	2.37%	4.37%	82.55%	10.71%
Total	28.67%	17.63%	29.35%	24.35%

Figure 6 shows how the type distribution changes with age. With age, the overall type II, III, and IV proportions decrease, whereas the type I proportion increases. The type changes are driven primarily by increasing levels of conscientiousness, with age and with education (recall Table 8).

Table 10 shows the fraction of individuals who change types between ages 15 and 44, conditional on their initial type. As seen in the table, the percentage of individuals who do not change type is 52.1% for type I, 38.2% for type II, 42.5% for type III, and 45.3% for type IV. As seen in the first row and column of numbers, it is more common for individuals to switch from another type to being type I than vice versa. Individuals who are type II at age 15 are the ones most likely to switch to a different type.

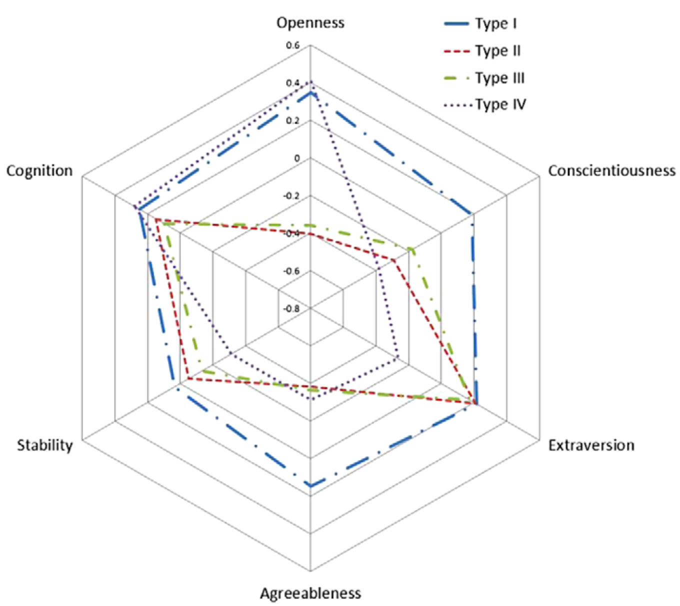

FIGURE 5. Average personality traits and cognitive ability by type. *Note*: This radar chart provides a graphical depiction of the average personality trait and cognition levels by type. Each equiangular spoke ("radii") represents one dimension of personality traits. All values of personality traits and cognitive score are standardized to be zero mean and unit standard deviation.

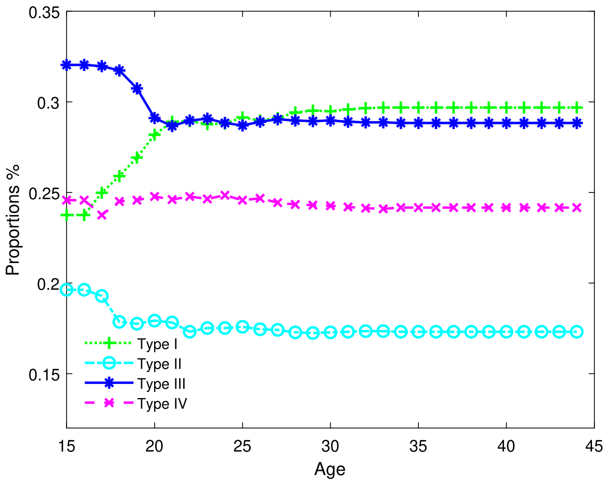

FIGURE 6. How the type distribution changes with age.

8.2 *Understanding the effect of personality traits on education, earnings, and ex ante lifetime utility*

To explore the importance of cognitive and noncognitive traits in affecting educational and labor market outcomes, we report in Table 11 the effects of a one-unit standard deviation increase in initial personality traits and cognitive abilities on earnings, education, occupational sector, ex ante utility, and type probabilities (at ages 35–40). Increases in the initial level of openness to experience, conscientiousness, or agreeableness generate the largest increases in log earnings and make it more likely that a person has a white-collar occupation. These traits also increase the sample fraction of type I's. An increase in conscientiousness and in agreeableness lead to higher levels of education, perhaps because these traits facilitate school success. An increase in the cognitive score also increases earnings, but decreases the fraction white collar and years of education. This is because the return to cognitive ability is higher in blue-collar occupations than its return in white-collar occupations. As a result, individuals are more likely to choose the blue-collar option rather than white collar or schooling options.

We find our estimated returns to personality traits and cognitive ability are comparable to the existing literature estimating the returns of noncognitive skills. For example,

TABLE 10. Type transition probabilities between ages 15 and age 44.

	Final (at Age 44)			
Initial	Type I	Type II	Type III	Type IV
Type I	0.521	0.106	0.184	0.190
Type II	0.190	0.382	0.230	0.197
Type III	0.243	0.166	0.425	0.166
Type IV	0.244	0.112	0.191	0.453

Note: Each row sums to 1.

Table 11. Simulated effects of one-unit standard deviation increase in initial personality traits and cognitive abilities.

	(Log) Annual Earning	White-Collar Occupation	Blue-Collar Occupation	Education Years	Ex Ante Utility	Type Proportions at Age 15			
						I	II	III	IV
Baseline	11.133	0.594	0.306	5.328	855,560	0.252	0.179	0.317	0.252
Opn (+1 SD)	0.039	0.096	−0.076	0.716	25,031	0.046	−0.061	−0.109	0.124
Cos (+1 SD)	0.037	0.017	−0.014	0.033	12,869	0.062	−0.024	0.009	−0.046
Ext (+1 SD)	0.004	−0.036	0.034	−0.327	−1190	0.004	0.045	0.029	−0.078
Agr (+1 SD)	0.053	0.039	−0.028	0.177	21,901	0.063	−0.009	−0.028	−0.026
Stb (+1 SD)	0.011	0.012	−0.007	0.050	5512	0.006	0.012	−0.021	0.004
All (+1 SD)	0.170	0.124	−0.094	0.576	74,882	0.242	−0.061	−0.109	−0.072
Cog (+1 SD)	0.060	−0.052	0.088	−0.435	48,325	0.000	0.000	0.000	0.000

