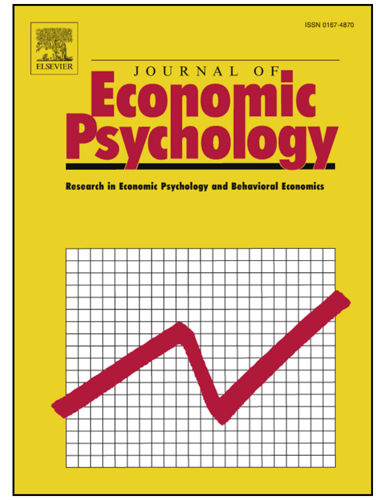

Personality Traits, Forgone Health Care and High School Dropout: Evidence from US Adolescents

Giuseppe Migali, Eugenio Zucchelli

PII: S0167-4870(16)30084-8
DOI: http://dx.doi.org/10.1016/j.joep.2017.06.007
Reference: JOEP 2020

To appear in: *Journal of Economic Psychology*

Received Date: 24 February 2016
Accepted Date: 16 June 2017

Personality Traits, Forgone Health Care and High School Dropout: Evidence from US Adolescents

Giuseppe Migali*
Eugenio Zucchelli†

June 23, 2017

JEL Classification: C21, D03, I12, I21
Keywords: personality traits; forgone health care; high school dropout; multiple treatment matching.

Abstract

There is sparse evidence on the effects of personality traits on high school dropout, especially on whether combinations of different traits may uniquely influence dropout decisions. We employ single and multiple treatment matching together with rich data on US adolescents to explore the relationship between personality traits and their combinations on school attrition. Using the Big Five inventory, we find that introversion, and to a lesser extent neuroticism, are individually associated with higher probabilities of dropping out from school. Multiple treatment estimates show that blends of low levels of conscientiousness and neuroticism present higher likelihoods of an early exit. Furthermore, we exploit information on forgone health care and explore its role as a predictor of dropout, potentially proxying relevant traits associated with psychological maturity of judgement such as responsibility, perspective and temperance. These traits refer to the capacity of assessing the long-term consequences of actions and may influence an individual's decision-making process, including dropout choices. Forgone health care appears to be a statistically significant predictor of dropout throughout our models. Individuals who forgo their health care and present low conscientiousness and introversion

*Corresponding author: g.migali@lancaster.ac.uk, Department of Economics, Lancaster University Management School, Lancaster LA1 4YX, UK. Tel: +44(0)1524594221. Dipartimento S.G.S.E.S., Universita' Magna Graecia, Catanzaro, Italy.

†Division of Health Research (DHR), Lancaster University, UK. Email: e.zucchelli@lancaster.ac.uk.

have the highest risk of dropout. Overall, our results are robust to alternative specifications and increasing levels of selection on unobservables. We suggest that given its predictive power, forgone health care could be used as a signalling device to help identifying individuals at higher risk of school dropout.

JEL Classification: C21, D03, I12, I21
Keywords: personality traits; forgone health care; high school dropout; multiple treatment matching.

1 Introduction

Dropping out of high school is still a major policy issue that affects more than 20% of young pupils in most OECD countries (OECD, 2012). Early dropout substantially increases the risk of unemployment, leads to lower lifetime earnings and is linked to a number of adverse outcomes later on in life, including poorer health status and behaviours such as violence and crime (e.g. Chapman et al., 2010; Thornberry et al., 1985). Heckman and LaFontaine (2010) find that high school dropout in the US might have been underestimated due to inconsistencies in the measurement of high school graduation rates. Comparable data and methods suggest that estimates on graduate rates have been substantially biased upward and the actual dropout rate in the US increased slightly during the last decades and should be around 12%.

The economic literature has identified a number of important determinants of high school dropout. These include demographic characteristics such as ethnicity and gender, parental background, cognitive skills and individual preferences. However, while an emerging literature indicates that noncognitive traits may influence individual behaviours related to education, there is sparse evidence on the effects of personality traits on high school dropout (Almlund et al., 2011). Furthermore, earlier contributions focus on the effects of single traits, such as locus of control, and produce contrasting results (e.g. Coleman and DeLeire, 2003; Cebi, 2007). In line with the most recent literature, we think that personality traits might play an important role in determining educational choices. However, differently from previous works, we focus on the effects of several traits and their combinations (personality types or styles) on high school dropout.

Moreover, while some of the skills and preferences which determine high school dropout could be observed and influenced, some important individual traits that may also predict dropout might be difficult to identify and tackle (Eckstein and Wolpin, 1999). In the presence of multiple and potentially unobserved noncognitive traits, it is important to find proxies of those traits that could be strong predictors of dropout decisions. More specifically, the personality psychology literature indicates that personality traits may be powerful predictors of health and health-related behaviours (Almlund et al., 2011; Kern and Friedman, 2011). For instance, previous studies (e.g. Vollrath et al., 1999) suggest that some of the Big Five personality traits, such as conscientiousness and agreeableness, are strongly correlated to decreased consumption of tobacco and alcohol. Therefore, finding proxies of further unobserved traits might help capturing relevant personality features and exploring their roles in determining educational choices.

Accordingly, we also extend our analysis on the effects of heterogeneous personality styles on high school dropout, by investigating a potential proxy of unobserved personality traits. To be specific, we exploit unique information on forgone health care, a widespread yet overlooked health-related behaviour among adolescents, and test it as a predictor of high school dropout.

According to the medical literature (e.g. Ford et al., 1999; Ginsburg et al., 1995), health can be forgone either by the inability to access health care or by voluntary avoidance, given there is a perceived need. Previous studies on health care utilisation suggest that forgone care might be the result of objective circumstances, such as access to health insurance and household financial burdens, or of an individual's predisposition to use health care services (Ford et al., 1999; Wisk and Witt, 2012). This individual predisposition depends on individual characteristics, including beliefs and noncognitive traits: some adolescents may lack the needed psychological maturity of judgement to effectively use health care (Ford et al., 1999; Dubow et al., 1990). Maturity of judgement has been linked to traits such as responsibility, temperance, perspective (Steinberg and Cauffman, 1996) as well as self-control (Donnellan et al., 2007). Since all these traits may be relevant to educational choices, we believe forgone health care could be an important predictor of dropout behaviour.

Hence, the objective of this paper is twofold. First, we estimate the effects of personality traits on early high school attrition. We exploit rich individual-level data on high school pupils, including the Big Five personality traits from the US National Longitudinal Study of Adolescent to Adult Health (Add Health). We employ linear probability models (LPM) as well as single and multiple treatment non-parametric propensity score matching (PSM). While LPM explore the relative roles of established determinants of dropout together with personality traits via standard correlations, PSM methods allow to produce more reliable estimates through the use of comparable treated and control groups without imposing stringent parametric assumptions. In our single treatment PSM models, introversion and, to a lesser extent, neuroticism are individually associated with a higher probability of dropping out from school. Multiple treatment estimates show that combinations of low conscientiousness and neuroticism present a higher likelihood of an early exit. Secondly, we extend these models by including forgone health care and consistently find it is a statistically significant predictor of high school dropout. Also, individuals who forgo their health care and present lower levels of conscientiousness and introversion have the highest risk of leaving school. We suggest that given its relevance, forgone health care could be potentially used as a signalling device to help targeting individuals at hight risk of dropping out. Overall, our results appear to be robust to selection on unobservables (Altonji et al., 2005).

This paper offers several contributions to the literature. To the best of our knowledge, this is the first paper focusing on the effects of a range of personality traits on high school dropout. While the psychology literature highlights the importance of assessing the effect of personality styles (Terracciano and Costa, 2004; Weiss et al., 2009), so far this has not been tested on school dropout, especially within the economic literature. Furthermore, we employ multiple treatment matching models to explore whether the impact on dropout varies according to heterogenous personality styles. Finally, we investigate the use of forgone health care as a predictor of dropout. Although previous studies have analysed the relationship between health-behaviours and educational outcomes (Suhrcke and de Paz Nieves, 2011), forgone health care has never been linked to dropout behaviour.

2 Background

Our work bridges three strands of literature: the economic determinants of high school dropout, studies on the effects of noncognitive skills on dropout decisions, and the literature on forgone health care among adolescents.

2.1 Determinants of high school dropout

In standard economics models, education is seen as an investment decision where individuals choose their optimal level of education by weighting the potential returns from obtaining a degree against the effort and costs needed to obtain it. The decision of dropping out from high school has been analysed mainly in the light of two competing theories: human capital model and signalling theory (Bedard, 2001). According to the human capital framework, education augments natural abilities that are subsequently sold in the labour market. Signalling models suggest that education could also act as a signalling or screening device for unobserved abilities: firms infer abilities from students' education levels. Hence, in signalling theory, the earnings reward from high school graduation is a combination of the increase in human capital and the effects of being identified as a graduate (or 'higher-ability' student signal in the labour market). From an individual's perspective, whether education enhances human capital or acts as a signal, dropout decisions would still depend on the established positive correlation between education and earnings upon which both theories are based on. As a result, human capital and signalling mechanisms could equally support our analysis on the individual determinants of dropout.

We choose to focus on high school dropout specifically because of the size of the phenomenon in OECD countries as well as its potential long-term negative effects at individual and societal level (OECD, 2012). A large body of evidence (De Witte et al., 2013) finds that adolescents who do not complete high school have substantially lower lifetime earnings, resulting in significant lost tax revenues. School dropout is also frequently associated with serious physical and mental health issues in adulthood and is consistently linked to higher crime rates.

The empirical literature has identified a number of important determinants of high school dropout. These are mainly: gender; ethnicity; time preferences; parental characteristics such as parents' educational attainment, social status and single parenthood (Bratti, 2007; Ermish and

Francesconi, 2001; Mocetti, 2008; Oreopoulos, 2007).[1] According to these studies, individuals at higher risk of high school dropout are males, Hispanic or black; have low academic or cognitive skills; come from disadvantaged or low educated parental backgrounds and heavily discount future consequences of present choices. Eckstein and Wolpin (1999) develop and estimate a sequential structural model of school attendance and work decisions. Importantly, they find that pupils who drop out from high school appear to have different traits than those who graduate. These include: lower school ability and motivation; lower expectations from graduation; higher value of leisure and a lower consumption value of school attendance. The majority of these studies focus on either observable sociodemographic characteristics or cognitive skills and very few of them include noncognitive abilities.

2.2 Noncognitive skills and high school dropout

An increasing number of studies are exploring the role of noncognitive skills in predicting relevant economic outcomes, including academic achievement, labour supply and earnings (Almlund et al., 2011). Whilst there is no standard terminology across disciplines, most economists employ terms such as noncognitive traits to broadly define individual skills other than cognitive abilities (Heckman and Rubinstein, 2001; Borghans et al., 2008) and often use noncognitive skills and personality traits interchangeably.[2] For the purpose of this study, we focus on the effects of personality traits as defined by the Big Five inventory (Goldberg, 1993; Barenbaum and David, 2008) on high school completion. This is because, despite a growing literature in both psychology and economics on the relationship between noncognitive skills and educational attainment, evidence on the effects of personality traits on the specific decision to drop out from high school is still sparse and mixed.[3]

Heckman and Rubinstein (2001) point out that in the US low returns of General Education Development tests (GED) may be due to the lack of noncognitive skills of GED holders, as defined by the high incidence of behaviours such as drug use, violence and shoplifting. They suggest that, given the quantitative importance of noncognitive traits, social policies should actively attempt to alter them. Furthermore, the authors suggest that standard signalling models in economics should

[1] More recently, Stinebrickner and Stinebrickner (2012) also show the importance of revised expectations in dropout decisions at college level: learning about academic ability and performances through grades may increase the probability of college dropout.

