Economics Working Paper Series

2014/026

The Relationship between Forgone Health Care and High School Dropout: Evidence from US Adolescents

Giuseppe Migali and Eugenio Zucchelli

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

LUMS home page: http://www.lums.lancs.ac.uk/

The Relationship between Forgone Health Care and High School Dropout: Evidence from US Adolescents

Giuseppe Migali*
Eugenio Zucchelli†

December 18, 2014

Abstract

High school dropout is an important policy issue and its determinants are a long-standing interest of economics. However, very little is known on the roles of noncognitive traits in influencing school dropout decisions. We employ voluntary forgone health care as a proxy for the underlying noncognitive traits that may induce adolescents to dropout and estimate its effects on early school attrition. We exploit data from the US National Longitudinal Study of Adolescent to Adult Health (Add Health) and employ a series of flexible specifications with school fixed effects and cohort effects. Our models account for well-established determinants of dropout, including individual and parental characteristics, together with personality traits. Forgone health care consistently appears to be a statistically significant and substantial predictor of dropout among adolescents. We suggest that forgone health care could be used as a signalling device for policy makers targeting potential high school dropouts.

JEL Classification: I1, I2, I18
Keywords: forgone health care; high school dropout; Add Health.

1 Introduction

Dropping out of high school is still a major policy issue that affects more than 20% of young pupils in the majority of OECD countries (OECD, 2012). Early dropout substantially increases the risk of unemployment, leads to lower lifetime earnings and is linked to a number

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

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

of adverse outcomes later on in life including poorer health status and behaviours such as violence and crime (Chapman et al., 2010; Thornberry et al., 1985). Heckman and LaFontaine (2010) find that high school dropout in the U.S. 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 U.S. 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 some of these skills and preferences could be observed and influenced, some important individual traits that may predict dropout are difficult to identify and tackle (Eckstein and Wolpin, 1999).

Noncognitive traits appear to influence individual behaviours related to education (Almlund et al., 2011). Despite an emerging literature on the relationship between personality traits and education, very little is known on the effects of noncognitive traits on the decision to drop out from high school. In the presence of persistent heterogeneity in multiple noncognitive traits, it is important to find proxies of those traits that could be strong predictors of dropout decisions.

Noncognitive traits may also be powerful predictors of health and health-related behaviours (Almlund et al., 2011). Forgone health care is a widespread phenomenon among adolescents and a well-known subject within the medical literature (e.g. Ford et al., 1999; Ginsburg et al., 1995). Health can be forgone either by an 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 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 related to persistency, dutifulness or the inability to commit. Accordingly, forgone health care could be an important predictor of other behaviours involving these traits such as high school dropout decisions.

The main objective of this paper is to employ voluntary forgone health care as a proxy for the underlying noncognitive traits that may induce adolescents to dropout and estimate its effects on early school attrition. We exploit rich individual-level data on high school pupils, including unique information on forgone health care from the US National Longitudinal Study of Adolescent to Adult Health (Add Health). We estimate the effects of forgone health care on school dropout using a series of flexible specifications with fixed effects. We find that forgone health care appears to be a consistent and strong predictor of dropout. This is confirmed by several specifications that include all of the well-established determinants of high school dropout and employ school fixed effects and cohort effects. The relevant quantitative effect of forgone health care appears to be robust to the inclusion of a wide range of physical and mental health variables that account for the direct effect of ill-health on dropout, access and types of health insurance and alternative definitions of personality traits. Also, the influence of forgone health care on high school dropout does not appear to be context dependent as it does not vary by parental background characteristics. Our robustness checks also suggests that forgone health care does not appear to depend on

cognitive abilities but it is likely to be driven by (unobserved) noncognitive traits. Overall, we find that the presence of forgone health care increases school dropout by between 1 to 3.5 percentage points. This corresponds to a percentage variation of the average dropout rate between 12.7 and 21.5.

This paper offers several contributions to the literature. To the best of our knowledge, this is the first paper that employs forgone health care as a potential new predictor of high school dropout. Secondly, this is also among the very few studies that focus on the influence of noncognitive traits on the decision to drop out from high school. We simultaneously explore whether high school dropout is influenced by personality traits as well as if the relationship between forgone health care and dropout is mediated by known noncognitive traits. Thirdly, we examine whether the effect of forgone health care on dropout is context-dependent and affected by socioeconomic status. Finally, we contribute more broadly to the literature concerned with the determinants of educational attainment by bridging the medical literature with the economics literature.

The rest of the paper proceeds as follow. Section 2 provides some background on the determinants of high school dropout, their relationship with noncognitive traits as well as forgone health care. Section 3 focuses on the description of data and econometric methods. Section 4 presents main results and robustness checks and Section 5 concludes.

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 medical literature on forgone health care.