Note: The expected ex ante utility is an Australian dollar equivalent measure at age 15. The first row shows the simulated levels under baseline model. Rows 2–8 display the deviation from baseline levels from a one standard deviation unit increase in each and all personality traits and in cognitive ability.

a one SD increase in cognitive ability increases the ex ante utility by AU\$48,325 and a one SD intervention in all of the "big-five" personality traits increases the ex ante utility by AU\$74,882. In comparison, Cunha, Karahan, and Soares (2011) find the economic value of one SD increase in cognitive skills has a range between \$25,000–135,000 and the economic value of one SD increase in noncognitive skills has a range between \$55,000–135,000.[36]

8.3 *Testing for type stability*

As previously noted, our model allows the unobserved types to change in a way that may depend on age and on personality traits. However, the Markov specification we adopted nests a model with fixed types. In this section, we perform two model specification tests:

1. Allow the type probabilities to depend on initial personality traits and on family background but restrict them not to vary with age (a restriction on the Markov matrix that $p(a)$ be an Identity matrix.)

2. Assume the types are drawn from a multinomial distribution that does not depend on regressors and that types are fixed with age (a restriction on the initial type probability distribution function and also on the Markov matrix).

We test these restrictions using a Wald test. As seen in Table 12, the fixed type models are rejected with a p-value less than 0.01.

[36]In their paper, the money values of cognitive and noncognitive skills vary with different education levels and age cohorts.

Table 12. Model specification Wald tests.

	Types Are Fixed and Determined by Initial Personality Traits	Types Are Fixed and Do Not Depend on Personality Traits
Null Hypothesis	H_0: $\frac{1}{\gamma_7}=0$	H_0: $\frac{1}{\gamma_7}=0$, $\gamma_{4kn}=0$
Wald test	157.54	1310.1
The number of restrictions	1	16
$\chi^2(0.01)$ criteria	6.64	32.00

Note: According to equation (4), $\frac{1}{\gamma_7}=0$ is a sufficient and necessary restriction to guarantee that $p(a)=0, \forall a$.

9. Two education policy experiments: Compulsory senior secondary school and a college subsidy

We next use the estimated dynamic discrete choice models, both the variable-type and the fixed-type variants, to evaluate the effects of two education policies, a college tuition subsidy program, and a compulsory schooling policy.

9.1 *Using the model to simulate the effects of educational policies*

In the late 1980s, the Australian government started providing financial assistance to students through a program called the Higher Education Contribution Scheme (HECS) and, after 2005, the Higher Education Loan Programme (HELP). With the goal of relieving the financial burden of a university education, those eligible for HECS-HELP can either receive no interest student loans or get a 10% discount on the upfront payment. Some students also receive direct financial help to cover living expenditures through a means-tested programs (such as Austudy or Youth Allowance). Motivated by these financial aid programs, we use our estimated model to simulate the effects of a hypothetical policy that reduces the college tuition cost by 50%.[37]

Our second policy experiment is motivated by the spatial variation in compulsory schooling requirements across different states and territories. The compulsory education policy in Australia is age-based. In 2009, the minimum school leaving age in Queensland, Western Australia, South Australia, and Tasmania was 17, whereas the minimum age in other areas was between 15–16.[38] In 2010, areas with lower compulsory school attendance ages came up with plans to increase the compulsory schooling level.[39] As a result, students in all states and territories are now required to stay in school until age 17 (National Report on Schooling in Australia 2011). Inspired by these policies,

[37] Our estimated annual cost of attending college is AU$39,199, which includes not only the tuition fee but also the living cost as well as potential psychic cost. We set the tuition fee to be AU$24,000 based on various reports (see footnote 30). The 50% tuition subsidy is a conditional transfer with an annual value of AU$12,000.

[38] Source: National Report on Schooling in Australia 2009.

[39] As of 2010, New South Wales, Victoria, Northern Territory, and Australian Capital Territory have requirements that local students need to complete Year 10 and then participate in education, training or employment until they turn 17.

we consider the imposition of a perfectly enforced national compulsory schooling rule that mandates individuals to stay in school until at least age 17.[40]

We next use the model to simulate the effects of the two education policies, both mean and distributional effects. To understand the importance of allowing for time-varying types, we compare the simulated policy effects for 2934 individuals obtained under the baseline model to those obtained under a restricted "fixed type" model. Table 13 shows the simulated policy effects.[41] Specifically, we examine effects on (1) the percentage high school graduates; (2) the percentage college graduates; (3) the average years of education; (4) the annual earnings for workers; and (5) the expected lifetime utility gain. In each of these categories, we first present the values under baseline model in the row labeled as "benchmark." The two rows labeled "50% college subsidy" and "compulsory senior secondary school" show the deviations from baseline values under two separated policy experiments.

Comparing the effects of two policies, two features stand out. First, the compulsory schooling policy has the most direct positive effect on the high school completion rate (+8.1 pp), whereas the college subsidy has the largest positive impact on the fraction of college graduates (+12.2 pp). Second, these two policies affect different types of individuals. The college subsidy increases the average years of completed education by 0.3–0.6 years, with the largest increase observed for type I. The compulsory school policy increases years of education by 0.24 and 0.45 years for types II and III, respectively, but has almost no effect for types I and IV.

We observe a similar pattern for labor market outcomes. Under the college subsidy intervention, types I and IV experience an average increase in annual earnings of AU$2259 and AU$2445. The increases observed for types II and III are AU$2234 and AU$1640. When implementing the compulsory schooling policy, types II and III benefit the most. The annual earnings increases of those two types are AU$1425 and AU$2197, whereas the changes for other two types are only AU$58 and AU$134. High school completion is already so prevalent among types I and IV, so few individuals of those types are affected by the compulsory schooling policy. These individuals are more likely to face the trade-off between finishing college or not and are most strongly influenced by the college subsidy policy. In terms of utility, types I and IV benefit the most from the college subsidy policy, although types II and III also benefit. All types have a negative utility change from the compulsory schooling requirement, as it represents a constraint on their choices at early ages.[42]

[40] Individuals who are younger than age 18 after the year 2009 in HILDA data should be already subject to the compulsory education policy. However, currently, the policy is not strictly enforced. The school enrollment rates for teenagers ages 15–18 are 84.9% (175/206) in year 2010, 90.0% (226/251) in the year 2011, 89.8% (211/235) in the year 2012 and 83% (176/212) in the year 2013. These enrollment rates are stable and do not significantly differ for years prior to 2009. Our baseline model estimation assumes no compulsory schooling law and we simulate the effects of a compulsory schooling law that is strictly enforced.