[2] However, personality psychology appears to employ a more precise definition of personality traits and define them as "patterns of thought, feelings, and behavior", see e.g. Roberts (2009). Noncognitive skills and personality traits are also thought to be closely related to individual preferences such as time and risk preferences, although preference parameters should formally fulfil the criteria for rationality within economic models. For a detailed discussion on the complex relationships between personality traits, preferences and constrains within economics models as well as its mathematical formalisation, see section 3 of Almlund et al. (2011).

[3] For a recent and comprehensive overview on the relationship between noncognitive skills and economic outcomes, including educational attainment, also see Almlund et al. (2011).

also account for noncognitive skills. Accordingly, we follow Heckman and Rubinstein (2001) and include GED holders in our defintion of dropouts due to their similarities in noncognitive skills.

Previous studies have explored the effects of locus of control on years of schooling.[4] Coleman and DeLeire (2003) find that locus of control measured in 8th grade affects high school graduation by influencing an individual's expectations on the returns to human capital investment. Their results imply that adolescents with an internal locus of control (i.e. who believe they have some degree of control over life events) should be more likely to invest in higher education. However, Cebi (2007) finds that locus of control does not appear to be an important determinant of educational outcomes once cognitive abilities are controlled for. Barón and Cobb-Clark (2010) observe that 18-years old individuals with more locus of control have a higher probability of completing secondary school. Coneus et al. (2011) consider the joint effects of academic skills and internal locus of control and show that both reduce school dropout. Interestingly, they also observe that the influence of locus of control appears to increase with age. Although these papers explore the relevance of noncognitive traits on educational choices, they focus exclusively on the effects of locus of control, present overall mixed results and consider only a relatively limited number of covariates.[5]

Heckman and other authors have also produced a series of studies on the economics of cognitive and noncognitive skills formation (e.g. Cunha and Heckman, 2008; Cunha et al., 2010), as well as on the role of noncognitive skills in the development of health inequalities (e.g. Heckman and Kautz, 2012; Heckman and Masterov, 2007). These suggest that noncognitive skills promote the development of cognitive abilities, whereas the impact of cognitive abilities on noncognitive skills seems more limited. Furthermore, over an individual's life-cycle, noncognitive skills are malleable for longer periods than cognitive skills. Remediation policies, for example those aimed at individuals from disadvantaged family backgrounds, and early life interventions — the Perry Preschool project (Schweinhart, 2003), or EPIS programme, a Portuguese intervention aimed at boosting the noncognitive skills of students between 7-9th grade, (Martins, 2010) — should focus on developing noncognitive skills and should be preferred to interventions later on in life.[6]

[4]Locus of control is a trait often related to emotional stability/neuroticism and measures the extent to which an individual believes his actions would affect life events (Rotter, 1966).

[5]More specifically, Coleman and DeLeire (2003) employ US data from the National Educational Longitudinal Study and consider basic demographic characteristics, cognitive skills and family characteristics. Cebi (2007) uses the US National Longitudinal Survey of Youth (NLSY) and exploits demographic variables, cognitive skills, family characteristics, home life and geographical variables. Barón and Cobb-Clark (2010) use information from the Australian Youth in Focus (YIF) Project and account for demographic, cognitive, and family-related variables together with year of birth. Coneus et al. (2011) draw data from the German Socio-Economic Panel (SOEP) and pay particular attention to variables defining family background.

[6]According to these studies, it would be more efficient to invest into individual noncognitive skills at early 'critical' periods of human development. Policy interventions implemented later on in the life-cycle, and hence during potentially less critical periods of human development, would not be able to compensate for the loss in skills development.

Also of relevance to this paper, a separate strand of studies within the psychology literature highlights the importance of exploring the effects of combinations of personality traits, known as personality styles or types. For example, Terracciano and Costa (2004) examine the effects of different blends of the Big Five on smoking. They find that individuals with personality styles combining low conscientiousness and high neuroticism are more likely to smoke. Similarly, according to Weiss et al. (2009) high neuroticism with (low or high) extraversion or high openness are highly correlated with major depression. We build on these findings and investigate whether the interaction between traits may uniquely influence school dropout decisions.

2.3 Forgone health care and psychological maturity of judgement

Forgone health care

According to the medical literature, between 17 to 20 percent of adolescents worldwide do not access health care when needed (Ford et al., 1999; Denny et al., 2013). Forgone health care has been associated with objective circumstances that may restrict an individual's ability to access health care services, such as economic deprivation (low household income) and health insurance type, especially in the US.

Other studies, e.g. Lehrer et al. (2007), link forgone health care to confidentiality concerns and risky health behaviours, such as birth control non-use. Ford et al. (1999) employ data from Add Health to analyse forgone health care among adolescents in the US. They conclude that together with continuous access to health insurance, age and ethnicity, there are other important factors that increase the probability of reporting forgone health care. These are individual behaviours, such as daily cigarette use, frequent alcohol consumption and sexual intercourse.

Although results from previous studies on the size and significance of the main determinants of forgone health care may vary, they all recognise the relevance of an individual's propensity to access health care services. This propensity appears to be determined by both the ability to secure access to health care service (circumstances) and individual beliefs and traits which translate into a series of behaviours. Notably, Dubow et al. (1990) and Ford et al. (1999) also suggest that adolescents might not be psychologically mature enough to make effective decisions about health care use. In this paper, since we are able to control for a comprehensive set of variables concerned with circumstances, including access to health care and health insurance, we employ forgone health care as a proxy of psychological maturity of judgement.

Psychological maturity of judgement

Within the specialised literature (e.g. Cauffman and Steinberg (2000)), psychological maturity of judgement refers to a individual's decision-making process and how this is affected by a range of cognitive and psychosocial factors. In this case, ”judgement” concerns the decision-making process per se rather than a specific outcome, while ”maturity” refers to how this process changes with an individual’s overall development. As such, an individual can reveal poor judgment because of either intellectual or psychological shortcomings or both. Accordingly, a young individual who forgoes health care may exhibit low psychological maturity of judgement because of cognitive or psychosocial limitations.

Importantly, maturity of judgement has been previously associated with psychosocial traits such as responsibility, perspective and temperance (Steinberg and Cauffman, 1996; Cauffman and Steinberg, 2000) and with personality traits like self-control (Donnellan et al., 2007). Responsibility and perspective encompass independence and the capacity of considering situations from different viewpoints, including placing choices in broader societal and temporal contexts. Temperance includes tendencies to limit impulsivity and evaluating situations before making decisions. While we are not able to directly observe these traits, they may be relevant predictors of educational outcomes as they all refer to the propensity of considering the long-term consequences of actions. Similarly, aspects of self-control relate to relevant parameters within economic models of investment in education (e.g. myopia and time preferences). As a result, here we investigate whether lower levels of psychological maturity of judgement, proxied by forgone health care, might have an effect on high school dropout.

More specifically, since we account for a standard measure of cognitive skills, forgone health care should help identifying further psychosocial factors driving poor judgement. Moreover, because we also include proxies for time and risk preferences, forgone health care should represent the remaining traits suggested by the literature, i.e. responsibility, perspective and temperance.

3 Data and descriptive statistics

We employ data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a panel study of a nationally representative sample of US high school students in grades 7-12. This cohort has been followed through adolescence and transition into adulthood, until individuals in the sample are aged 24-32, with four in-home interviews.

The Add Health sample has a school-based design and includes 132 schools stratified by region, urban area, size, school type and ethnicity to ensure representativeness among US schools.[7] Add Health includes detailed information on respondents’ social, economic, psychological and physical well-being together with data on family, neighbourhood, community and schools. Data are collected through four main questionnaires: the school questionnaire, the school-administrator questionnaire, the in-home questionnaire and the parent questionnaire.

[7] For further details on the sampling strategy, see Harris et al. (2009).

The in-school questionnaire was administrated between 1994-1995 to 90,000 students from all schools in the sample and includes information mainly on school context and activities, friendship networks and a series of health conditions. Further school context data (e.g. school policies) were collected in the same period and are included in the school-administrator questionnaire (reported usually by schools' principals).

Our main source of information is the core sample of the wave I in-home questionnaire. This includes 12,105 students randomly selected from the 132 schools (approximately 200 students per school). Supplemental samples based on ethnicity (Cuban, Puerto Rican and Chinese), adoption status and disability were also added to this core sample using information from the in-school questionnaire. Furthermore, African-American students with highly educated parents were also oversampled leading to a total of 20,745 individuals for the wave I in-home sample. The latter forms the basis for the subsequent longitudinal follow ups (wave II: 1996; wave III: 2001/2002; wave IV: 2008). The parent questionnaire provides data on marriage, health-related behaviours, education, employment, and household income and is completed by around 80 per cent of the parents (usually the resident mother) of adolescents responding to the wave I in-home questionnaire.

For the purpose of our analysis on dropout behaviour, we focus on pupils enrolled in high school. This corresponds to all the adolescents interviewed in wave I. In addition, we combine data from wave I with retrospective information on education decisions and degrees completion collected in waves II, III and IV. We then merge these individual-level data with school-level characteristics (school size and type) in order to separately identify each school and their main characteristics. We draw information on parental background, relatives' health-behaviours and health insurance from the parent questionnaire.

3.1 Dependent variable

Our dependent variable is High School (HS) dropout:

$$\begin{cases} Y_i = 1, & \text{if individual i is a HS dropout;} \\ Y_i = 0, & \text{if individual i is a HS graduate.} \end{cases}$$

More specifically, our definition of dropout includes adolescents that report 'dropout' or 'other non-graduate' in their high school exit status in wave III. We cross-check this information with the highest education level reported in wave IV, and further include in our definition of dropout those pupils who indicated '8th grade or less', 'some high school', 'did not earn diploma, GED or equivalent certificate'. Since GED holders do not complete high school graduation and so they must have dropped out earlier, we include them in our definition of dropout. However, as a robustness check we also perform our analysis by excluding GED holders and restricting our sample to 'pure' dropout (see Table 7). Our reference category, HS graduates, includes adolescents that have at least a high school diploma.[8]

3.2 Personality traits and forgone health care

Definitions of personality traits

Following Young and Beaujean (2011) we combine information from 13 questions in wave I to build three of the Big Five PT: conscientiousness (a measure of reliability and dutifulness), neuroticism (a measure of anxiety and emotional liability) and extraversion (a measure of enthusiasm toward life's circumstances). Young and Beaujean systematically searched the wave I in-home and in-school questionnaires in Add Health for items that matched statements from the International Personality Item Pool (IPIP) version of the NEO Personality Inventory (NEO-PI-R) (see Costa and McCrae, 1992; Goldberg et al., 2006). [9]

In our models, conscientiousness is defined through a dummy variable taking value 1 if an individual answers 'disagree/strongly disagree' to at least three out of four questions related to positive aspects of conscientiousness and 0 otherwise.[10] This variable identifies low levels of conscientiousness for ease of interpretation and comparison with the other variable of interest, FHC. That is, based on the existing literature on PT and education, we expect low conscientiousness to be positively associated with dropout behaviour. Similarly, the binary indicator for neuroticism becomes 1 when a pupil answers 'agree /strongly agree' to at least five out of six items as identified by Young and Beaujean. As a result, this variable identifies emotional instability (or high neuroticism). Extraversion (defined negatively as low extraversion or introversion for ease of comparison with the effect of FHC) is obtained through a dummy variable taking value 1 when an individual

[8] We have cross-checked high school exit status, (year of) high school diploma and year of GED/Equivalent certificate/high school degree in wave III, with highest education level, high school graduation status and most recent degree/certificate in wave IV. We have identified some inconsistencies in the education data across waves and decided to exclude all adolescents with inconsistently reported education status (i.e. this occurred when we were not able to match between waves consistent information on high school dropout or graduation). In particular, we have excluded individuals who did not report their high school completion status by wave III. Most of them are likely to be 'stopouts', that is individuals who interrupted their studies and eventually obtained a qualification afterwards.