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 rewards from obtaining a degree against the effort needed to obtain it. The decision of dropping out from high school has been analysed mainly in the light of two competing theories: the 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). 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

[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

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 high 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 socio-demographic characteristics or cognitive skills and very few of them include noncognitive abilities.[2]

2.2 Noncognitive skills and high school dropout

Although there is a growing literature in both psychology and economics on the relationship between noncognitive skills and educational attainment, less is known on the effects of noncognitive skills on the specific decision to drop out from high school.[3]

Heckman and Rubinstein (2001) point out that in the US low returns to General Education Degrees (GEDs) 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 also account for noncognitive skills.

Three recent studies explore whether personality traits may play an important role in determining educational attainment by focusing on 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. Although these three papers explore the relevance of noncognitive traits on educational choices, they focus exclusively on the effects of locus of control and consider only a relatively limited number of covariates.[5]

increase the probability of college dropout.

[2]There is no standard terminology across disciplines in the definition of individual skills other than cognitive abilities. Economists tend to use the terms noncognitive skills (Heckman and Rubinstein, 2001) or personality traits (Borghans et al., 2008). Psychologists employ a broader range of definitions including cognitive self-regulation and self-discipline (Blair and Diamond, 2008; Blair and Razza, 2007). We follow the standard terminology in economics and adopt the terms noncognitive skills or noncognitive traits interchangeably.

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

[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 Longitu-

More recently, DiPrete and Jennings (2012) use data from the US Early Child Longitudinal Study Kindergarten Cohort and find that noncognitive abilities such as social and behavioural skills affect academic outcomes from kindergarten through fifth grade. Moreover, they find that girls begin school with higher noncognitive abilities and that the female advantage grows over time. Mendolia and Walker (2014) examine the relationships between personality traits, subject choice and performance in high school. Using the Longitudinal Study of Young People in England (LSYPE) they find that individuals with external locus of control and low self-esteem are less likely to achieve good performances in test scores, especially in mathematics and science. These effects seem to be stronger for adolescents from disadvantaged backgrounds.

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). The key messages from these streams of research suggest that noncognitive skills promote the development of cognitive abilities, whereas the impact of cognitive abilities on noncognitive skills seems more limited. Over an individual's life-cycle, noncognitive skills are malleable for longer periods than cognitive skills. Remediation policies (e.g. those aimed at individuals from disadvantaged family backgrounds) and early life interventions (e.g. Perry Preschool project) should focus on developing noncognitive skills and should be preferred to interventions later on in life.[6]

2.3 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) employs 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, other important factors that increase the probability of reporting forgone care are individual behaviours, such as daily cigarette use, frequent alcohol consumption and sexual intercourse. Although results from previous studies on size and significance of the main determinants of forgone health care among adolescents may vary, they all recognise

dinal 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 (urbanity). 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.

[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.

the relevance of an individual's propensity to access health care services. The latter appears to be determined partly by the ability to secure access to health care service (circumstances) and partly by a set of individual beliefs and traits that translate into a series of behaviours.

In this paper, we hypothesise that the individual noncognitive traits which, among other factors, induce adolescents to voluntary forgone health care may also be relevant determinants of high school dropout. Accordingly, information on forgone health care should help in predicting dropout decisions.

3 Data and Econometric Methods

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.

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 behaviors, 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 also draw information on parental background,

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

parents and relatives' health-behaviours and health insurance from the parent questionnaire.

Our dependent variable is High School dropout:

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

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.[8] Our reference category, HS graduates, includes adolescents that have at least a high school diploma.

After a series of further checks,[9] and the identification of several inconsistencies in the education data across waves, we have 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 [10].

Our variable of interest is Foregone Health Care (FHC):

$$\begin{cases} FHC_i = 1, & \text{if individual i reports forgone health care;} \\ FHC_i = 0, & \text{otherwise} \end{cases}$$

Following the medical literature, Foregone 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). In order to capture voluntary forgone health care 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 voluntary Forgone Health Care all the remaining answers that did not relate to strictly objective circumstances and that 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 Forgone Health Care should exclude serious health conditions. Furthermore, we also remove from our analysis all adolescents with chronic medical conditions (e.g. diabetes) as we believe that their behaviour should be systematically different. In addition, we control

[8] See Heckman and LaFontaine (2010) for a discussion on GED and HS graduates.

[9] 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 and highest education level, high school graduation status and most recent degree/certificate in wave IV.

[10] Most of them are likely to be "stopouts", that is individuals who interrupted their studies and eventually obtained a qualification afterwards.

for a number of specific physical and mental health issues to account for the direct effect of ill-health on dropout decisions. Since the two questions on Forgone Health Care were asked in wave I, all respondents were still enrolled at school.