[41] Because the types change over time and are potentially influenced by education, we classified agents according to their initial type at age 15.

[42] We used the estimated wage equation to determine what fraction of the earnings increase is attributable to policy effects on years of schooling verses effects coming through changes in personality traits. For the college tuition subsidy policy, 20.7% comes from personality trait changes. For the compulsory

Table 13. The effect of educational policies on schooling and labor market outcomes, by type.

Model	Type I	Type II	Type III	Type IV	Total
			Percentage finishing high school (%)		
Benchmark	99.1%	91.1%	80.8%	99.5%	91.9%
50% college subsidy	0.0%	0.8%	0.1%	0.0%	0.2%
Compulsory schooling	0.9%	8.9%	19.2%	0.5%	8.1%
			Percentage college graduates		
Benchmark	49.9%	32.1%	26.2%	54.8%	40.5%
50% college subsidy	14.3%	13.5%	9.1%	12.9%	12.2%
Compulsory schooling	0.0%	2.3%	2.4%	0.3%	1.2%
			Years of education		
Benchmark	15.035	13.812	13.343	15.226	14.328
50% college subsidy	0.530	0.549	0.363	0.543	0.484
Compulsory schooling	0.012	0.243	0.446	0.018	0.193
			Annual earnings (for workers, unit: AU$)		
Benchmark	91,026	75,998	65,409	78,804	77,487
50% college subsidy	2259	2234	1640	2445	2142
Compulsory schooling	58	1425	2187	134	909
			Fraction blue collar (%)		
Benchmark	21.5%	36.3%	39.1%	22.7%	30.0%
50% college subsidy	−2.9%	−2.9%	−2.0%	−3.1%	−2.7%
Compulsory schooling	0.0%	−1.4%	−1.3%	−0.1%	−0.7%
			Fraction white collar (%)		
Benchmark	72.8%	56.2%	46.5%	70.2%	60.8%
50% college subsidy	3.4%	3.3%	1.9%	3.7%	3.0%
Compulsory schooling	0.0%	2.0%	2.7%	0.2%	1.3%
			Utility change(Unit: AU$10,000)		
Benchmark	88.932	83.129	83.24	86.822	85.556
50% college subsidy	0.950	0.696	0.723	0.965	0.836
Compulsory schooling	−0.525	−0.644	−0.662	−0.548	−0.596

Note: The rows labeled "50% college subsidy" and "compulsory senior secondary school" show deviations from benchmark values under two separate policy experiments. The annual earnings, the fraction blue collar and the fraction white collar are for workers ages 35 to 40.

Table 14 reports the effects of two policies on personality traits at age 30 (when most have completed their education). The "benchmark" row shows the average trait score of each type. The rows "50% college subsidy" and "compulsory senior secondary school" report the changes under these two policies. Of the two policies, the college tuition subsidy has the greatest effect, increasing openness to experience, conscientiousness, and emotional stability. The effects of the compulsory schooling requirement are focussed on individuals of type II and III, particularly in the areas of openness to experience, conscientiousness, and emotional stability. The estimated effects on extraversion and agreeableness are negligible.

schooling policy, 11.5% comes from personality trait changes. Decomposition results are available upon request.

Table 14. The effect of educational policies on personality traits, by type.

Model	Type I	Type II	Type III	Type IV	Total
			Openness to Experience (at age 30)		
Benchmark	4.779	3.998	4.085	4.747	4.411
50% college subsidy	0.013	0.011	0.007	0.013	0.011
Compulsory schooling	0.000	0.008	0.013	0.001	0.006
			Conscientiousness (at age 30)		
Benchmark	5.293	4.808	4.894	4.804	4.956
50% college subsidy	0.011	0.01	0.006	0.011	0.009
Compulsory schooling	0.000	0.007	0.011	0.001	0.005
			Extraversion (at age 30)		
Benchmark	4.635	4.609	4.475	4.090	4.442
50% college subsidy	0.009	0.008	0.005	0.009	0.007
Compulsory schooling	0.000	0.005	0.009	0.000	0.004
			Agreeableness (at age 30)		
Benchmark	5.483	4.966	4.996	5.142	5.150
50% college subsidy	−0.006	−0.005	−0.003	−0.005	−0.005
Compulsory schooling	0.000	−0.003	−0.006	0.000	−0.003
			Emotional Stability (at age 30)		
Benchmark	5.106	5.153	5.029	4.942	5.049
50% college subsidy	0.027	0.025	0.015	0.027	0.023
Compulsory schooling	0.001	0.017	0.028	0.001	0.012

Note: The rows labeled "50% college subsidy" and "compulsory schooling" show the deviations from benchmark values under two separate policy experiments. The calculation simulates personality traits at age 30.

Table 15 simulates the per capita cost and benefits of the two policies and explores how these policies affect earnings inequality and government tax revenues. The tuition subsidy leads to a decrease in inequality as measured by the 50/10 quantile earnings ratio and the 90/10 quantile ratios. In the model, the estimated utility is measured in Australian dollars, so we can compare the average utility gain (loss) under these two policies. The college subsidy policy increases the expected utility by AU$8400. The compulsory secondary school policy decreases the expected utility by AU$6000, because it distorts individuals' optimal choices. From the government's perspective, however, a policy that increases education will tend to increase tax revenue. We obtain a rough estimate of the additional tax revenue using the Australian income tax scheme for the year 2005–2006.[43] The row in the table labeled "tax revenue" reports the present value of average tax revenue at age 15 that the government could collect from each individual over the lifetime.[44] The additional tax revenue is AU$7600 for the college tuition subsidy intervention and AU$6400 for compulsory schooling intervention. The presented value

[43] More specifically, no tax for income below $6000; 15% for income between $6000 and $25,000; 30% for income between $25,000 and $75,000; 40% for income between $75,000 and $150,000; 45% for income over $150,000.