[9] These 13 items show internal consistency estimates similar to the ones of the shortened version of the IPIP NEO-PI-R, i.e. the NEO-Five Factor Inventory (NEO-FFI), and good predictive validity. All items and corresponding questions in the IPIP NEO-PI-R are reported in details in Appendix B.

[10] We employ binary variables to define PT in order to use single and multiple treatment matching and further explore the relationships between single PT and their combinations on dropout. These matching methods are based on binary treatments and for consistency we use the same definitions of personality traits in our linear probability models. Further details about matching can be found in section 4.2. Furthermore, we have also employed definitions of PT based on a continuous scale. These were built by summing row values from answers of the items within each question. In this case, estimates of the magnitude of the effects of PT would be more difficult to interpret (as these would not be based on a cardinal scale) and we would not be able to use our dummy variable adjustment approach (described in section 4.1), therefore loosing observations.

reports 'disagree/strongly disagree' in at least two of three questions related to extraversion.[11]

Following recent developments within the personality psychology literature (e.g Weiss et al., 2009), we also investigate the effects of personality styles on high school dropout. Although there is currently no evidence on their influence on dropout behaviour, based on previous studies on the impact of single traits, we hypothesise that combinations including low conscientiousness could increase the risk of dropout.

As an alternative to the wave I PT, we also estimate models with the full set of the Big Five using information available in wave IV (2008). We follow Donnellan et al. (2006), and more recently Lundberg (2013), and employ the 'mini-IPIP' 20-items measure of the Big Five.[12] Accordingly, we build dummies for the Big Five using negative answers to at least three out of four questions for each trait. This produces further variables for agreeableness (coded negatively as antagonism/hostility) and openness (coded as closeness to experience). We believe that for the purpose of our analysis, wave I personality traits are the most appropriate measures of personality, because they have the advantage of being collected before the decision of dropping out from high school (i.e. when pupils are still enrolled in high school).[13]

Definition of forgone health care

Following the medical literature, forgone health care is defined using the wave I question '*Has there been any time over the past year whether you thought you should get medical care but you did not?*' (yes/no).[14] We also use the follow-up question '*If yes, what kept you from seeing a health professional when you really needed to?*', and exclude all individuals that reported "objective circumstances" (117 among the 5,448 individuals who forgo health care) and reported *'no transport'; 'no one to go with'; 'parents would not go'; 'could not pay'*. We include in our definition of FHC all the remaining answers that did not relate to strictly objective circumstances and imply some degree of choice. These are: *'did not want my parents to know'; 'afraid of what the doctor will do'; 'I thought that the problem will go away'; 'did not know who to see'; 'hard to make appointment';* and *'other reasons'*. Accordingly, this definition of FHC should exclude objective impediments and serious health conditions. Furthermore, we remove from our analysis all

[11]Previous evidence on the effects of neuroticism and introversion on human capital accumulation present mixed results and may depend on the specific educational outcome considered, see e.g. (Almlund et al., 2011). Therefore, we do not have firm ex-ante expectations about the direction of their impact on high school dropout.

[12]The 'mini-IPIP' uses 20 of the original 50-items (i.e four for each of the five traits) to define the Big Five. This is considered a reliable alternative to the full definition.

[13]The literature suggests that personality traits may still be malleable at younger ages. Hence, personality traits measured in adulthood may have been partly altered by experience and personal development. However, recent empirical studies find that personality traits should be stable among working adults, see e.g. Cobb-Clark and Schurer (2012), Brown and Taylor (2014).

[14]Since this question does not specify which type of medical care should have been sought, we assume that this refers to any form of health care, whether formal or informal.

adolescents with chronic health conditions (e.g. diabetes) as we believe that their behaviour should be systematically different. Since the two questions on FHC were asked in wave I, all respondents were still enrolled at school.

In addition, we include in our regression models a wide range of relevant controls. These are presence and type of health insurance, ease of access to medical care and existence of school-based or -linked health clinics, which should also help accounting for practical constraints in accessing health care.[15] We control for a series of physical and mental health conditions to account for the direct influence of ill-health on high school dropout. We include variables defining time preferences (high discount rate), an individual's general attitude towards risk as well as risky-health behaviours (for both adolescents and members of their households). All these controls may be correlated with forgone health care and have also been found to be important determinants of dropout. Furthermore, since forgone health care might be potentially influenced by parental background, we account for education levels and job type of both fathers and mothers. Finally, as health care consumption might be related to the ability of processing health information and advice (sometimes defined as "health literacy", e.g. Nutbeam (2008)), we include a general measure of cognitive skills as well the availability of school based health education courses.

3.3 Descriptive statistics

Descriptive statistics for the full set of individual, family and school-level variables included in our models are reported in Table 1 and Table 2, respectively. Appendix B reports detailed definitions of all variables, including personality traits.

Our initial sample used for the estimation of LPM includes 18,765 observations. Table 1 reports descriptive statistics for the dependent variable (HS dropout) and our main main variables of interest, PT and FHC. This table also includes standard t-tests for the equality of means of all variables between dropout and graduates. These tests appear to indicate significant differences between the two samples for virtually all variables apart from receiving counselling and being obese. HS dropouts are around 8% of our sample.[16] Looking at wave I PT, we notice that a slightly higher percentage of dropouts present low levels of conscientiousness if compared to graduates (16.8%

[15]As reported in Table 1 and Appendix B, we include in our models variables for not being covered by health insurance, being under Medicaid or Medicare support and other health insurance covers. Medicare is a federal health insurance programme designed for people over 65 years old or with disabilities and people with end stage renal disease (ESRD) which covers specific hospital treatments for children. Medicaid is a social health care programme for individuals and families with limited income and resources that include dependent children. Standard insurance plans (used as baseline category in our models) purchased from health insurance companies offer family and dependent children coverage. For further information, see www.medicare.gov and www.medicaid.gov.

[16]It should be noted that this figure cannot be directly compared with annual average dropout rates in the US as students in our sample belong to different cohorts and do not all drop out in the same year. Furthermore, in the wave I in-home sample a number of ethnic minority adolescents were purposely oversampled, including 1,547 African-American students with highly educated parents. This may also reduce the total number of dropouts in Add Health.

versus 14.6%) and that dropouts are more likely than graduates to be neurotic (11.3% vs 7.7%). Low levels of extraversion (introversion) are also more prevalent among dropouts (22.3% vs 17%). Wave IV traits variables confirm similar percentages for the corresponding PT among dropouts together with lower levels of openness to experience and agreeableness (antagonism/hostility). Overall, slightly more than 27% of students do not seek health care when needed. This percentage is substantially higher among dropouts (around 36%).[17]

We account for cognitive abilities using the Add Health Picture Vocabulary Test (AHPVT). The AHPVT is an abridged computerised version of the Peabody Vocabulary Test (PVT), a well-established measure of general cognitive skills. Moreover, we control for learning disabilities (suffered by 22.3% of dropouts) and risky health-behaviours. The latter are captured by frequent smoking, drinking, marijuana use and consumption of other drugs.[18] As expected, all risky health-behaviors consistently present higher percentages among dropouts. We also include obesity (body bass index, BMI, greater than 30), which is concentrated among individuals who do not complete high school (8.1 vs 6.3%). In the attempt to fully capture individuals' preferences towards risks, we further control for a more general variable defining an individual's attitude towards risk (risky attitude) via non-use of seat-belt and/or birth controls. We notice that a risky attitude is more prevalent among dropouts (around 54.4%).

Following Oreopoulos (2007), who suggests that adolescents may heavily discount the future consequences of dropping out from school, we use a dummy variable capturing high rates of time preferences/discounting. We observe that around 26% of dropouts place a larger value on current as opposed to future utility, a percentage twice as high as that of graduates.

As religion might be associated with relevant aspects of high school completion such as conscientiousness (Saroglou, 2002) and effort, we employ a dummy variable (religious) to control for pupils reporting any religion. In our sample, dropouts are slightly less religious than graduates (7.8% versus 8.6%).

We control for the direct effects of both physical and mental health on dropout decisions using a wide range of health conditions. It is interesting to note that while only around 12% of dropouts report general fair or poor health, nearly 18% report being depressed. We further include another binary measure of general health based on the frequency of school absences due to health or

[17] We have also implemented simple (t-test) to explore pairwise correlations between PT and FHC. The resulting correlations show only a small degree of association between FHC and three of the Big Five personality traits, built using wave I information. Interestingly, low conscientiousness presents the smallest correlation with FHC (around 2%). Low extraversion is also only marginally correlated with FHC (around 4.3%). Neuroticism shows a slightly higher correlation (around 9.7%).

[18] As the main objective of the paper is to identify variables which are relevant for dropout decisions, we chose to focus on frequent or heavy consumption of tobacco and drugs. These should also reflect individual time and risk preferences and are thus relevant in economic models of dropout decisions. Detailed definitions can be found in Appendix B.

emotional problems. This should capture any relevant physical or psychological issue preventing regular school attendance (about 13% among dropouts).

Since health care utilisation may depend on health insurance, we control for parents' access to health insurance and its type. A substantial number of dropouts' parents (around 27%) are covered by Medicare/Medicaid while about 20.5% do not have any health insurance.

T-tests for the equality of means in Table 2 also indicate significant differences for the majority of variables related to parental socioeconomic background (parents' educational attainment and job status) and health as well as school characteristics. Table 2 shows that more than 47% of dropouts have low educated mothers and fathers. We also observe that between 29 and 41% of dropouts have either an unemployed/at home mother or father.

Similarly to the variable defining general health for pupils, we include two dummy variables that capture self-reported ill-health of the main parent and an assessment of his/her partner's health. We observe that nearly 21% of dropouts have their main parent in poor health conditions. Moreover, we include in our models parents' difficulties in accessing health care and health-behaviours of family members, which concern 58% of dropouts.

We also make use of school contextual data and include information on school grade spans (i.e. grades offered) and school size, type, school health policies and specific health related programmes.

4 Econometric methods

Our empirical strategy exploits the wealth of observables in Add Health through linear probability models (LPM) and propensity score matching (PSM), but also accounts for selection on unobservables by employing a test suggested by Altonji et al. (2005).

First, we estimate a series of LPM to explore the effects of PT together with established determinants of dropout on high school attrition. While these models have the advantage of including a wide range of observables based on individual, parental and school-level characteristics, they are only capable of identifying standard linear correlations and do not account directly for reverse causality or unobservables. Therefore, we employ non-parametric single treatment matching methods to further investigate the relationship between each personality trait and school dropout. These techniques improve over LPM as they do not impose assumptions concerning functional forms, i.e. they do not restrict the relationship between PT and dropout to be linear. They also produce more reliable estimates based on samples of treated and control groups which are as comparable as possible in terms of observable characteristics. Furthermore, we estimate multiple treatment matching specifications to establish whether the estimated effects on dropout vary by heterogeneous personality styles, that is by different combinations of PT. Secondly, we also separetely estimate all LPM and matching models by including FHC to examine its role as a proxy of psychological maturity of judgement. Finally, since LPM and matching solely exploit observables, we follow the recent literature (e.g. Johnston et al., 2013) and test whether our results are affected by the presence of selection on unobservables.

4.1 Linear probability models

We first estimate a succession of more comprehensive LPM with an incremental number of explanatory variables that may affect dropout behaviour. Our basic specification is:

$$E[Y_i|PT, X] = a + \alpha_s + \alpha_t + \beta_j PT_{ij} + \gamma \mathbf{X_i}. \tag{1}$$

where $i = 1 \dots N$ are HS pupils; $j = 1 \dots J$ are personality traits, PT; $s = 1 \dots S$ are school identifiers; and $t = 1 \dots T$ correspond to years of birth. α_s are school fixed effects which control for any school-specific factor that might be also correlated with our variables of interest.[19] α_t are cohort effects, which are built using individuals' years of birth (i.e. from 1975 to 1983). These account for any cohort-specific effects that relate to the number of births, economic context and resources available, and that may uniquely shape an individual's school experience and also impact on PT. $\mathbf{X}$ is a vector that includes the full set of observed variables described above. We focus on the estimation of the parameters β_j, the ones capturing the effects of PT. We first estimate models which only include PT and subsequently identical specifications augmented by FHC.[20]

4.2 Propensity score matching

Our main empirical strategy is based on PSM methods with single and multiple treatment (see Imbens, 2000; Lechner, 2001). This approach exploits the wealth of observables in Add Health and does not require any exclusion restriction. In our single treatment PSM models we consider the impact of low levels of conscientiousness, extraversion and neuroticism, separately.