Our final sample includes 19,038 observations. Table 1 reports basic descriptive statistics of our variables of interest, including dependent variable (HS dropout) and Forgone Health Care (FHC). HS dropouts are around 8% of our sample.[11] Overall, slightly more than 27% of students do not seek health care when needed. This percentage is substantially higher among dropouts (around 36%).

3.1 Explanatory Variables

Due to the richness of information included in Add Health, we have combined our determinants of HS dropout in four broad categories. Descriptive statistics for these variables are reported in Table 1 and Table 2.

Individual Characteristics

Our basic specification includes a wide range of demographic and other relevant individual-level characteristics. We control for an individual's year of birth (or cohort fixed effect), gender (including a dummy variable for males and using females as baseline) and ethnicity (categorised as African-Americans; Asians, Hispanics; using whites as baseline). In Table 1, we notice that in our sample the majority of dropouts are male (57%) and white (nearly 49% versus 27.5% African-American, 19.4% Hispanic and 4.4% Asian).

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. As expected, dropouts in our sample present lower average scores of the AHPVT if compared to high school graduates (Table 1). Moreover, we control for learning disabilities using a dummy variable based on the (parent questionnaire) question: "Does (he/she) have a specific learning disability, such as difficulties with attention, dyslexia, or some other reading, spelling, writing, or math disability?". Around 22% of dropouts appear to suffer from this problem.

Health-behaviours are accounted for using a variable identifying at least one of the following bad behaviours: heavy consumption of tobacco (smoking 20 or more cigarettes in the days you smoke); alcohol (having 5 or more drinks every time you drink); marijuana (30 or more joints in the last 30 days); cocaine (10 or more times in the last 30 days); or inhalants (10 or more times in the last 30 days). Among dropouts, 28% report having at least one of these behaviours. An individual's attitude towards risk is defined through a binary indicator capturing at least one of the following behaviours: "no use of seat belts" or "no use of birth controls". We notice that around 55% of adolescents who drop out from school present at least one of these risky behaviours.

[11] 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 overall number of dropouts in Add Health.

Following Oreopoulos (2007), who suggests that adolescents may heavily discount the future consequences of dropping out from school, we use self-reported information to define a dummy variable capturing high rates of time preferences/discounting.[12] We observe that 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 to control for whether pupils report any religion (however, we do not distinguish between religious denominations). 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. We employ a binary measure of self-assessed general health (which equals 1 if health is reported as fair or poor versus excellent, very good and good health) and two variables concerning mental health. A first dummy variable identifies adolescents who received "counseling, 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 counseling for mental health reasons, the second variable should be a proxy for the intensity of mental health problems (depression). It is interesting to note that while only 12% of dropouts report general fair or poor health, around 17% report being depressed.

We further include another binary measure of general health based on the frequency of school absences due to health or emotional problems (which equals 1 when the pupil was absent from school at least once a week in the last month). This should capture any relevant physical or psychological issue preventing regular school attendance. We are also able to control for more specific health conditions such as migraine, asthma, physical disabilities (walking difficulties) and obesity. The prevalence of these health conditions appears to be higher among dropouts than graduates.[13]

Since health care utilisation may depend on health insurance, we control for whether parents have access to health insurance as well as the type of insurance. We include a variable defining three categories: not being covered by health insurance, being under Medicaid or Medicare support, other health insurance covers (baseline category). A substantial number of dropouts' parents (around 27%) are covered by Medicare/Medicaid while about 21% report not having any health insurance.

Family Characteristics

Our second set of models includes a series of relevant family characteristics. We define parental socioeconomic background by including categorical variables for both parents ed-

[12]We employ information on individuals who "agree" or "strongly agree" to the sentence "I live my life without much thought for the future".

[13]This is because chronic conditions such as asthma and migraine (Rees and Sabia, 2011), physical disabilities and other conditions that may cause stigma such as obesity could have an impact on school attendance and ultimately dropout decisions. Further, walking difficulties could also be considered an objective circumstance for not accessing health care when needed and confound the effect of forgone health care. It should be noted that obesity is defined using parents' assessment on a child's obesity.

ucational attainment (primary/middle school; high school; higher education; with no education as the baseline category) and job status (routine occupation/technical occupation; small employers/intermediate occupation; managerial and professional occupation versus unemployed/ home maker as baseline category). Table 2 shows that more than 75% of dropouts have low educated mothers and fathers. We also observe that more than 40% of dropouts have an unemployed/at home mother and around 49% have a father with a routine/technical job.

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 (identifying in both cases a poor health condition). We observe that around 21% of dropouts have their main parent in poor health conditions. Moreover, we include in our models parent's difficulties in accessing health care by identifying those parents that 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". This variable coupled with the one on health insurance should control for objective difficulties in accessing medical care. for 21% of dropouts have parents who have difficulties in accessing medical care. Health-behaviours of family members are defined by 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. In our sample, family "bad" health-behaviours concern almost 58% of dropouts as opposed to only 38% of graduates.