[44] We set the real interest rate to be 3.48% per year, which is the average interest rate over 2001–2013 (Source:world bank). We further assume each individual is capable to work until age 65 and the annual income is not changeable after age 45.

Table 15. Cost-benefit analysis of the two educational policies.

	Benchmark Case (no Policy)	50% College Subsidy	Compulsory Senior Secondary School
Earning inequality (for workers) at age 40			
50/10 earnings ratio	2.31	2.28	2.25
90/10 earnings ratio	3.77	3.67	3.66
Expected utility (Unit: AU$10,000)	85.56	86.39	84.96
Expected utility change (Unit: AU$10,000)	–	0.84	−0.60
Government expenditure (Unit: AU$10,000)	–	1.51	0.00
Exp. utility - gov exp. (Unit: AU$10,000)	–	−0.67	−0.60
Tax revenue (Unit: AU$10,000)	29.78	30.54	30.41
Increase in tax revenue (Unit: AU10,000)	–	0.76	0.64

Note: Inequality is measured by the 90/10 and 50/10 percentile earnings ratios. The row "Expected utility" reports the expected lifetime utility at age 15. The extra gain (loss) under the two policies are reported in the next row "Expected utility change." The row "Government expenditure" reports the average subsidy the government needs to pay for each individual. The differences between expected utility gain (loss) and government expenditure are reported in the next row "Exp. utility - gov exp." The row 'tax revenue" shows the expected value of average tax revenue that government would collect from each individual over the lifetime (until age 65). The changes under the two policies are reported in the next row "Increase in tax revenue."

of tuition subsidy policy is AU$15,100 per person on average (reported in the row "government expenditure").

9.2 *Exploring the importance of personality traits in explaining ex-ante lifetime utility heterogeneity*

We next explore which personality traits are the most important determinants of the variation in ex ante lifetime utility at age 15. First, we simulate lifetime ex-ante utility $\mathrm{var}(V(s(15)))$ for each individual given their age 15 initial conditions $s(15)$. Then we redo the calculation where we eliminate the variation in all except for one of the personality traits/cognitive ability (setting the values for other components at the mean sample values). Let $\mathrm{var}(V(\bar{s}(15))$ represent the variance when the heterogeneity is restricted in this fashion. In Table 16, the importance of each personality trait is represented by the fraction of variance accounted for when incorporating the variance only of that trait θ:

$$R^2_\theta = \frac{\mathrm{var}\big(V\big(\bar{s}_{-\theta}(15), \theta(15)\big)\big) - \mathrm{var}\big(V\big(\bar{s}(15)\big)}{\mathrm{var}\big(V\big(s(15)\big)\big)}.$$

As seen in the table, openness to experience and agreeableness are the most important personality traits in accounting for ex ante lifetime utility variation. In total, initial personality trait heterogeneity explains about 23.6% of the ex-ante lifetime utility variation. Doing the same calculation for cognitive ability shows that cognitive skill explains 22.9% of the total variation of ex ante utility. The combination of personality traits and cognitive ability explains 51.8% of the total variation of ex ante utility.

TABLE 16. Percent variation in ex-ante lifetime utility explained by personality and cognitive traits.

	Openness to Experience	Conscientiousness	Extraversion	Agreeableness	Emotional Stability	Big-Five in Total	Cognitive Ability	All Included
R^2_θ	0.149	0.026	0.008	0.070	0.001	0.236	0.229	0.518

Note: The table shows the percentage of variance in ex ante lifetime utility (at age 15) accounted for when we allow for heterogeneity only in one personality trait or cognitive ability (setting the values for other components equal to the mean sample values).

9.3 *Heterogeneous policy effects by family background social-economic status (SES)*

Lundberg (2013) emphasized the importance of family background in understanding the correlation between personality traits and college graduation. We therefore examine how individuals from different socioeconomic status (SES) backgrounds respond to the policy interventions. SES is defined in terms of parents' educational attainment. In group I, both parents have education equal to high school or less. In group II, one parent has some college, and in group III, both parents have some college.[45] We find the personality trait differences between individuals from different backgrounds are similar to those reported in Lundberg (2013). Individuals from more advantaged family backgrounds tend to have higher scores for conscientiousness, openness to experience, and emotional stability.

Table 17 summarizes the effects of both the college subsidy policy and the compulsory senior secondary school policy. The effect of the college tuition subsidy is substantial across all SES groups in terms of increasing education and earnings. We also observe an increase in the percentage of white-collar workers and a decrease in the percentage blue collar for all SES groups. The compulsory schooling policy similarly affects all SES groups in terms of education and earnings, although the effects are smaller in magnitude than for the tuition subsidy.

10. CONCLUSIONS

This paper develops a dynamic model of schooling and occupational choices that incorporates personality traits. As is common in the discrete choice literature, we introduce unobservable types' to capture agents' heterogeneous comparative advantages in schooling and particular occupational sectors. In line with some recent papers in the literature (Hu et al. (2015) and Arcidiacono and Miller (2011)), we adopt a specification with time-varying types, where the probability of changing type can depend on age and on personality traits. We perform a test of the assumption that types are fixed, which is rejected. Our estimates show that types are malleable when agents are young but become stable by age 30. Another finding is that high levels of cognitive skills and high personality trait scores, in all five dimensions, tend to be clustered in a certain type of

[45]We did not consider the family intactness as an additional dimension, because the majority of the sample (82.89%) grew up with both biological parents.

Table 17. The effect of educational policies on labor market outcomes by SES background.