[19] We also estimate equation 1 by replacing school fixed effects with variables that account for school-level heterogeneity (e.g. type, size, location etc.).

[20] We deal with missing values in our sample by employing a dummy variable adjustment method (Allison, 2000). This simply translates into adding a dummy variable that equals 1 when the observations for a variable are missing, 0 otherwise. We repeat this method for each of the categorical variables presenting a large portion of missing observations. More specifically, we apply this method to the variables drawn from the parent questionnaire and the ones defining waves I and IV noncognitive traits (see Tables 1 and 2). The advantage of this approach is twofold: it allows us to retain a consistent sample size throughout different specifications while simultaneously controlling for additional sources of noise. It should be noted that this method may produce biased estimates if data are not missing at random (Allison, 2000; Cohen et al., 2013). However, a simple t-test for the coefficients of the missing data dummies points out if data are genuinely missing at random. In any case, we have also estimated the full battery of models by dropping all missing observations and results are available upon request.

Multiple treatment models allow investigating heterogenous personality styles or types.[21] The main purpose of PSM is building a group of non-treated pupils (our control group) who are similar to the treated in all relevant pretreatment observable characteristics, X, the only remaining difference being the treatment. To allow for the presence of one or multiple traits, the treatment variable can take one of K discrete values. Hence, this multiple treatment approach can be seen as a combination of several binary treatments. We estimate the $K(K-1)/2$ binary conditional probabilities that are used as propensity scores in the matching model. The reference category for all single binary treatments and their combinations corresponds to the case where individuals report having all "positive" PT traits (conscientiousness, no neuroticism, extraversion).

Matching methods are based on two fundamental assumptions: (i) *common support*, i.e. pupils with the same pretreatment characteristics have a positive probability of reporting a personality trait of type k; (ii) *conditional independence assumption, CIA*, i.e. selection into treatment (i.e. PT, whether single or combined) is based solely on observables characteristics. The latter means that conditional on a set of pretreatment observable characteristics, our potential outcomes (dropout vs non dropout) are independent of assignment to treatment. If these assumptions hold, we can estimate an average treatment effect, which is obtained as the mean difference in outcomes over the common support, weighted by the propensity score distribution of pupils:

$$\theta^{k,j}_{ATE} = \sum_{i \in k} (Y_{ik} - \sum_{h \in j} W_{ih} Y_{hj}) \, w_i \tag{2}$$

where W_{ih} and w_i are weights used to construct the counterfactuals and the outcome distribution for the treated, respectively. Matching estimators differ in the way they construct the weights, W_{ih}: in this case, we implement a nearest neighbour algorithm and provide analytical standard errors (Abadie and Imbens, 2008).[22]

4.3 Test for selection on unobservables

Although based on a wide range of observables, LPM and matching models do not explicitly account for individual-level unobserved heterogeneity that may potentially bias our estimates. Therefore, we check the robustness of all our results by employing a standard test for selection on unobservables (Altonji et al., 2005; Chatterji et al., 2007; Johnston et al., 2013).[23] This test essentially

[21] We present results for single and multiple PSM models first for PT and subsequently for PT and FHC jointly.

[22] We have also estimated PSM models using a kernel algorithm and bootstrapped standard errors. Since these results were virtually identical, in this paper we only show those obtained by nearest neighbour matching.

[23] This should account for remaining (unobserved) individual characteristics that were not captured by our specifications and may still influence the relationship between PT and dropout. For example, additional omitted variables related to innate abilities, motivation and effort or family connections, but also measurement errors in the included covariates.

consists in making different assumptions about the correlation between the unobservables that determine high school dropout and the ones influencing PT and FHC. This simply translates into the estimation of a series of bivariate probit models. A first probit model defines dropout as a function of the full battery of observable determinants, whereas a second probit defines sequentially each PT and FHC as a function of the same covariates. Here, identification relies on the constrained correlation coefficient as well as on functional form. In each bivariate model we impose increasing levels of correlation between the unobservables of the two equations. We start from no correlation and impose increasing levels of correlations up to a threshold, which corresponds to the correlation between HS dropout and each specific PT (or FHC) in linear models without other covariates (Johnston et al., 2013).

5 Results

5.1 Personality traits and school dropout

Linear probability models

Table 3 reports our initial set of estimates on the effects of PT on high school dropout. These are obtained from LPM that comprise individual characteristics, school fixed effects and cohort effects. The first column of Table 3 includes only PT. The effects of neuroticism and low extraversion on dropout are positive, statistically significant and substantial in size (3.0 and 1.6 percentage points, henceforth pp, respectively). However, low conscientiousness does not appear to be statistically significant. Hence, in this model neuroticism and introversion are important risk factors for dropout decisions.

The second column adds controls for demographic characteristics and skills, behaviours and preferences. Here, the estimated coefficients for neuroticism and introversion are still positive, both statistically significant at 5% and similar in magnitude (1.6pp). Moreover, health-behaviours such as frequent/heavy consumption of tobacco and marijuana, and to a smaller extent obesity, appear to be relevant and statistically significant risk factors for school dropout (around 10pp for smoking and 6pp for marijuana use). The third column presents a specification which includes health conditions to control for the direct effect of ill-health on dropout and information on health insurance. While still statistically significant, the effect of low extraversion reduces to around 1.2pp and the one of neuroticism becomes not statistically significant. The last column also contains parental and family characteristics and none of the PT are statistically significant.[24]

[24]For editing purposes, the effects of parental variables are not displayed in this table. These are virtually identical to the ones subsequently presented in Table 6 which includes the same specification with the sole addition of FHC. In any case, full results of both specifications without and with FHC can be found in Table A1.

Most of the remaining covariates have the expected signs and their quantitative effects are similar throughout all models. For example, in line with the literature males have a higher probability of dropping out compared to females (between 1.3 and 2.1pp depending on the models). Individuals with higher levels of cognitive skills (scores of PVT) have a slightly lower probability of dropping out (0.2pp) whereas the ones with learning disabilities are at a much higher risk of leaving school (between 3.9 and 4.7pp). Adolescents with a high discount factor (i.e. that 'do not give much thought about the future') and those with a propensity toward risk ('no birth control use' and/or 'no seat belt use') are also more likely to drop out.

Compared to the previous literature, the negative effect of being African-American on dropout may appear counterintuitive. However, Add Health purposely over-sampled African-American students from highly educated families and this may justify the direction of the effect.[25]

Similarly to previous evidence, individuals with an Asian background show a lower probability of dropping out while the opposite is observed for Hispanics (although the corresponding estimate is only statistically significant in the first model). Focusing on the full specification (column four), we find that self-reported ill-health does not appear to play a major role in dropping out decisions, whereas depression (2.6pp), migraine (1.6pp) and school absence due to health reasons (4.7pp) increase the probability of dropping out and are all statistically significant.

The effects of absence and type of health insurance are also significant and quantitatively relevant: not having a health insurance increases the probability of early attrition by 2.9pp while being covered by either Medicare or Medicaid by 6.0pp. This is not surprising as these variables may be also considered proxies of low income and job status.[26]

As a robustness check and to exploit the full extent of information available in Add Health on the Big Five, models in Table 4 include the wave IV full set of PT. All specifications account for cohort and school fixed effects and are presented without (first column) and with individual and family characteristics as additional covariates. Most of the Big Five have the expected effects on HS dropout: positive for neuroticism, low extraversion, openness to experience (defined as closeness) and (low) agreeableness. These effects are all statistically significant and relatively large in size, although they seem to decrease slightly with the addition of individual and family-level covariates (except for introversion, constant at 1.4pp). Low conscientiousness does not seem to play a substantial role in high school dropout decisions. In any case, these effects should be considered with caution, as the questions used to define these PT are collected when individuals have already

[25] In Table A2 we report the estimates of our preferred model by employing the Add Health weight components.

[26] Since some of the variables included in these models, especially those derived from the parent questionnaire, have a large number of missing observations, we adjust our models including additional categories for missing values. We observe that most of these dummies are not statistically significant. This suggests that observations are likely to be missing at random (Cohen et al., 2013). For ease of interpretation, we did not report estimates for the dummy variables used to control for missing values. The full set of estimates is available upon request.

entered adulthood (i.e. between 2008-2009). Hence, wave IV PT may potentially differ if compared to their wave I counterparts.

Matching models

Table 5 presents population average treatment effects (ATE) for single and multiple treatment matching models with wave I PT. In this case, ATE are preferred to average treatment effects on the treated (ATT) as they allow estimating the effects of PT on dropout at the mean of the outcome distribution. Thus, matching models yield estimates that can be compared to the ones of the LPM.

Single and double treatments are reported in the upper part of the table. More specifically, single treatment effects can be found on the main diagonal that runs from the top left to the bottom right of the upper part of the table, while double treatments fill the remaining cells.

The lower part of the table reports a triple treatment for the joint effect of the three wave I PT. Estimates for multiple treatment models were obtained by employing different combinations of binary treatments, with each treatment representing a personality trait. Hence, multiple treatments correspond to heterogeneous personality styles/types.

To evaluate the quality of the matching, we report in the first three columns of Table A3 the results of balancing tests for the pretreatment characteristics included in the propensity score (estimated using a logit model) for low levels of conscientiousness. More specifically, we include gender; ethnicity; cognitive abilities; type of health insurance coverage; family and parental characteristics (i.e. father's and mother's education and job type together with family health-behaviours and access to health care). Overall, standardised differences and variance ratios indicate a good balance for the matched sample.[27]

Looking at single treatment models, only introversion shows a positive and highly statistically significant ATE of 2.3pp. Therefore, this appears to be an individually relevant trait for the decision to drop out from school. The single treatment effect of neuroticism is 2pp and only weakly significant while the one for low conscientiousness is not statistically significant. Overall, single treatment effects seem in line with those of LPM, although they show slightly larger quantitative effects. Interestingly, the double treatment effect combining low conscientiousness with neuroticism is positive and statistically significant (4.1pp). This suggests that while these two traits do not appear to be strong predictors if considered separately, their combination increase the risk of dropping out. The triple treatment effect for the joint effects of all PT is not statistically significant. This underlines once again the relevance of considering the effects of personality styles on dropout instead of analysing PT in isolation.

[27] Variance ratios of treatment over control, in Table A3, should equal 1 in the presence of perfect balance. Ratios appear close to unity for all variables after matching. Results of the logit models and for balancing tests for the other PT and their multiple combinations are similar and available upon request.

5.2 Forgone health care and school dropout

This section includes LPM and PSM specifications which test FHC as a predictor of dropout and further robustness checks, each comprising FHC as well. These include the most complete LPM with all individual level and family characteristics; a stratified analysis by parental socioeconomic status; models which replace school fixed effects with schools' observable characteristics; a specification that excludes GED holders; and further single and multiple treatment PSM models to explore personality styles which account for FHC.

Table 6 displays our most comprehensive LPM specification (see column 4 of Table 3) augmented by FHC (full specification is available in Table A1, column 2). Although this model confirms the limited effects of wave I PT, it also shows a positive and highly statistically significant effect of FHC (1.3pp). In this case, FHC appears to be an important determinant of high school dropout. In line with the latest literature, adolescents with higher educated mothers (high school: -2.8pp; higher education: -4.3pp) and fathers (high school: -2.9pp; higher education: -3.1pp) are less likely to drop out from school. We also observe an increase in the probability of dropping out when at least one of the following behaviours is present: main parent is a smoker or a heavy drinker or another member of the family is a smoker (family health-behaviours, 2.4pp).