Personality Traits

Since our main objective is to examine the role of forgone health care (and indirectly the role of noncognitive traits related to it) on dropout decisions, our third set of models includes measures of personality traits. Following Young and Beaujean (2011), we combine information from 13 questions in wave I to build three of the Big Five personality traits: 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 Add Health wave I in-home questionnaire and in-school questionnaire for items that matched statements from the International Personality Item Pool (IPIP, Goldberg et al. (2006)) version of the the NEO Personality Inventory (NEO-PI-R, Costa and McCrae (1992)). 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.[14] This variable identifies lack of conscientiousness for ease of comparison with our main variable of interest (FHC, forgone health care). That is, a positive effect of this variable would imply that lack of conscientiousness is positively associated with high school dropout. We notice in Table 1 that a slightly higher percentage of dropouts present lack of conscientiousness if compared to graduates (16.6% versus 14.6%).

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.[15] This

[14]The questions considered are: paying attention to details, coming up with good solutions, doing things according to plans and doing more to what is expected to me.

[15]These relate to having a low opinion of myself; feeling comfortable with myself; being very pleased with myself; finding it difficult to approach others; and disliking myself.

variable identifies emotional instability, and we observe that dropouts are more likely than graduates to be neurotic (11.3% vs 7.8%).

Finally, extraversion is defined through a dummy variable taking value 1 when an individual reports "disagree/strongly disagree" in at least two of three questions related to extraversion[16]. This dummy identifies the negative side of the trait (introversion). Introversion appears to be more prevalent among dropouts than (22.2%) than high school graduates (16.9%).

As an alternative, we also estimate models with the full set of the Big Five personality traits 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.[17] 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).[18]

School Characteristics

Our final specifications also make use of school contextual data and include information on school grade spans (i.e. grades offered) and school size (four sizes: less than 126 students, baseline; between 126-350 students; 351-775; and 776 or more). We also control for school type and we notice fewer dropouts in catholic and private schools, whereas most of dropout decisions appear to be concentrated in public schools (see Table 2). We exploit information on school locations (urban, suburban, rural) and geographical position (West, Midwest, South and Northeast). In addition, we employ a school identifier that allows us to include school fixed effects.

3.2 Econometric methodology

Our basic specification is the following linear probability model:

$$E[Y_i|FHC, X] = \gamma + \mu_s + \theta_t + \alpha FHC_i + \beta \mathbf{X_i}. \tag{1}$$

where μ_s are school fixed effects which control for any school-specific factors that might be also correlated with our variable of interest, forgone health care (FHC). As a robustness check we re-estimate equation 1 by replacing school fixed effects with variables that account for school-level heterogeneity (e.q. type, size, location etc.). θ_t are cohort effects, which are built using individuals' years of birth (i.e. year dummies 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

[16]Making friends easily, warming up quickly to others, feel comfortable around people

[17]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.

[18]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.

also impact on FHC. **X** is a vector that includes (incrementally) the full set of observed variables described above. We focus on the estimation of the parameter α (FHC). The hypothesis we want to test is whether FHC is an independent predictor of HS dropout. For this purpose, we estimate a succession of more comprehensive models with an incremental number of explanatory variables that may affect dropout behaviour. Our approach exploits the richness of observables available in Add Health to reduce the omitted variable bias, while simultaneously accounting for as many sources of noise as possible that may confound the impact of FHC on high school dropout. Further, we test whether the association between FHC and dropout holds once we also account for observed personality traits. This should help evaluating whether known noncognitive traits are mediating the relationship between FHC and high school dropout choices.

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 personality 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 appear to point out that data are genuinely missing at random. In any case, we also estimate the full battery of models by dropping all missing observations including our preferred specifications whose results are reported among the robustness checks. This implies a substantial reduction of our sample size. However direction, statistical significance and size of our main results are confirmed.

4 Results

Individual Characteristics

Table 3 reports our first set of estimates on the determinants of high school dropout. These are obtained from linear probability models that includes individual characteristics, school fixed effects and cohort effects. Mod.1a only includes demographic variables and the effect of forgone health care (FHC) on dropout is positive, highly statistically significant and substantial in size (3.4 percentage points, henceforth pp). This implies that in this initial model, FHC appears to be an important risk factor for dropout decisions. Mod.1b adds controls for cognitive abilities, preferences and behaviors. Here, the estimated coefficient for FHC is still highly statistically significant while decreasing slightly in size to 2.8 pp. Mod.1c presents the richest specification and further includes a number of health conditions and information on health insurance. The effect of FHC reduces to around 2 pp and is still highly statistically significant. Notice that the standard errors of FHC are not affected by the inclusion of additional variables, leaving the residual variance also unaffected.