Model Simulation	Socioeconomic Status (SES)			
	I	II	III	Total
	Percentage finishing high school			
Benchmark	89.4%	89.4%	89.4%	89.4%
50% college subsidy	0.3%	0.1%	0.2%	0.2%
Compulsory senior secondary school	10.6%	10.3%	4.3%	8.1%
	Percentage college graduates			
Benchmark	35.1%	36.3%	47.9%	40.5%
50% college subsidy	12.8%	11.9%	12.0%	12.2%
Compulsory senior secondary school	1.4%	1.1%	1.2%	1.2%
	Education years			
Benchmark	13.980	14.021	14.844	14.328
50% college subsidy	0.484	0.486	0.482	0.484
Compulsory senior secondary school	0.256	0.219	0.126	0.193
	Annual earnings (for workers)			
Benchmark	76,047	72,764	82,644	77,487
50% college subsidy	1836.9	2357.7	2137.1	2142.1
Compulsory senior secondary school	1171.7	984.26	705.83	908.67
	Fraction blue collar (%)			
Benchmark	29.1%	33.7%	27.1%	30.0%
50% college subsidy	−1.8%	−2.9%	−3.1%	−2.7%
Compulsory senior secondary school	−0.9%	−0.7%	−0.5%	−0.7%
	Fraction white collar (%)			
Benchmark	60.0%	55.6%	66.4%	60.8%
50% college subsidy	2.3%	3.1%	3.3%	3.0%
Compulsory senior secondary school	1.7%	1.5%	0.8%	1.3%
	Utility gain (Unit: AU$10,000)			
Benchmark	82.541	83.092	89.841	85.556
50% college subsidy	0.725	0.752	0.988	0.836
Compulsory senior secondary school	−0.689	−0.689	−0.447	−0.596

Note: The rows labeled "50% college subsidy" and "compulsory senior secondary school" show the deviations from baseline values under the two separate policy experiments.

individual, type I in our analysis. This type also acquires more schooling and tends to work in the white-collar sector.

Much of the prior economics literature emphasizes the role of cognitive skills in determining lifetime outcomes.[46] Our analysis shows that cognitive skills are important but also that having high cognitive skills, on average, goes hand-in-hand with having high noncognitive skills. We find that initial personality trait heterogeneity at age 15 explains roughly the same percentage of ex ante lifetime utility variation (24%) as does cognitive ability (23%). Because cognitive and noncognitive attributes tend to be positively correlated, studies that only focus on cognitive traits likely overstate their importance as a determinant of labor market success.

[46] See, for example, Neal and Johnson (1996).

Using the estimated dynamic discrete choice model, we evaluate two education policies: compulsory senior secondary school and a 50% college tuition subsidy. Both policies increase educational attainment, but their distributional effects are very different. The compulsory school policy is effective for individuals at risk for not finishing high school, represented by types II and III in the model. The college tuition subsidy effects are more evenly distributed, affecting all types. When the data are divided by SES family background, we see that both educational policies benefit individuals from all SES backgrounds.

We found that this channel is empirically important to consider when evaluating the distributional effects of education policies. The simulated policy responses are greater and the effects more evenly distributed across individuals in our sample in a varying-type model than in a fixed-type model. Moreover, we find that individuals from lower SES backgrounds are the ones most likely to change their types.

In summary, our results indicate that one of the benefits of attending school is that it changes some personality attributes, which along with increased schooling, enhances earnings. A caveat to our findings is that personality traits in our data are measured as of age 15, and they likely reflect parental investment and life experience from conception to age 15. As emphasized in Cunha, Heckman, and Schennach (2010), the most cost effective policies for fostering the accumulation of desirable personality traits may be policies that are targeted during early childhood years rather than high school or post-secondary schooling interventions. Nonetheless, we find that policies that encourage secondary school and college attendance can be effective in enhancing lifetime earnings and ameliorating inequality.

Appendix A: Identification of the transition matrix

We describe here how the Markov law of motion $\Pr(d_m(a), s_{-k}(a), k(a)|d_m(a-1), s_{-k}(a-1), k(a-1))$ can be identified based on observations from at least three consecutive time periods $\{d_m(a), s_{-k}(a), d_m(a-1), s_{-k}(a-1), d_m(a-2)\}$. The proof follows Hu et al. (2015). The main difference is that Hu et al. (2015) make a stationary assumption, whereas our assumptions are conditional on age. The proof shows that the law of motion can be uniquely decomposed into the following three components. Put it in a different way, the proof shows that the following three components are identified separately.

$$\begin{aligned}
&\Pr\bigl(d_m(a), s_{-k}(a), k(a)|d_m(a-1), s_{-k}(a-1), k(a-1)\bigr) \\
&\quad = \underbrace{\Pr\bigl(d_m(a)|s_{-k}(a), k(a)\bigr)}_{\text{CCP}} \underbrace{\Pr\bigl(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1)\bigr)}_{\text{Law of motion for type } k} \\
&\qquad \times \underbrace{\Pr\bigl(s_{-k}(a)|d_m(a-1), s_k(a-1), k(a-1)\bigr)}_{\text{Law of motion for } s_{-k}}.
\end{aligned}$$

First, we make the following assumption, which is satisfied under our model specification:

Assumption 1 (Limited feedback). (1) $\Pr(d_m(a)|s_{-k}(a), k(a), s_{-k}(a-1), k(a-1), d_m(a-1)) = \Pr(d_m(a)|s_{-k}(a), k(a))$. (2) $\Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1), d_m(a-1)) = \Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1))$.

In addition, we assume the following.

Assumption 2 (Full rank). *For any* $\{s_{-k}(a), d_m(a-1), s_k(a-1)\}$,

$$\left[\Pr\left(d_m(a)=i, s_{-k}(a), d_m(a-1)|s_{-k}(a-1), d_m(a-2)=j\right)\right]_{i,j}$$

is invertible.

Lastly, we require assumptions that given the same state space, different types have different choice probabilities.

Assumption 3 (Distinctive types). *For any two different types* k_1 *and* k_2 *at age* $a-1$, $\forall k_1, k_2 \in k(a-1)$

$$\Pr\left(d_m(a-1)|s_{-k}(a), s_{-k}(a-1), k_1(a-1)\right) \neq \Pr\left(d_m(a-1)|s_{-k}(a), s_{-k}(a-1), k_2(a-1)\right).$$

Assumption 4 (First-order stochastic dominance). $\Pr(d_m(a-1)|s_{-k}(a), s_{-k}(a-1), k(a-1))$ *is stochastically increasing in the sense of first-order stochastic increasing in* $k(a-1)$ *for fixed* $(s_{-k}(a), s_k(a-1))$.