In the remaining columns of Table 6, we test whether the effects of FHC and PT are exacerbated by different socioeconomic status. We do this by stratifying our sample by parental education (mother's low and high education, columns 3 and 4, and father's low and high education, columns 5 and 6).[28] These estimates show stronger effects of FHC for pupils whose parents have no education or primary/middle school degrees (for low educated mothers, the effect of FHC is positive, statistically significant and 3.9pp; for low educated fathers, the quantitative effect of FHC is 4.7pp). Thus, the effect of FHC is larger for adolescents with a disadvantaged parental background. This may imply that, although we include in our specifications comprehensive information on access to health care and health insurance, these effects might still be influenced by further financial difficulties in accessing health care. Finally, introversion and neuroticism appear to become significant risk factors in some of these sub-samples: introversion becomes positive and significant (1.5pp, high educated fathers) while neuroticism only weakly significant (-3.6pp, low educated fathers).

To explore further the roles of both FHC and PT, we estimate LPM models which make use of school contextual data and employ a school identifier to include school fixed effects.[29] The

[28]Estimates for samples stratified by type of parental employment are available upon request. Results for models with interactions between PT and FHC with each level of parental education are qualitatively identical and are also available upon request.

[29]These include information on school grade spans (i.e. grades offered); school size; school locations and geographical position; school health policies and programmes; and the provision of state-funded community health and school-based or school-linked health clinics. School programmes include state-recommended courses on either dietary behaviours and nutrition; disease prevention and control; emotional and mental health; HIV prevention; human sexuality; injury prevention and safety; personal health; pregnancy prevention; or sexually transmitted disease prevention.

first column of Table 7 presents FHC and wave I PT together with the observed school-level variables and cohort fixed effects. We find that the probability to drop out decreases if adolescents are enrolled in schools of larger size (size II and III, that often provide more options and services to students), in non-public schools (e.g. private) and in those located in suburban areas. The probability of dropping out from school decreases by 1.3pp in the presence of school-based health clinics (although this estimate is only significant at 10%) while it is not affected by state programmes of health education. The estimated coefficient of FHC is positive, statistically significant and its quantitative effect is 3.3pp. Neuroticism is also highly significant and its effect is around 2.6pp. We find a positive effect of low levels of extraversion (1.2pp), yet this is only weakly statistically significant. Conscientiousness does not appear to be an important predictor of dropout. In the second column we further add controls for individual and parental characteristics. In this model, only FHC is statistically significant and, consistently with the estimates of the previous models, it increases the probability of dropping out by 1.7pp. Overall, this suggests once more that access to health care (school health clinics) does not greatly influence the effect of forgone health care on dropout decisions. None of the PT appear to be statistically significant.

Since there is still an open debate and some sources of data (e.g. US Census Bureau) do not discriminate between GED holders and regular HS graduates, we test the robustness of our findings by excluding future GED holders and focussing only on 'pure' dropout (see third column in Table 7).[30] Only FHC is statistically significant with a quantitative effects of around 1pp. This might also support the idea that FHC does not depend on cognitive skills.[31]

To evaluate the effects of personality styles further, we provide estimates for single and multiple treatment PSM models with FHC as an additional treatment (Table 8). The upper part of the table includes a single treatment model for FHC and double treatments of FHC paired with all wave I PT. The middle part shows triple treatments obtained from combinations of FHC and wave I PT. As a further robustness check, the lower part of the table presents a single treatment model for FHC estimated in the absence of health-behaviours. This should test whether the effects of FHC is affected, or partially driven, by the presence of other risky health-behaviours.[32] The ATE for FHC in the single treatment model is positive and highly statistically significant with a magnitude of 2.1pp. All double treatment models with FHC present positive and statistically significant ATE,

[30] The total number of GED in the full sample is 611, this implies a substantial reduction in the number of dropouts, from 1,513 to 902.

[31] Moreover, to further explore whether the noncognitive determinants of dropouts are similar to those of GED recipients, we have also estimated models just for GED holders, i.e. where our binary dependent variable equals 1 for GED holders and 0 for HS graduates. Although based on few GED holders (only 611), these models find that while all PT present no statistically significant effects, FHC has a small (0.008) but significant (at 5%) effect on the probability to obtain a GED. Full results are available upon request.

[32] We could not perform multiple treatment models with FHC and other PT after excluding all health-behaviours because we do not have enough individuals in the treatment group (i.e. presenting simultaneously FHC, a given PT and no other risky health-behaviours).

with the largest effect observed for the combination of FHC and neuroticism (2.9pp). Conversely, the blend of FHC, low conscientiousness and introversion is the only triple treatment effect which is statistically significant and also presents the largest quantitative effect (5.0pp). Finally, the effect of FHC does not appear to be related to the presence of other risky health-behaviours: when we exclude them the ATE of FHC is 2.8pp and still highly statistically significant.

5.3 Selection on unobservables

It is important to test whether our results are robust to the potential presence of unobserved heterogeneity. Thus, Table 9 reports a test that measures the effects on our main variables of interest (PT and FHC) of different degrees of selection on unobservables. We follow Johnston et al. (2013) and vary the level of the correlation coefficients (ρ) between the unobservables in a series of binary probit models. We first set ρ at 0 (no correlation), and then increase this up to the estimated correlation between dropout and each PT or FHC in linear models without other covariates (our upper threshold, see the OLS estimates in the first column of Table 9).

In the absence of correlation between the unobservables ($\rho = 0$), the effect of FHC is positive and statistically significant at 3.3pp, while the effects of PT vary from 1.4 to 3.1pp. For increasing levels of correlation, which imply larger effects of unobservables, all estimated effects of FHC and PT, apart from the one of low conscientiousness, are still strongly statistically significant although slightly decreasing in magnitude. The disappearance of the effect of low conscientiousness for higher levels of unobservables is in line with our previous results. These showed a limited and inconsistent effect of low conscientiousness on dropout which seemed to be present only for double and triple treatment matching models, i.e. in the presence of FHC and other PT.

Finally, we impose the more stringent condition that the amount of selection on unobservables equals the one on observables. As shown in the last column of the table, this corresponds to levels of ρ around three times the upper bound. In this case, all the effects disappear, however these levels of ρ would normally be considered unrealistic. Thus, we conclude that our main findings are robust to 'reasonable' amounts of selection on unobservables.

6 Conclusions

This paper examines the effects of personality traits on the decision to drop out from high school. We exploit rich data from Add Health and extend the literature by employing single and multiple treatment matching models to examine the relationship between heterogeneous personality styles and school dropout. Furthermore, we make use of specific information on forgone health care and explore its role as a predictor of dropout potentially related to psychological maturity of judgement and its associated traits such as responsibility, perspective and temperance.

We find that introversion and, to a lower degree, neuroticism increase the probability of dropping out from school in single treatment matching models. Multiple treatment matching shows higher likelihoods of an early exit for combinations of low levels of conscientiousness and neuroticism. Throughout all our models, we find that forgone health care is a strong and highly statistically significant risk factor for dropout decisions. In our triple treatment matching models, individuals who forgo their health care and present lower levels of conscientiousness and introversion seem to have the highest risk of dropout. Robustness checks confirm that overall the effects of personality traits and forgone health care are unaffected by realistic variations of selection on unobservables.

Interestingly, our estimates also highlight the limited role of conscientiousness if considered in isolation. In linear probability and single matching models, its effect on dropout does not appear to be statistically significant and this is confirmed at low levels of selection on unobservables. This may also suggest the importance of employing combinations of the Big Five facets when considering the relationship between personality and dropout behaviour.

It should be also noted that our estimates show larger effects of forgone health care on dropout for pupils from low socioeconomic backgrounds. Whilst we control for a wide range of variables concerning access to health care, including detailed information on health insurance, these effects might be still partly driven by restricted access to health care due to financial constraints. As such, forgone health care may be a more suited proxy of psychological maturity of judgement for individuals in middle and higher socioeconomic groups. Furthermore, as in any other study concerned with adolescents' personality traits, it should be borne in mind that noncognitive traits may still malleable at this stage of an individual's life. Hence, while these results might be helpful in building more specific profiles of pupils at risk of dropping out from high school, they may not necessarily apply to other types of behaviours such as college dropout.

We can conclude that forgone health care appears to be a relevant predictor of high school dropout and this might be driven by psychosocial and personality traits related to low psychological maturity of judgement. However, while we are aware that further research is needed to more precisely identify the noncognitive traits proxied by forgone health care, we argue that given its predictive power, the signalling feature of forgone health care could be exploited to refine the profiling of dropouts. That is, information on forgone health care could be used, together with other known determinants, as an additional device to identify adolescents at higher risk of school dropout.

Our work may have potentially important policy implications. First, evidence on the impact of personality styles could be directly exploited by governments and policy makers to refine the targeting of potential dropouts. Secondly, given its importance in explaining dropout decisions, we suggest that information on forgone health care among adolescents could be more routinely collected. This could be done for example by integrating students' medical records with educational records, or by adding specific questions on pre-school ability tests and future surveys on adolescents' educational attainment.

After the identification of students at higher risk of dropping out, specific policies such as monitoring and mentoring programmes could be implemented. These would have the objective to raise students and families' awareness on the possible long-term consequences of dropout, such as higher risk of unemployment, ill-health and lower lifetime earnings.[33] Broader societal benefits of reduced dropout rates would include higher tax revenues (due to the higher earnings of high school graduates) and lower spending on public assistance as well as lower crime rates.

[33]The National Dropout Prevention Center/Network web-site lists a number of alternative dropout prevention programmes, see http://www.dropoutprevention.org/modelprograms.

References

Abadie, A. and Imbens, G. (2008). Notes and comments on the failure of the bootstrap for matching estimators. *Econometrica*, 76(6).

Allison, P. D. (2000). *Missing data*. Sage Thousand Oaks, CA.

Almlund, M., Duckworth, A. L., Heckman, J. J., and Kautz., T. D. (2011). Personality psychology and economics. handbook of the economics of education. In Hanushek, E. A., Machin, S. J., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 4. Elsevier.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151.

Barenbaum, N. B. and David, G. (2008). History of modern personality theory and research. *Handbook of personality: Theory and research*, pages 3–28.

Barón, J. D. and Cobb-Clark, D. A. (2010). Are young people's educational outcomes linked to their sense of control? *Available at SSRN 1594792*.

Bedard, K. (2001). Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy*, 109(4).

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

Bratti, M. (2007). Parents' income and children's school drop-out at 16 in england and wales: evidence from the 1970 british cohort study. *Review of Economic Household*, 5:515–540.

Brown, S. and Taylor, K. (2014). Household finances and the 'big five'personality traits. *Journal of Economic Psychology*, 45:197–212.

Cauffman, E. and Steinberg, L. (2000). (im) maturity of judgment in adolescence: Why adolescents may be less culpable than adults. *Behavioral sciences & the law*, 18(6):741–760.

Cebi, M. (2007). Locus of control and human capital investment revisited. *Journal of Human Resources*, 42(4):919–932.

Chapman, C., Laird, J., and Kewal Ramani, A. (2010). Trends in high school dropout and completion rates in the united states: 1972-2008. compendium report. nces 2011-012. *National Center for Education Statistics*.

Chatterji, P., Alegria, M., Lu, M., and Takeuchi, D. (2007). Psychiatric disorders and labor market outcomes: evidence from the national latino and asian american study. *Health Economics*, 16(10):1069–1090.

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

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Coleman, M. and DeLeire, T. (2003). An economic model of locus of control and the human capital investment decision. *Journal of Human Resources*, 38(3):701–721.