All the other covariates present the expected signs. For example, in line with the literature males have a higher probability of dropping out compared to females (from about 2.3 to 1.9

pp depending on the models), whereas individuals with higher levels of cognitive skills (scores of PVT) have a lower probability of dropping out. Adolescents with a high discount factor (i.e. that "do not give much thought about the future"), those consuming heavily tobacco or other drugs, and the ones with a propensity toward risk ("no birth control use" and/or "no seat belt use") are also more likely to drop out.

If compared to the previous literature, the negative effect of being African-American on dropout may appear counterintuitive. However, recall that Add Health purposely over-sampled African-American students from highly educated families and this may justify the direction of the effect. In accordance with previous evidence, individuals with an Asian background present a lower probability of dropping out while the opposite is observed for Hispanics. As expected, suffering from ill-health (1.8 pp), depression (3.3 pp), migraine (2.2 pp), school absence due to health reasons (5.3 pp) increase the probability of dropping out and are all statistically significant predictors.

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 5.4 pp while being covered by either Medicare or Medicaid by 8.5 pp. This is not surprising as these variables may be also considered proxies of low income and job status. More specifically, Medicaid provides coverage to individuals with limited incomes or disabilities. Medicare covers those aged 65 years old or older and individuals under 65 years who are disabled and have been receiving different types of social security benefits. The absence of any health insurance might be related to the lack of an employer-based insurance or the inability to purchase individual coverage.

Family Characteristics

Mod.2a in Table 4 extends Mod.1c (Table 3) by adding variables that capture parental characteristics including health status, health-behaviours and access to medical care. We notice that in this model, the effect of FHC on HS dropout is stable: the estimated coefficient for FHC remains positive, highly statistically significant and around 1.8 pp in size. We also observe an increase in the probability of dropping out (3.3 pp) when at least one of the following conditions is present: main parent is a smoker or a heavy drinker or when another member of the family is a smoker.

Mod.2b further includes controls for mother's and father's education and occupation. Here, FHC is still highly significant and its positive effect on dropout is just slightly lower at 1.7 pp. In line with the latest literature, adolescents with higher educated mothers (high school: -2,7 pp; higher education: -4.4 pp) and fathers (high school: -2.7; higher education: -2.8 pp) are less likely to drop out from school.

Since some of the variables, 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).[19]

Personality Traits

[19]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.

Mod.3a-3b include measures of the Big Five personality traits. The relationship between education and personality is receiving increasing attention among economists. Previous studies have indicated that two of the Big Five, conscientiousness and neuroticism, have some predictive power on educational attainment. A recent review from Almlund et al. (2011) highlights a positive effect of conscientiousness and a (mostly) negative effect of neuroticism on education. Closer to our work, Lundberg (2013) employs the "mini-IPIP" in wave IV of Add Health. He finds that conscientiousness has a positive effect on college graduation, especially among individuals from advantaged socioeconomic backgrounds, whereas openness to experience has positive effects on college completion among those from less educated families. In line with the recent literature we think that personality traits are potentially important determinants of HS dropout.

The specific purpose of the analysis in Mod.3a-.3b is to evaluate whether the Big Five are influencing the relationship between FHC and HS dropout.

Simple (t-test) correlations show only a small degree of association between FHC and three of the Big Five personality traits, built using wave I information. Interestingly, (lack of) conscientiousness presents the smallest correlation with FHC (around 2%). Extraversion (defined negatively as introversion for ease of comparison with the effect of FHC) is also only marginally correlated with FHC (around 4.3%). Neuroticism shows a slightly higher correlation (around 9.7%). This may provide some support for our assumption that forgone health care might be a behaviour driven by further (unobserved) noncognitive traits.

Mod.3a in Table 5 includes among the determinants of dropout only FHC, wave I personality traits, school fixed effects and cohort effects. In this model, the estimated coefficient for FHC is positive, highly statistically significant and relatively large in size (3.0 pp). Hence, even controlling for the presence of standard personality traits, FHC still increases the probability of HS dropout. Neuroticism and introversion also appear to be positive and statistically significant predictors of HS dropout, although their coefficients are smaller in size (2.8 pp and 1.5 pp). Lack of conscientiousness does not seem to have a statistically significant effect. Importantly, when we add individual and parental background characteristics (Mod.3b), while FHC is still statistically significant, neuroticism and introversion cease to be important factors in determining dropout. We also notice that the quantitative effect of FHC is stable at 1.7 pp (the same as in Mod.2b, Table 4), which corresponds to a 21.5% increase of the average dropout rate.[20]