Theorem 5. *Under Assumptions* 1, 2, 3, 4, *the density function* $\Pr(d_m(a), s_{-k}(a), d_m(a-1), s_{-k}(a-1), d_m(a-2))$ *uniquely determines the conditional probability function* $\Pr(d_m(a)|s_{-k}(a), k(a))$, *the law of motion for* k $\Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1))$, *and the law of motion for the rest state variables* $\Pr(s_{-k}(a)|d_m(a), s_{-k}(a-1), k(a-1))$.

Proof. Our proof of Theorem 5 follows Hu et al. (2015). Their proof is for the stationary Markov case, but we the theorem still holds when the conditional choice probability is age-dependent. We assume that the discrete values $\{d_m(a-2), d_m(a-1), d_m(a), s_o(a), s_o(a-1)\}$ share the common support $\{1, 2, \dots, J\}$, then introduce the following notation of J-dimensional square matrices:

$$A = \left[\Pr\left(d_m(a)=i, s_{-k}(a), d_m(a-1)|s_{-k}(a-1), d_m(a-2)=j\right)\right]_{i,j};$$

$$B = \left[\Pr\left(d_m(a)=i|s_{-k}(a), s_{-k}(a-1), k(a-1)=k\right)\right]_{i,k};$$

$$C = \left[\Pr\left(k(a-1)=k|s_{-k}(a-1), d_m(a-2)=j\right)\right]_{k,j};$$

$$D_1 = \text{diag}\left\{\left[\Pr\left(d_m(a-1)\right)|s_{-k}(a), s_{-k}(a-1), k(a-1)=k\right]_k\right\};$$

$$D_2 = \text{diag}\left\{\left[\Pr\left(s_{-k}(a)|s_{-k}(a-1), k(a-1)=k\right)\right]_k\right\};$$

$$E = \left[\Pr\left(d_m(a)=i, s_{-k}(a)|s_{-k}(a-1), d_m(a-2)=j\right)\right]_{i,j};$$

$$F = \left[\Pr\left(k(a)=l|s_{-k}(a), s_{-k}(a-1), k(a-1)=k\right)\right]_{l,k};$$

$$G = \big[\Pr\big(d_m(a) = i | s_{-k}(a), k(a) = l\big)\big]_{i,l},$$

$$H = \big[\Pr\big(s_{-k}(a) | s_{-k}(a-1), k(a-1) = k\big) \cdot \Pr\big(k(a-1) = k | s_{-k}(a-1), d_m(a-2)\big)\big]_{k,j}.$$

From the above matrices, only matrices A and E are observed. Given the matrix definitions, the following equation:

$$\begin{aligned}
&\Pr\big(d_m(a), s_{-k}(a-1), d_m(a-1) | s_{-k}(a-1), d_m(a-2)\big) \\
&\quad = \sum_{k(a-1)} \Pr\big(d_m(a) | s_{-k}(a), s_{-k}(a-1), k(a-1)\big) \\
&\qquad \times \Pr\big(d_m(a-1), s_{-k}(a) | s_{-k}(a-1), k(a-1)\big) \\
&\qquad \times \Pr\big(k(a-1) | s_{-k}(a-1), d_m(a-2)\big)
\end{aligned} \tag{14}$$

can be written as

$$A = B \cdot D_1 \cdot D_2 \cdot C.$$

Integrating over $d_m(a-1)$ in equation (14) yields

$$\begin{aligned}
&\Pr\big(d_m(a), s_{-k}(a) | s_{-k}(a-1), d_m(a-2)\big) \\
&\quad = \sum_{k(a-1)} \Pr\big(d_m(a) | s_{-k}(a), s_{-k}(a-1), k(a-1)\big) \\
&\qquad \times \Pr\big(s_{-k}(a) | s_{-k}(a-1), k(a-1)\big) \\
&\qquad \times \Pr\big(k(a-1) | s_{-k}(a-1), d_m(a-2)\big),
\end{aligned} \tag{15}$$

which is equivalent to the following matrix notation equation:

$$E = B \cdot D_2 \cdot C.$$

Given the assumption that E is invertible (Assumption 2), we can get

$$A \cdot E^{-1} = B \cdot D_1 \cdot B^{-1}.$$

Using Assumption 3, the eigenvalue-eigenvector of $A \cdot E^{-1}$ should be unique,[47] thus B is identified as the eigenvector and D_1 is identified as eigenvalues. Assumption 3 also infers that B and D_1 are both invertible, thus we have the identification of H:

$$H \equiv D_2 \cdot C = D_1^{-1} \cdot B^{-1} \cdot A.$$

Therefore, D_2 and C are identified separately under Assumption 4. □

Corollary 6. *The age-dependent conditional choice probability* $\Pr(d_m(a) | s_{-k}(a), k(a))$ *and law of motion for* $s_{-k}(a)$ *are identified nonparametrically.*

[47] The summation of each column of B should be equal to one. Thus the decomposition is unique up to this normalization constraint.