Coneus, K., Gernandt, J., and Saam, M. (2011). Noncognitive skills, school achievements and educational dropout. *Schmollers Jahrbuch*, 131(4):547–568.

Costa, P. T. and McCrae, R. R. (1992). Revised neo personality inventory (neo-pi-r) and neo five-factor inventory (neo-ffi) professional manual.

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

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

De Witte, K., Cabus, S., Thyssen, G., Groot, W., and van den Brink, H. M. (2013). A critical review of the literature on school dropout. *Educational Research Review*, 10:13–28.

Denny, S., Farrant, B., Cosgriff, J., Harte, M., Sheridan, J., and Robinson, E. (2013). Forgone health care among secondary school students in new zealand. *Journal of Primary Health Care*, 5(1):11–18.

Donnellan, M. B., Conger, R. D., and Burzette, R. G. (2007). Personality development from late adolescence to young adulthood: Differential stability, normative maturity, and evidence for the maturity-stability hypothesis. *Journal of personality*, 75(2):237–264.

Donnellan, M. B., Oswald, F. L., Baird, B. M., and Lucas, R. E. (2006). The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. *Psychological assessment*, 18(2):192.

Dubow, E. F., Lovko Jr, K. R., and Kausch, D. F. (1990). Demographic differences in adolescents' health concerns and perceptions of helping agents. *Journal of Clinical Child Psychology*, 19(1):44–54.

Eckstein, Z. and Wolpin, K. (1999). Why youths drop out of high school: the impact of preferences, opportunities, and abilities the impact of preferences, opportunities, and abilities. *Econometrica*, 67(6):1295–1339.

Ermish, J. and Francesconi, M. (2001). Family matters: Impacts of family background on educational attainment. *Economica*, 68:137–156.

Ford, C. A., Bearman, P. S., and Moody, J. (1999). Foregone health care among adolescents. *Journal of the American Medical Association*, 282(23):2227–2234.

Ginsburg, K. R., Slap, G. B., Cnaan, A., Forke, C. M., Balsley, C. M., and Rouselle, D. M. (1995). Adolescents' perceptions of factors affecting their decisions to seek health care. *Journal of the American Medical Association*, 273(24):1913–1918.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist*, 48(1):26.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1):84–96.

Harris, K. M., Florey, F., Tabor, J., Bearman, P. S., Jones, J., and Udry, J. R. (2009). The national longitudinal study of adolescent health: research design.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Heckman, J. J. and LaFontaine, P. A. (2010). The american high school graduation rate: Trends and levels. *The Review of Economics and Statistics*, 92(2):244–262.

Heckman, J. J. and Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29(3):446–493.

Heckman, J. J. and Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, pages 145–149.

Imbens, G. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, (87):706–710.

Johnston, D. W., Schurer, S., and Shields, M. A. (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of Health Economics*, 32(6):1077–1089.

Kern, M. L. and Friedman, H. S. (2011). Personality and pathways of influence on physical health. *Social and Personality Psychology Compass*, 5(1):76–87.

Lechner, M. (2001). *Econometric Evaluation of Labour Market Policies*, chapter Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption, pages 43–58. Heidelberg: Physica-Verlag.

Lehrer, J. A., Pantell, R., Tebb, K., and Shafer, M.-A. (2007). Forgone health care among us adolescents: associations between risk characteristics and confidentiality concern. *Journal of Adolescent Health*, 40(3):218–226.

Lundberg, S. (2013). The college type: personality and educational inequality. *Journal of Labor Economics*, 31(3):421–441.

Martins, P. S. (2010). Can targeted, non-cognitive skills programs improve achievement? evidence from epis.

Mocetti, S. (2008). Educational choices and the selection process before and after compulsory schooling. *Banca D'Italia Eurosistema*, 691. Working Paper.

Nutbeam, D. (2008). The evolving concept of health literacy. *Social science & medicine*, 67(12):2072–2078.

OECD (2012). *Equity and Quality in Education*. OECD Publishing.

Oreopoulos, P. (2007). Do dropouts drop out too soon? wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91:2213–2229.

Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality*, 43(2):137–145.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, 80(1):1.

Saroglou, V. (2002). Religion and the five factors of personality: A meta-analytic review. *Personality and individual differences*, 32(1):15–25.

Schweinhart, L. J. (2003). Benefits, costs, and explanation of the high/scope perry preschool program.

Steinberg, L. and Cauffman, E. (1996). Maturity of judgment in adolescence: Psychosocial factors in adolescent decision making. *Law and Human Behavior*, 20(3):249.

Stinebrickner, T. R. and Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748.

Suhrcke, M. and de Paz Nieves, C. (2011). The impact of health and health behaviours on educational outcomes in high-income countries: a review of the evidence.

Terracciano, A. and Costa, P. T. (2004). Smoking and the five-factor model of personality. *Addiction*, 99(4):472–481.

Thornberry, T. P., Moore, M., and Christenson, R. (1985). The effect of dropping out of high school on subsequent criminal behavior. *Criminology*, 23(1):3–18.

Vollrath, M., Knoch, D., and Cassano, L. (1999). Personality, risky health behaviour, and perceived susceptibility to health risks. *European journal of personality*, 13(1):39–50.

Weiss, A., Sutin, A. R., Duberstein, P. R., Friedman, B., Bagby, R. M., and Costa, P. T. (2009). The personality domains and styles of the five-factor model are related to incident depression in medicare recipients aged 65 to 100. *The American Journal of Geriatric Psychiatry*, 17(7):591–601.

Wisk, L. E. and Witt, W. P. (2012). Predictors of delayed or forgone needed health care for families with children. *Pediatrics*, pages peds–2012.

Young, J. K. and Beaujean, A. A. (2011). Measuring personality in wave i of the national longitudinal study of adolescent health. *Frontiers in Psychology*, 2.

Table 1: Descriptive Statistics - Dropout and Individual Characteristics

	All	*Dropout*	*Graduate*	*t-test*	*N*
HS dropout	0.078	1.000	0.000	.	18765
Low conscientiousness (wave I)	0.148	0.168	0.146	0.025	18660
Neuroticism (wave I)	0.080	0.113	0.077	0.000	18714
Low extraversion (wave I)	0.173	0.223	0.170	0.000	13121
Low conscientiousness (wave IV)	0.099	0.120	0.096	0.005	14384
Neuroticism (wave IV)	0.169	0.272	0.158	0.000	14385
Low extraversion (wave IV)	0.225	0.282	0.219	0.000	14388
Low openness to experience (wave IV)	0.042	0.081	0.038	0.000	14380
Low agreeableness (wave IV)	0.048	0.104	0.043	0.000	14386
FHC	0.274	0.363	0.266	0.000	18749
Males	0.492	0.573	0.485	0.000	18762
Whites	0.536	0.492	0.539	0.001	18765
Hispanics	0.152	0.189	0.149	0.000	18765
African-Americans	0.229	0.274	0.225	0.000	18765
Asians	0.084	0.044	0.087	0.000	18765
Religious	0.861	0.780	0.868	0.000	18765
Peabody test (score)	100.207	91.171	100.976	0.000	17870
Learning disability	0.105	0.223	0.095	0.000	18765
Risky attitude	0.354	0.544	0.337	0.000	18765
High discount factor	0.129	0.261	0.118	0.000	18765
Ill-health	0.068	0.118	0.064	0.000	18765
Counseling	0.087	0.096	0.086	0.185	18765
Depression	0.100	0.177	0.094	0.000	18765
Asthma	0.117	0.135	0.115	0.037	15963
Migraine	0.113	0.171	0.108	0.000	15862
Obesity	0.064	0.081	0.063	0.011	15994
School absence (health reasons)	0.060	0.129	0.054	0.000	18765
Heavy smoker	0.024	0.078	0.019	0.000	18765
Heavy drinker	0.172	0.223	0.167	0.000	18765
Heavy marijuana	0.015	0.040	0.013	0.000	18765
Heavy other drugs	0.046	0.080	0.043	0.000	18765
Private/prepaid/other cover	0.775	0.525	0.796	0.000	16056
Medicare/medicaid	0.105	0.270	0.091	0.000	16056
No insurance	0.120	0.205	0.113	0.000	16056

T-test of equality of means between graduate and dropouts. P-value reported.

Table 2: Descriptive Statistics - Family and School Characteristics

	All	*Dropout*	*Graduate*	*t-test*	*N*
Mother					
No education	0.043	0.080	0.040	0.000	17667
Middle/primary	0.193	0.394	0.176	0.000	17667
High school	0.487	0.447	0.490	0.002	17667
Higher education	0.278	0.078	0.294	0.000	17667
Unemployed/at home	0.294	0.413	0.287	0.000	12315
Routine/technical	0.154	0.214	0.150	0.000	12315
Small employer/intermediate	0.242	0.185	0.246	0.000	12315
Managerial/professional	0.310	0.188	0.317	0.000	12315
Father					
No education	0.055	0.123	0.050	0.000	13738
Middle/primary	0.176	0.351	0.164	0.000	13738
High school	0.455	0.437	0.456	0.266	13738
Higher education	0.315	0.089	0.330	0.000	13738
Unemployed/at home	0.194	0.291	0.189	0.000	9909
Routine/technical	0.420	0.494	0.416	0.001	9909
Small employer/intermediate	0.119	0.114	0.119	0.699	9909
Managerial/professional	0.267	0.102	0.276	0.000	9909
Ill-health (main parent)	0.123	0.209	0.115	0.000	18765
Ill-health (partner)	0.085	0.125	0.082	0.000	18765
Access to medical care	0.125	0.210	0.117	0.000	18765
Family 'bad' behaviours	0.398	0.583	0.383	0.000	18765
School Characteristics					
Size: <= 125	0.019	0.029	0.018	0.002	18676
Size I: 126-350	0.069	0.077	0.068	0.185	18676
Size II: 351-775	0.237	0.219	0.238	0.086	18676
Size III: 776	0.676	0.675	0.676	0.944	18676
Public	0.930	0.984	0.925	0.000	18676
Catholic	0.028	0.010	0.030	0.000	18676
Private	0.042	0.006	0.045	0.000	18676
Urban	0.294	0.296	0.294	0.866	18676
Suburban	0.544	0.497	0.548	0.000	18676
Rural	0.161	0.206	0.157	0.000	18676
West	0.239	0.192	0.243	0.000	18676
Midwest	0.239	0.258	0.237	0.069	18676
South	0.374	0.440	0.368	0.000	18676
Northeast	0.148	0.110	0.151	0.000	18676
School services	0.726	0.779	0.721	0.000	18765
State health education	0.778	0.816	0.774	0.000	18657
School community health	0.483	0.524	0.479	0.002	15707
School health clinic	0.805	0.761	0.809	0.000	18413

T-test of equality of means between graduate and dropouts. P-value reported.