4.1 Robustness checks

In order to further assess the validity of our results, we perform a series of robustness checks. For comparative purposes, our reference model is Mod.3b in Table 5. This model includes individual and family characteristics, wave I personality traits, cohort effects and schools fixed effects. Mod.4a-b in Table 6 are identical to Mod.3a-b, respectively, except for personality traits that have been replaced by the wave IV full set of Big Five. Estimates of Mod.4a show

[20]We have also estimated models with alternative definitions of wave I personality traits. These include measures of personality traits both on a continuous scale built by simply summing the row values of the answers for each of the item/questions, as well as categorical variables taking into account the intensity of each trait. Results were very similar to those reported here and are available upon request.

that FHC is a statistically significant determinant of dropout with a quantitative effect of 2.6 pp. Most of the Big Five have the expected effects on HS dropout: positive for neuroticism, extraversion (defined negatively as introversion), openness (defined as closeness) and (lack of) agreeableness. These effects are statistically significant and relatively large in size. However, conscientiousness still does not appear to play a substantial role in high school dropout decisions.

Mod.4b adds controls for individuals and parents and the effect of FHC decreases to 1 pp, while remaining positive and statistically significant. In this model, all the effects of the other personality traits are confirmed, although quantitatively reduced. Overall, the importance of FHC in determining HS dropout is substantially confirmed. However, the effects of the Big Five should be considered with caution, as the questions used to define them are collected when individuals have already entered adulthood (2008-2009).

Mod.5a replaces school fixed effects with direct measures of school characteristics. In Table 7, Mod.5a includes FHC, 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 (that often provide more options and services to students), non-public schools (catholic and private), those located in suburban areas, in the Midwest, South or Northeast (compared to the West). The effect of our variable of interest appears to be unaffected by the inclusion of observed school characteristics: the estimated coefficient for FHC is positive, statistically significant and its quantitative effect is 3.5 pp. Mod.5b further adds controls for individual and parental characteristics and wave I personality traits. In this model, FHC is still statistically significant and, consistently with the estimates of previous models, increases the probability of dropping out by 1.8 pp.

Prior to the work of Cameron and Heckman (1993) GED recipients were considered as high school graduates. Cameron and Heckman show that although GED recipients appear to have similar cognitive skills to graduates, their social behaviours and economic outcomes resemble those of dropouts (see also Heckman and LaFontaine, 2006). Since this is still an open debate and some sources of data (e.g. US Census) do not distinguish between GED and regular HS graduates, we test the robustness of our findings by excluding the future GED-holders and focussing only on 'pure' dropout (see Mod.5c in Table 7).[21] The effect of FHC is highly significant although smaller in size (1 pp), if compared to the coefficient obtained from our preferred specification that includes all dropouts (Mod.3b). However, the effect of FHC corresponds to a 20% variation of the average "pure" drop rate. This also supports the idea that FHC does not depend on cognitive skills.

We also estimate our reference model (Mod.3b) by excluding all missing observations. The sample size decreases to 9,983 observations and the effect of FHC is positive, statistically significant and, consistently with our previous estimates, 1.9 pp in size.[22] This suggests that the dummy variable adjustment approach employed in previous models does not affect our results.

We test whether forgone health care is context dependent and exacerbated by lower socioeconomic status. We do this by adding interactions between FHC and parental education

[21]This implies a loss of 615 observations, and a substantial reduction in the number of dropouts, from 1,513 to 898. The average "pure" dropout rate is 4.8%.

[22]Results are available upon request.

and FHC and job status variables. Our model, excluding all missing observations, shows that the interaction terms between FHC and parental variables are not statistically significant, whereas the effect of FHC is positive, significant and around 3 pp. This implies that the effect of forgone health care should not depend on the socioeconomic status of the family of origin.

5 Conclusions

This paper draws from and merges the medical literature with the economics literature and examines the effects of forgone health care on the decision to drop out from high school. We exploit rich data and detailed information on forgone health care on recent cohorts of young pupils from Add Health. We estimate a series of flexible specifications that progressively account for established determinants of HS dropout together with personality traits, school-level and cohort effects. Throughout these models, forgone health care appears to be a strong and highly statistically significant risk factor for dropout decisions. In the most comprehensive specification the quantitative effect of forgone health care is around 1.7 percentage points. This translates into an increase of the average probability of dropping out of 21.5 percent. Our more conservative estimates find an effect of around 1 percentage points. This also translates into a substantial increase of the average probability of dropping out of 12.7 percent.

Estimates and robustness checks confirm the role played by forgone health care as a signalling device for high school dropout. Our findings suggest that voluntary forgone health care does not depend on cognitive skills. Moreover, the effect of forgone health care on dropout does not appear to be exacerbated by family (parents) socioeconomic status and it is also not affected by observed personality traits, as defined by alternative definitions of the Big Five. We conclude that forgone health care is a relevant predictor of high school dropout and it is likely to be driven mainly by individual unobserved noncognitive traits.