Given the identification of D_1 and D_2, we can identify

$$\begin{aligned}&\Pr\bigl(s_{-k}(a), d_m(a-1)|s_{-k}(a), k(a-1)\bigr)\\&\quad=\Pr\bigl(d_m(a-1)|s_{-k}(a), s_{-k}(a-1), k(a-1)\bigr)\Pr\bigl(s_{-k}(a)|s_{-k}(a-1), k(a-1)\bigr).\end{aligned}$$

Then the age-dependent conditional choice probability and law of motion for $s_{-k}(a)$ are two marginal distributions of $\Pr(s_{-k}(a), d_m(a-1)|s_{-k}(a), k(a-1))$:

$$\begin{aligned}&\Pr\bigl(s_{-k}(a), d_m(a-1)|s_{-k}(a), k(a-1)\bigr)\\&\quad=\underbrace{\Pr\bigl(d_m(a-1)|s_{-k}(a), k(a-1)\bigr)}_{\text{CCP}}\underbrace{\Pr\bigl(s_{-k}(a)|d_m(a-1), s_{-k}(a-1), k(a-1)\bigr)}_{\text{law of motion for } s_{-k}}.\end{aligned}$$

Corollary 7. *The law of motion for types* $\Pr(k(a)|s_{-k}(a), s_{-k}(a-1), k(a-1))$ *is also identified nonparametrically.*

F is the law of motion for k, and G is the conditional choice probability we just identified. Given $B = G \cdot F$, F can be recovered by the equation $F = G^{-1} \cdot B$. The conclusions from Corollary 6 and Corollary 7 complete the proof.

Appendix B: Method used to impute initial age 15 personality traits

In many cases, individuals are sampled for the first time at an age older than 15, so we do not directly observe their initial personality traits. The data contain up to three measures of personality traits, each measured at a time 4 years apart. We next describe the method that we use to impute the initial personality traits $\theta_n(15)$ based on these three measures, $\theta_n^{M1}(a_1)$, $\theta_n^{M2}(a_2)$, $\theta_n^{M3}(a_3)$, observed at ages a_1, a_2, a_3 and using the structure of our model. Given the current trial parameter values Ω, personality trait n at age 15 ($\theta_n(15)$) is obtained as follows:

1. From equation (2) in Section 4.1, we solve the projection of initial personality $\theta_n(15)$ based on the measures on age a_1 $\theta_n(a_1)$:

$$\theta_n(15) = \theta_n(a_1) - \bigl(\gamma_{0n} + \gamma_{1n} g(a_1) + \gamma_{2n}(a-15)g(a) + \gamma_{3n}(a-15) + \gamma_{4n}(a-15)^2\bigr),$$

where a_1 is the age when individual's personality trait θ_n is surveyed and $g(a)$ is the accumulative education years at age a_1.

2. Substituting $\theta_n(a_1) = \theta_n^{M1}(a_1) - \zeta_n(a_1)$, where $\zeta_n(a_1)$ is the unobserved measurement error at age a_1 with mean 0. Then

$$\theta_n(15) = \theta_n^{M1}(a_1) - \bigl(\gamma_{0n} + \gamma_{1n} g(a_1) + \gamma_{2n}(a-15)g(a) + \gamma_{3n}(a-15) + \gamma_{4n}(a-15)^2\bigr) - \zeta_n(a_1).$$

3. Define $\theta_n^{M1}(15) \equiv \theta_n(15) + \zeta_n(a_1)$, its value could be directly calculated by

$$\theta_n^{M1}(15) = \theta_n^{M1}(a_1) - \bigl(\gamma_{0n} + \gamma_{1n} g(a_1) + \gamma_{2n}(a-15)g(a) + \gamma_{3n}(a-15) + \gamma_{4n}(a-15)^2\bigr).$$

4. For the other two personality measurements at age a_2 and age a_3, ($\theta_n^{M2}(a_2)$ and $\theta_n^{M3}(a_3)$) repeat steps (1)–(3) to get

$$\theta_n^{M2}(15) \equiv \theta_n(15) + \zeta_n(a_2), \theta_n^{M3}(15) \equiv \theta_n(15) + \zeta_n(a_3).$$

5. This procedure provides three different imputed values of initial personality traits, each with a measurement error that is assumed to be mean 0 drawn from an i.i.d. distribution. We obtain our measure of the personality trait at age 15 $\theta_n(15)$ as the mean of these three values:

$$\theta_n(15) = \frac{1}{3}\big(\theta_n^{M1}(15) + \theta_n^{M2}(15) + \theta_n^{M3}(15)\big).$$

References

Agostinelli, F. and M. Wiswall (2016), "Estimating the technology of children's skill formation." NBER Working Paper No. 22442. [235]

Almlund, M., A. L. Duckworth, J. J. Heckman, and T. D. Kautz (2011), "Personality psychology and economics." NBER Working Paper No. 16822. [232, 239]

Altonji, J. G. and L. M. Segal (1996), "Small-sample bias in gmm estimation of covariance structures." *Journal of Business & Economic Statistics*, 14 (3), 353–366. [250]

Arcidiacono, P. and R. A. Miller (2011), "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity." *Econometrica*, 79, 1823–1867. [233, 242, 266]

Becker, G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education*. London. [232]

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008), "The economics and psychology of personality traits." *Journal of Human Resources*, 43 (4), 972–1059. [233]

Chan, M. K. (2013), "A dynamic model of welfare reform." *Econometrica*, 81 (3), 941–1001. [249]

Cobb-Clark, D. A. and S. Schurer (2012), "The stability of big-five personality traits." *Economics Letters*, 115 (1), 11–15. [239]

Cubel, M., A. Nuevo-Chiquero, S. Sanchez-Pages, and M. Vidal-Fernandez (2016), "Do personality traits affect productivity? Evidence from the laboratory." *The Economic Journal*, 126 (592), F654–F681. [234]

Cunha, F. and J. Heckman (2007), "The technology of skill formation." *American Economic Review*, 97 (2), 31–47. [235, 239]

Cunha, F. and J. J. Heckman (2008), "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." *Journal of Human Resources*, 43 (4), 738–782. [235, 240]

Cunha, F. and J. J. Heckman (2009), "Investing in our young people." *Rivista Internazionale di Scienze Sociali*, 117 (3/4), 387–417. [240]

Cunha, F., J. J. Heckman, and S. M. Schennach (2010), "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica*, 78 (3), 883–931. [235, 239, 268]

Cunha, F., F. Karahan, and I. Soares (2011), "Returns to skills and the college premium." *Journal of Money, Credit and Banking*, 43, 39–86. [260]

Dahmann, S. and S. Anger (2014), "The impact of education on personality: Evidence from a German high school reform." IZA Discussion Paper No. 8139. [235]

Del Boca, D., C. Flinn, and M. Wiswall (2014), "Household choices and child development." *The Review of Economic Studies*, 81 (1), 137–185. [250]