Table 3: Dropout and Personality Traits

Low conscientiousness	0.009	-0.002	-0.002	-0.003
	(0.006)	(0.006)	(0.005)	(0.005)
Neuroticism	0.030***	0.016**	0.007	0.006
	(0.007)	(0.007)	(0.007)	(0.007)
Low extraversion	0.016***	0.016**	0.012**	0.009
	(0.006)	(0.006)	(0.006)	(0.006)
Demographic Characteristics				
Male		0.013***	0.019***	0.021***
		(0.004)	(0.004)	(0.004)
Hispanic		0.018**	0.011	-0.001
		(0.007)	(0.007)	(0.007)
Afro-American		-0.005	-0.014**	-0.012*
		(0.007)	(0.007)	(0.007)
Asian		-0.018**	-0.021**	-0.016*
		(0.009)	(0.009)	(0.009)
Skills, Behaviors and Preferences				
Religion		-0.029***	-0.024***	-0.019***
		(0.006)	(0.006)	(0.006)
Peabody test		-0.002***	-0.002***	-0.002***
		(0.000)	(0.000)	(0.000)
Learning disability		0.047***	0.039***	0.038***
		(0.007)	(0.007)	(0.007)
Risky attitude		0.027***	0.023***	0.020***
		(0.004)	(0.004)	(0.004)
High discount factor		0.050***	0.047***	0.045***
		(0.006)	(0.006)	(0.006)
Smoking		0.102***	0.097***	0.089***
		(0.013)	(0.013)	(0.013)
Drinking		0.004	0.003	0.002
		(0.006)	(0.006)	(0.006)
Obesity		0.013*	0.009	0.007
		(0.007)	(0.007)	(0.007)
Marijuana		0.060***	0.056***	0.054***
		(0.017)	(0.016)	(0.016)
Other drugs		0.002	0.002	0.003
		(0.012)	(0.012)	(0.012)
Health				
Ill-health			0.009	0.007
			(0.008)	(0.008)
Counselling			0.006	0.005
			(0.007)	(0.007)
Depression			0.029***	0.026***
			(0.007)	(0.007)
Migraine			0.021***	0.016**
			(0.007)	(0.007)
School absence health reasons			0.049***	0.047***
			(0.008)	(0.008)
Medicare/medicaid			0.082***	0.060***
			(0.007)	(0.008)
No insurance			0.050***	0.029***
			(0.007)	(0.007)
Family characteristics	no	no	no	yes
constant	0.043***	0.306***	0.244***	0.230***
	(0.008)	(0.021)	(0.021)	(0.025)
N	18139	18139	18139	18139
F	29.903	47.446	39.561	30.776

Significance levels: *** 1% ** 5% * 10%. Cohort effects and school fixed effects included in all models.

Missing values dummy variables included in all models. Asthma, Obesity and Walking difficulties included in col.3 but not significant.

Table 4: Dropout and Personality traits - Wave IV

Low conscientiousness	0.011	0.005
	(0.008)	(0.007)
Neuroticism	0.058***	0.038***
	(0.006)	(0.006)
Low extraversion	0.014***	0.014***
	(0.005)	(0.005)
Low agreeableness	0.094***	0.065***
	(0.010)	(0.010)
Low openness to experience	0.074***	0.052***
	(0.011)	(0.011)
Individual characteristics	no	yes
Family characteristics	no	yes
N	18139	18139
F	37.746	37.420

Significance levels: *** 1% ** 5% * 10%.

Table 5: Personality Traits - Single and Multiple Treatment Matching, ATE

Single/Double treatment	Low conscientiousness	Neuroticism	Low extraversion
Low conscientioness	0.010	0.041**	0.004
	(0.007)	(0.020)	(0.012)
Neuroticism	.	0.019*	0.021
	.	(0.010)	(0.021)
Low extraversion	.	.	0.023***
	.	.	(0.008)
Triple treatment			
Low conscientioness & Neuroticism	.	.	0.009
	.	.	(0.035)

Significance levels: *** 1% ** 5% * 10%.

Table 6: Dropout, FHC and Family Characteristics

	All(FHC)	*Moth ed (-)*	*Moth ed (+)*	*Fath ed (-)*	*Fath ed (+)*
FHC	0.013***	0.039***	0.007	0.047***	0.009**
	(0.004)	(0.012)	(0.004)	(0.013)	(0.005)
Low conscientiousness	-0.003	-0.016	0.002	-0.005	0.008
	(0.005)	(0.016)	(0.005)	(0.017)	(0.006)
Neuroticism	0.005	0.000	0.011	-0.036*	0.013
	(0.007)	(0.020)	(0.008)	(0.022)	(0.008)
Low extraversion	0.009	0.020	0.008	-0.000	0.015**
	(0.006)	(0.018)	(0.006)	(0.019)	(0.007)
Middle/primary$_{moth}$	0.018			0.030	0.025
	(0.011)			(0.021)	(0.018)
High school$_{moth}$	-0.028***			-0.016	-0.006
	(0.011)			(0.022)	(0.017)
Higher education$_{moth}$	-0.043***			-0.030	-0.025
	(0.011)			(0.029)	(0.017)
Routine/technical$_{moth}$	0.003	0.024	-0.003	0.032	-0.009
	(0.008)	(0.018)	(0.008)	(0.021)	(0.008)
Small employer/intermediate$_{moth}$	-0.001	-0.013	-0.003	0.009	-0.004
	(0.007)	(0.022)	(0.006)	(0.022)	(0.007)
Managerial/professional$_{moth}$	0.003	-0.003	-0.005	0.007	0.002
	(0.006)	(0.023)	(0.006)	(0.023)	(0.007)
Middle/primary$_{fath}$	-0.009	-0.018	0.012		
	(0.012)	(0.020)	(0.015)		
High school$_{fath}$	-0.029***	-0.039*	-0.014		
	(0.011)	(0.021)	(0.014)		
Higher education$_{fath}$	-0.031***	-0.057*	-0.028*		
	(0.011)	(0.030)	(0.014)		
Routine/technical$_{fath}$	-0.002	-0.028	0.007	-0.016	0.004
	(0.007)	(0.020)	(0.007)	(0.019)	(0.007)
Small employer/intermediate$_{fath}$	0.001	-0.021	0.005	0.002	-0.001
	(0.010)	(0.030)	(0.010)	(0.029)	(0.009)
Managerial/professional$_{fath}$	-0.002	-0.047	0.007	-0.066**	0.002
	(0.008)	(0.030)	(0.008)	(0.031)	(0.008)
Ill-health (main parent)	-0.004	-0.034**	0.005	-0.026	0.016**
	(0.006)	(0.015)	(0.007)	(0.017)	(0.008)
Ill-health (partner)	0.006	0.028	-0.000	0.010	0.004
	(0.007)	(0.019)	(0.007)	(0.017)	(0.007)
Access to medical care	0.008	0.017	0.007	0.010	0.013
	(0.007)	(0.016)	(0.007)	(0.018)	(0.009)
Family health-behaviors	0.024***	0.043***	0.025***	0.038***	0.020***
	(0.004)	(0.013)	(0.004)	(0.014)	(0.004)
Individual characteristics	yes	yes	yes	yes	yes
constant	0.229***	0.277***	0.132***	0.259***	0.098***
	(0.025)	(0.054)	(0.026)	(0.061)	(0.028)
N	18139	4019	13094	3047	10240
F	30.435	7.559	17.069	5.094	14.891

Significance levels: *** 1% ** 5% * 10%.

Cohort effects, school fixed effects and individual characteristics included in all models.

Missing values dummy variables included in all models.

Table 7: School Characteristics and Non-GED holders

	Sch char 1	*Sch char 2*	*Non-GED*
FHC	0.033***	0.017***	0.010**
	(0.005)	(0.005)	(0.004)
Low conscientiousness	0.008	-0.001	-0.005
	(0.007)	(0.006)	(0.005)
Neuroticism	0.026***	0.004	0.011
	(0.009)	(0.008)	(0.007)
Low extraversion	0.012*	0.007	0.003
	(0.007)	(0.007)	(0.006)
Size I	-0.072**	-0.046*	-0.021
	(0.034)	(0.027)	(0.023)
Size II	-0.124***	-0.072***	-0.026
	(0.034)	(0.027)	(0.023)
Size III	-0.107***	-0.059**	-0.017
	(0.034)	(0.027)	(0.023)
Catholic	0.004	0.026	0.024*
	(0.012)	(0.016)	(0.013)
Private	-0.089***	-0.012	-0.026
	(0.021)	(0.021)	(0.018)
Suburban	-0.016***	0.000	-0.000
	(0.006)	(0.006)	(0.005)
Rural	-0.013	0.001	-0.006
	(0.010)	(0.008)	(0.007)
Midwest	0.048***	0.035***	0.035***
	(0.010)	(0.010)	(0.009)
South	0.018*	0.018	0.012
	(0.011)	(0.011)	(0.009)
Northeast	0.017	0.003	-0.003
	(0.011)	(0.012)	(0.009)
School health services	0.011	0.003	0.006
	(0.014)	(0.014)	(0.011)
State health education	0.017	0.019	0.010
	(0.015)	(0.015)	(0.012)
School community health	0.006	-0.002	-0.005
	(0.011)	(0.011)	(0.009)
School health clinic	-0.013*	-0.019***	-0.014**
	(0.007)	(0.007)	(0.006)
Individual characteristics	no	yes	yes
Family characteristics	no	yes	yes
constant	0.109***	0.241***	0.208***
	(0.034)	(0.038)	(0.052)
N	14693	14693	14493
F	13.214	23.322	18.585

Significance levels: *** 1% ** 5% * 10%.

Cohort effects and school grades included in all models.

Size I: 126-350 students. Size II: 351-775 students.

Size III: 776 or more students. Reference category: less than 126 students.

Col.3 excludes future GED-holders.

Table 8: FHC, Single and Multiple Treatment Matching, ATE

Single/Double treatment	FHC	Low conscientiousness	Neuroticism	Low extraversion
FHC	0.021***	0.028**	0.029**	0.028**
	(0.005)	(0.013)	(0.014)	(0.014)
Triple treatment				
FHC& Low conscientioness	.	.	0.012	0.050**
	.	.	(0.018)	(0.026)
FHC&Neuroticism	.	.	.	0.102
	.	.	.	(0.066)
Excluding risky health-behaviours				
FHC	0.028***	.	.	.
	(0.006)	.	.	.

Significance levels: *** 1% ** 5% * 10%.

Table 9: Selection on unobservables

	OLS	$\rho = 0$	$\rho = 0.01$	$\rho = 0.03$	$\rho = \beta_{ols}$	$\rho_{unob=obs}$	
FHC	0.033***	0.023***	0.021***	0.018***	0.017***	-0.000	$\rho = 0.132$
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	
Low conscientiouness	0.014**	0.007	0.005	0.001	0.004	0.000	$\rho = 0.037$
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	
Neuroticism	0.031***	0.025***	0.022***	0.017***	0.017***	0.000	$\rho = 0.109$
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	
Low extraversion	0.020***	0.014***	0.012***	0.009**	0.010**	-0.000	$\rho = 0.096$
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	

Significance levels: *** 1% ** 5% * 10%.

Marginal effects reported.

ρ is the correlation between the unobservables in a bivariate probit model.

OLS reports the estimate of FHC in a linear model without covariates.

Sample size: FHC models 18,139. Conscientiousness models 18,056.

Sample size: Neuroticism models 18,108. Extraversion models 12,521 observations.

Acknowledgements

We are grateful to David Figlio, Todd Stinebrickner, Bruce Hollingsworth, Colin Green, Chris Hatton, Maarten Lindeboom, Emma Tominey, Ian Walker, participants to the 2015 Royal Economic Society Conference, the 2015 Congress of the International Health Economics Association, the 2014 International Workshop on Applied Economics of Education and seminars at the universities of Ferrara, Stirling and York for useful comments.

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

This study was funded by an ESRC 'Radical Futures in Society' grant. The authors declare that they have no conflict of interest.

A Appendix

We estimate our most complete model (Table 6, column 1) with PT and FHC, by employing the Add Health weight components to account for the booster samples (i.e. supplemental samples based on ethnicity, genetic relatedness to siblings, adoption status, and disability). We find that the effect of FHC is 2.1pp, see Table A2, column 1. This result is confirmed when we further estimate our main model only on the wave I core sample (i.e. by excluding all booster samples, column 2, 1.8 pp). Furthermore, when we exclude all missing observations (column 3), the effect of FHC is positive, statistically significant and, consistently with our previous estimates, 1.9 pp in size. Neuroticism is never statistically significant, however in the specifications reported in columns 2 and 3, the other two PT show some statistically significant effects. More specifically, when excluding all missing values (column 3) we notice a reduction in the sample size to around 7476 observations, and that the effects of low conscientiousness and low extraversion become significant (whereas in our most complete models, the same variables were not significant). We believe that this apparently significant effects of low conscientiousness and low extraversion might be due to a potential sample selection bias, which we are not able to account for since we are not directly controlling for the noise generated by the missing values. As a result, we prefer models which account directly for missing values (and resulting selection bias) through the dummy variable adjustment approach.