Our work has potentially important policy implications. Given its quantitative relevance, forgone health care may be used as a signalling device to identify and target people at higher risk of dropping out. Information on forgone health care could be easily collected both by integrating students' medical records with educational records, and by adding specific questions on pre-school ability tests or in future surveys concerning adolescents' educational attainment. Finally, after the identification of students at higher risk of dropping out, specific policies could be implemented to raise students and families' awareness on the possible long-term consequences of high school dropout such as higher risk of unemployment and lower lifetime earnings.

References

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.

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).

Blair, C. and Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and psychopathology*, 20(03):899–911.

Blair, C. and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child development*, 78(2):647–663.

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.

Cameron, S. V. and Heckman, J. J. (1993). The nonequivalence of high school equivalents. *Journal of Labor Economics*, 11(1, part 1):1–47.

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*.

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.

Costa, P. T. and McCrae, R. R. (1992). Revised Neo Personality Inventory (NEO-PI-R) and neo five-factor inventory (NEI-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

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.

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

DiPrete, T. A. and Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, 41(1):1–15.

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. *Jama*, 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. *JAmA*, 273(24):1913–1918.

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. (2006). Bias-corrected estimates of GED returns. *Journal of Labor Economics*, 24(3):661–700.

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.

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.

Mendolia, S. and Walker, I. (2014). The effect of personality traits on subject choice and performance in high school: Evidence from an english cohort. *Economics of Education Review*, 43:47–65.

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

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.

Rees, D. I. and Sabia, J. J. (2011). The effect of migraine headache on educational attainment. *Journal of Human Resources*, 46(2):317–332.

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.

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

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.

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.

Acknowledgements

We are grateful to David Figlio, Todd Stinebrickner, Bruce Hollingsworth, Colin Green and Ian Walker for useful comments. We gratefully acknowledge funds from an ESRC "Radical Futures in Society" grant.

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.

Table 1: Descriptive Statistics - Dropout and Individual Characteristics

	All	*Dropout*	*Graduate*	*N*
HS dropout	0.079			19038
FHC	0.274	0.364	0.266	19019
Males	0.494	0.574	0.487	19035
Whites	0.534	0.488	0.538	19038
Hispanics	0.153	0.194	0.150	19038
African-Americans	0.229	0.275	0.226	19038
Asians	0.084	0.044	0.087	19038
Religious	0.861	0.781	0.868	19038
Peabody Test (score)	100.049	91.072	100.823	18107
Learning disability	0.107	0.220	0.097	19038
Health-behaviors	0.190	0.282	0.182	19038
Risky attitude	0.355	0.548	0.338	19038
High discount factor	0.130	0.260	0.118	19038
Ill-health	0.069	0.119	0.064	19038
Counseling	0.086	0.094	0.085	19038
Depression	0.101	0.176	0.094	19038
Asthma	0.117	0.136	0.115	16163
Migraine	0.114	0.172	0.109	16063
Obesity	0.065	0.081	0.063	16198
School absence (health reasons)	0.060	0.129	0.054	19038
Private/prepaid/other cover	0.771	0.519	0.793	16262
Medicare/medicaid	0.107	0.269	0.093	16262
No insurance	0.122	0.213	0.114	16262
Conscientiousness (wave I) (-)	0.148	0.166	0.146	18917
Neuroticism (wave I) (+)	0.080	0.113	0.078	18982
Extraversion (wave I) (-)	0.172	0.222	0.169	13236
Conscientiousness (wave IV) (-)	0.099	0.119	0.097	14566
Neuroticism (wave IV) (+)	0.169	0.271	0.158	14567
Extraversion (wave IV) (-)	0.226	0.287	0.219	14570
Openness to experience (wave IV)(-)	0.043	0.083	0.038	14562
Agreeableness (wave IV) (-)	0.049	0.104	0.043	14568