Department of Social Services, Melbourne Institute of Applied Economic and Social Research (2018), "The household, income and labour dynamics in Australia (hilda) survey, restricted release 16 (waves 1–16)." [235]

Elkins, R. K., S. C. Kassenboehmer, and S. Schurer (2017), "The stability of personality traits in adolescence and young adulthood." *Journal of Economic Psychology*, 60, 37–52. [240]

Fletcher, J. M. (2013), "The effects of personality traits on adult labor market outcomes: Evidence from siblings." *Journal of Economic Behavior & Organization*, 89, 122–135. [234]

Goldberg, L. R. (1992), "The development of markers for the big-five factor structure." *Psychological Assessment*, 4 (1), 26. [232]

Gosling, S. D., P. J. Rentfrow, and W. B. Swann (2003), "A very brief measure of the big-five personality domains." *Journal of Research in Personality*, 37 (6), 504–528. [232]

Gourieroux, C. and A. Monfort (1996), *Simulation-Based Econometric Methods*. Oxford University Press. [250]

Griliches, Z. (1977), "Estimating the returns to schooling: Some econometric problems." *Econometrica: Journal of the Econometric Society*, 1–22. [232]

Heckman, J. and B. Singer (1984), "A method for minimizing the impact of distributional assumptions in econometric models for duration data." *Econometrica: Journal of the Econometric Society*, 52, 271–320. [246]

Heckman, J. J. (1981), *The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process*. MIT Press. [246]

Heckman, J. J. and B. E. Honore (1990), "The empirical content of the Roy model." *Econometrica: Journal of the Econometric Society*, 1121–1149. [247]

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010), "The rate of return to the highscope perry preschool program." *Journal of Public Economics*, 94 (1), 114–128. [232]

Heckman, J. J. and L. K. Raut (2016), "Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model." *Journal of Econometrics*, 191 (1), 164–175. [235]

Heckman, J. J., J. Stixrud, and S. Urzua (2006), "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics*, 24 (3), 411–482. [235]

Horowitz, J. (1981), "Identification and diagnosis of specification errors in the multinomial logit model." *Transportation Research Part B: Methodological*, 15 (5), 345–360. [247]

Hu, Y., M. Shum, W. Tan, and R. Xiao (2015), "A simple estimator for dynamic models with serially correlated unobservables." *Journal of Econometric Methods*, 6, 345–360. [233, 242, 246, 247, 266, 268, 269]

Huggett, M., G. Ventura, and A. Yaron (2011), "Sources of lifetime inequality." *The American Economic Review*, 2923–2954. [232]

Kassenboehmer, S. C., F. Leung, and S. Schurer (2018), "University education and non-cognitive skill development." *Oxford Economic Papers*, 70 (2), 538–562. [235, 254]

Keane, M. P. and R. M. Sauer (2010), "A computationally practical simulation estimation algorithm for dynamic panel data models with unobserved endogenous state variables." *International Economic Review*, 51 (4), 925–958. [250]

Keane, M. P. and K. I. Wolpin (1997), "The career decisions of young men." *Journal of Political Economy*, 105 (3), 473–522. [232, 233, 234, 242, 244, 246]

Keane, M. P. and K. I. Wolpin (2001), "The effect of parental transfers and borrowing constraints on educational attainment." *International Economic Review*, 42 (4), 1051–1103. [250]

Kennan, J. and J. R. Walker (2011), "The effect of expected income on individual migration decisions." *Econometrica*, 79 (1), 211–251. [249]

Losoncz, I. (2009), "Personality traits in hilda." *Australian Social Policy*, 8, 169–198. [236]

Lundberg, S. (2013), "The college type: Personality and educational inequality." *Journal of Labor Economics*, 31 (3), 421–441. [235, 266]

Mullins, J. (2019), "Designing cash transfers in the presence of children's human capital formation." Job Market Paper. [235]

Neal, D. A. and W. R. Johnson (1996), "The role of premarket factors in black-white wage differences." *Journal of Political Economy*, 104 (5), 869–895. [267]

Nyhus, E. K. and E. Pons (2005), "The effects of personality on earnings." *Journal of Economic Psychology*, 26 (3), 363–384. [234]

Rust, J. (1987), "Optimal replacement of gmc bus engines: An empirical model of Harold Zurcher." *Econometrica: Journal of the Econometric Society*, 999–1033. [249]

Salgado, J. F. (1997), "The five factor model of personality and job performance in the European community." *Journal of Applied Psychology*, 82 (1), 30. [234]

Saucier, G. (1994), "Mini-markers: A brief version of Goldberg's unipolar big-five markers." *Journal of Personality Assessment*, 63 (3), 506–516. [232]

Schurer, S. (2017), "Does education strengthen the life skills of adolescents?" IZA World of Labor. [235]

Sullivan, P. (2010), "A dynamic analysis of educational attainment, occupational choices, and job search." *International Economic Review*, 51 (1), 289–317. [232, 233]

Terracciano, A., P. T. Costa, and R. R. McCrae (2006), "Personality plasticity after age 30." *Personality and Social Psychology Bulletin*, 32 (8), 999–1009. [233, 239]

Terracciano, A., R. R. McCrae, and P. T. Costa (2010), "Intra-individual change in personality stability and age." *Journal of Research in Personality*, 44 (1), 31–37. [233, 239]

Todd, P. E, and W. Zhang (2020), "Supplement to 'A dynamic model of personality, schooling, and occupational choice'." *Quantitative Economics Supplemental Material*, 11, https://doi.org/10.3982/QE890. [235, 236]

Wooden, M. (2013), "The measurement of cognitive ability in wave 12 of the hilda survey." Melbourne Institute Working Paper Series No. 44/13. [236]

Yamaguchi, S. (2012), "Tasks and heterogeneous human capital." *Journal of Labor Economics*, 30 (1), 1–53. [232, 233]

Co-editor Peter Arcidiacono handled this manuscript.

Manuscript received 26 May, 2017; final version accepted 27 May, 2019; available online 25 September, 2019.