Table A1: Dropout, Personality traits and FHC - full specifications

FHC		0.013***
		(0.004)
Low conscientiousnes	-0.003	-0.003
	(0.005)	(0.005)
Neuroticism	0.006	0.005
	(0.007)	(0.007)
Low extraversion	0.009	0.009
	(0.006)	(0.006)
Demographic Characteristics		
Male	0.021***	0.021***
	(0.004)	(0.004)
Hispanic	-0.001	-0.001
	(0.007)	(0.007)
Afro-American	-0.012*	-0.014**
	(0.007)	(0.007)
Asian	-0.016*	-0.017*
	(0.009)	(0.009)
Religion	-0.019***	-0.020***
	(0.006)	(0.006)
Skills, Behaviors and Preferences		
Peabody test	-0.002***	-0.002***
	(0.000)	(0.000)
Learning disability	0.038***	0.039***
	(0.007)	(0.007)
Risky attitude	0.020***	0.019***
	(0.004)	(0.004)
High discount factor	0.045***	0.044***
	(0.006)	(0.006)
Smoking	0.089***	0.089***
	(0.013)	(0.013)
Drinking	0.002	0.002
	(0.006)	(0.006)

Table A1: Dropout, Personality traits and FHC - full specifications

Obesity	0.007	0.007
	(0.007)	(0.007)
Marijuana	0.054***	0.054***
	(0.016)	(0.016)
Other drugs	0.003	0.003
	(0.012)	(0.012)
Health		
Ill-health	0.007	0.006
	(0.008)	(0.008)
Counselling	0.005	0.004
	(0.007)	(0.007)
Depression	0.026***	0.024***
	(0.007)	(0.007)
Migraine	0.016**	0.016**
	(0.007)	(0.007)
School absence health reasons	0.047***	0.046***
	(0.008)	(0.008)
Medicare/medicaid	0.060***	0.060***
	(0.008)	(0.008)
No insurance2	0.029***	0.029***
	(0.007)	(0.007)
Middle/primary$_{moth}$	0.019*	0.018
	(0.011)	(0.011)
High school$_{moth}$	-0.028***	-0.028***
	(0.011)	(0.011)
Higher education$_{moth}$	-0.042***	-0.043***
	(0.011)	(0.011)
Routine/technical$_{moth}$	0.003	0.003
	(0.008)	(0.008)
Small employer/intermediate$_{moth}$	-0.001	-0.001
	(0.007)	(0.007)
Managerial/professional$_{moth}$	0.003	0.003

Table A1: Dropout, Personality traits and FHC - full specifications

	(0.007)	(0.006)
Middle/primary$_{fath}$	-0.009	-0.009
	(0.012)	(0.012)
High school$_{fath}$	-0.029***	-0.029***
	(0.011)	(0.011)
Higher education$_{fath}$	-0.031***	-0.031***
	(0.011)	(0.011)
Routine/technical$_{fath}$	-0.002	-0.002
	(0.007)	(0.007)
Small employer/intermediate$_{fath}$	0.001	0.001
	(0.010)	(0.010)
Managerial/professional$_{fath}$	-0.002	-0.002
	(0.008)	(0.008)
Ill-health (main parent)	-0.004	-0.004
	(0.006)	(0.006)
Ill-health (partner)	0.007	0.006
	(0.007)	(0.007)
Access to medical care	0.008	0.008
	(0.007)	(0.007)
Family health-behaviors	0.024***	0.024***
	(0.004)	(0.004)
constant	0.230***	0.229***
	(0.025)	(0.025)
N	18139	18139

Significance levels: *** 1% ** 5% * 10%.

Cohort effects and school fixed effects included in both models.

Missing values dummy variables included in bothl models.

Table A2: Robustness Checks: Weights and Core Sample

	Restrictions on sample		
	Weighted	*Core*	*No missing*
FHC	0.021***	0.018***	0.019***
	(0.007)	(0.006)	(0.005)
Low conscientiousness	0.001	-0.005	0.017***
	(0.009)	(0.007)	(0.007)
Neuroticism	0.009	0.013	0.007
	(0.014)	(0.010)	(0.009)
Low extraversion	0.013	0.016**	0.017**
	(0.011)	(0.008)	(0.007)
Individual characteristics	yes	yes	yes
Family characteristics	yes	yes	yes
N	16850	10695	7476
F	12.899	17.390	11.348

Significance levels: *** 1% ** 5% * 10%.

Weighed: using sample weights and stratification by pseudo-states.

Core: core sample only.

No missing: excluding missing values and including school fixed effects.

Table A3: Balancing Test for Propensity Score Matching

	Low Conscientiousness			*FHC*		
	Std differences		*Var ratio*	*Std differences*		*Var ratio*
	Raw	*Matched*		*Raw*	*Matched*	
Males	-0.004	-0.014	0.999	-0.037	0.000	1.000
Hispanics	0.006	-0.006	0.987	0.080	0.025	1.052
African-Americans	-0.087	0.025	1.035	0.101	-0.002	0.998
Asians	-0.040	0.035	1.111	0.012	-0.010	0.970
Religious	-0.135	0.006	0.987	-0.104	0.005	0.990
Peabody test (score)	-0.055	-0.020	1.035	-0.118	0.005	0.965
Learning disability	0.191	0.013	1.035	0.082	-0.013	0.963
Medicare/medicaid	0.058	0.006	1.019	0.074	0.000	1.000
No insurance	0.017	0.025	1.072	0.093	-0.019	0.954
Mother						
Middle/primary	0.033	0.011	1.020	0.117	0.009	1.016
High school	0.005	0.000	1.000	-0.080	-0.022	0.997
Higher education	-0.102	-0.041	0.957	-0.086	0.011	1.012
Routine/technical	-0.058	0.009	1.022	-0.075	-0.014	0.966
Small employer/intermediate	-0.173	-0.027	0.959	-0.177	0.025	1.040
Managerial/professional	-0.230	0.001	1.001	-0.232	-0.007	0.992
Father						
Middle/primary	0.022	0.037	1.090	0.060	0.033	1.077
High school	-0.004	-0.046	0.972	-0.075	-0.020	0.987
Higher education	-0.087	-0.030	0.965	-0.124	-0.016	0.980
Routine/technical	-0.155	0.005	1.005	-0.193	-0.012	0.986
Small employer/intermediate	-0.122	-0.017	0.948	-0.132	-0.019	0.939
Managerial/professional	-0.179	-0.011	0.982	-0.236	-0.011	0.981
Access to medical care	0.014	0.009	1.023	0.098	0.000	0.999
Family 'bad' behaviours	0.118	-0.019	0.991	0.128	0.000	1.000

Variance ratio (in the matched sample) for variable x: $\frac{\sigma^2_x(T=1)}{\sigma^2_x(T=0)}$ where T=FHC, Low Conscientiousness

B Appendix: Variable definitions

Personality traits: we follow Young and Beaujean (2011) to define wave I personality traits. *Conscientiousness*: the items used to define wave I conscientiousness relate to the IPIP NEO-PI-R questions on paying attention to details; coming up with good solutions; doing things according to plans; and doing more to what is expected to me and are “when you have a problem to solve, one of the first things you do is to get as many facts about the problem as possible”; “when you are attempting to find a solution to a problem, you usually try to think of as many different ways to approach the problem as possible”; “when making decisions, you generally use a systematic method for judging and comparing alternatives”; and “after carrying out a solution to a problem, you usually try to analyze what went right and what went wrong”, respectively. Conscientiousness is defined through a dummy variable taking value 1 if an individual answers ‘disagree/strongly disagree’ to at least three out of these four questions and 0 otherwise.

Neuroticism: wave I items for Neuroticism relate to the NEO-PI-R questions about having a low opinion of myself; feeling comfortable with myself; being very pleased with myself; worry about things; finding it difficult to approach others; and disliking myself and are “you have a lot of good qualities”; “you have a lot to be proud of”; “you like yourself just the way you are”; “you feel like you are doing everything just about right”; “you feel socially accepted” (in-home questionnaire); “you feel wanted and loved”. The binary indicator for neuroticism equals 1 for a pupils answering ‘agree /strongly agree’ to at least five of these items and 0 otherwise.

Extraversion: the items for wave I extraversion concern the NEO-PI-R questions about make friends easily; warming up quickly to others and feeling comfortable around people and are ”I feel close to people at school”; “I feel like I am part of this school” and “I feel socially accepted” (school questionnaire). Introversion is defined through a dummy variable taking value 1 when an individual reports ‘disagree/strongly disagree’ in at least two of three questions and 0 otherwise.

Wave IV personality traits: we employ the standard “mini-IPIP” 20-items measure of the the Big Five to define the wave IV full set of personality traits.

Learning disabilities: we use a dummy variable which equals 1 in the presence of ‘specific learning disability, such as difficulties with attention, dyslexia, or some other reading, spelling, writing, or math disability’.

Risky health-behaviours: we employ several dummy variables which, respectively, equal 1 in the presence of heavy consumption of tobacco (”smoking”: smoking 20 or more cigarettes in the days you smoke); heavy alcohol consumption (”drinking”: having 5 or more drinks every time you drink), daily marijuana consumption (”marijuana”: 30 or more marijuana cigarettes in the last 30 days), ”other drugs”, i.e. cocaine (10 or more times in the last 30 days) or inhalants (10 or more times in the last 30 days); 0 otherwise.

Individual's attitude towards risk (risky attitude): we define a binary indicator capturing at least one of the following behaviours: 'no use of seat belts' or 'no use of birth controls'.

Time preferences/discounting: we include a dummy variable which equals 1 when an individual reports 'agree' or 'strongly agree' to the sentence 'I live my life without much thought for the future'; 0 otherwise.

Physical health: we employ a binary measure of self-assessed general health (which equals 1 if an individual reports fair or poor health; 0 otherwise). We also control for specific health conditions such as migraine, asthma, physical disabilities (walking difficulties) and obesity.

Obesity: we employ a binary variable equal if the BMI is greater than 30; 0 otherwise.

Mental health: we define mental health using two dummy variables. A first variable identifies adolescents who received 'counselling, psychological testing, or any mental health or therapy service within the last 12 months'. A second binary indicator identifies individuals feeling depressed all time or most of the time (against never depressed, rarely or sometimes). While the first mental health variable should identify adolescents who received any counselling for mental health reasons, the second variable should be a proxy for the intensity of mental health problems (depression).

Type of health insurance: we include a variable defining different types of health insurance. These are: not being covered by health insurance, being under Medicaid or Medicare support, other health insurance covers (baseline category).

School absences due to health or emotional problems: we employ a dummy variable which equals 1 when a pupil was absent from school at least once a week in the last month; 0 otherwise.

Parental education: we define a categorical variable that distinguishes between primary/middle school; high school; higher education; and use no education as the baseline category.

Parental job: we define a categorical variable that considers routine occupation/technical occupation; small employers/intermediate occupation; managerial and professional occupation versus unemployed/ home maker as baseline category.

Parents' difficulties in accessing health care: we use a dummy variable which equals 1 when parents answer 'hard' or 'somewhat hard' to the question *'in general, how easy or hard is it for you to get medical care for your family?*'; 0 otherwise.

Health-behaviours of family members: we use a binary indicator capturing at least one of the following behaviours: main parent is a smoker; another member of the family is a smoker; main parent drinks more than 5 five drinks at times at least 3 times a week.

Highlights

We explore the relationship between personality traits and school dropout

We employ multiple treatment propensity score matching

We use forgone health care as a proxy for psychological maturity of judgement

Forgone health care is a consistently significant predictor of dropout

Specific combinations of traits are associated with an increase in school dropout