Table 2: Descriptive Statistics - Family and School Characteristics

	All	*Dropout*	*Graduate*	*N*
Mother				
no education	0.043	0.083	0.040	17895
middle/primary	0.195	0.394	0.178	17895
high school	0.487	0.447	0.490	17895
higher education	0.276	0.076	0.292	17895
Unemployed/at home	0.296	0.414	0.288	12413
routine/technical	0.154	0.211	0.150	12413
small employer/intermediate	0.242	0.185	0.245	12413
managerial/professional	0.309	0.190	0.316	12413
Father				
no education	0.055	0.122	0.050	13909
middle/primary	0.177	0.353	0.165	13909
high school	0.455	0.434	0.457	13909
higher education	0.313	0.090	0.328	13909
Unemployed/at home	0.195	0.292	0.189	9983
routine/technical	0.421	0.493	0.417	9983
small employer/intermediate	0.118	0.111	0.119	9983
managerial/professional	0.266	0.103	0.275	9983
Ill-health (main parent)	0.123	0.211	0.116	19038
Ill-health (partner)	0.086	0.126	0.082	19038
Access to medical care	0.126	0.212	0.118	19038
Family 'bad' behaviours	0.398	0.579	0.382	19038
School Characteristics				
Size: <= 125	0.018	0.028	0.017	18949
Size I: 126-350	0.069	0.076	0.069	18949
Size II: 351-775	0.237	0.220	0.238	18949
Size III: 776	0.675	0.676	0.675	18949
Public	0.930	0.985	0.926	18949
Catholic	0.028	0.009	0.030	18949
Private	0.042	0.006	0.045	18949
Urban	0.296	0.298	0.295	18949
Suburban	0.543	0.496	0.547	18949
Rural	0.161	0.206	0.158	18949
West	0.239	0.192	0.243	18949
Midwest	0.237	0.256	0.236	18949
South	0.376	0.442	0.370	18949
Northeast	0.148	0.109	0.152	18949

Table 3: Dropout and Individual Characteristics

	Mod.1a	*Mod.1b*	*Mod.1c*
FHC	0.034***	0.028***	0.019***
	(0.004)	(0.004)	(0.004)
Demographic Characteristics			
Male	0.023***	0.013***	0.019***
	(0.004)	(0.004)	(0.004)
Hispanic	0.038***	0.015**	0.007
	(0.007)	(0.007)	(0.007)
Afro-American	-0.002	-0.018***	-0.026***
	(0.006)	(0.007)	(0.007)
Asian	-0.015*	-0.026***	-0.026***
	(0.009)	(0.009)	(0.009)
Skills, Behaviors and Preferences			
Religion	-0.049***	-0.037***	-0.029***
	(0.006)	(0.006)	(0.006)
Peabody Test		-0.002***	-0.002***
		(0.000)	(0.000)
Learning disability		0.053***	0.039***
		(0.007)	(0.007)
Health-behaviors		0.044***	0.037***
		(0.005)	(0.005)
Risky attitude		0.036***	0.030***
		(0.004)	(0.004)
High discount factor		0.054***	0.049***
		(0.006)	(0.006)
Health			
Ill-health			0.018**
			(0.008)
Counseling			0.008
			(0.007)
Depression			0.033***
			(0.007)
Migraine			0.022***
			(0.007)
School absence (health reasons)			0.053***
			(0.008)
Medicare/medicaid			0.085***
			(0.007)
No insurance			0.054***
			(0.007)
constant	0.106**	0.311***	0.199***
	(0.050)	(0.054)	(0.054)
N	19016	18093	18093
F	21.470	52.267	42.682

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 Mod.1c but not significant.

Table 4: Dropout and Family Characteristics

	Mod.2a	*Mod.2b*
FHC	0.018***	0.017***
	(0.004)	(0.004)
Ill-health (main parent)	0.006	-0.001
	(0.006)	(0.006)
Ill-health (partner)	0.005	0.009
	(0.007)	(0.007)
Access to medical care	0.011	0.005
	(0.007)	(0.007)
Family health-behaviors	0.033***	0.029***
	(0.004)	(0.004)
Mother		
Middle/primary		0.017
		(0.011)
High school		-0.027**
		(0.011)
Higher education		-0.044***
		(0.011)
Routine/technical		0.001
		(0.008)
Small employer/intermediate		-0.000
		(0.007)
Managerial/professional		0.006
		(0.007)
Father		
Middle/primary		-0.006
		(0.012)
High school		-0.027**
		(0.011)
Higher education		-0.028**
		(0.012)
Routine/technical		-0.004
		(0.007)
Small employer/intermediate		-0.002
		(0.010)
Managerial/professional		-0.003
		(0.008)
constant	0.177***	0.183***
	(0.054)	(0.055)
Individual characteristics	yes	yes
N	18093	18093
F	40.166	31.838

Cohort effects and school fixed effects included in all models.

Missing values dummy variables included in all models.

Table 5: Dropout and Personality traits - wave I

	Mod.3a	*Mod.3b*
FHC	0.030***	0.017***
	(0.004)	(0.004)
Conscientiousness (-)	0.005	-0.004
	(0.006)	(0.005)
Neuroticism	0.028***	0.010
	(0.007)	(0.007)
Extraversion (-)	0.015**	0.009
	(0.006)	(0.006)
Individual characteristics	no	yes
Family characteristics	no	yes
constant	0.030	0.187***
	(0.050)	(0.055)
N	19019	18093
F	28.690	28.939

Cohort effects and school fixed effects included in all models.

Missing values dummy variables included in all models.

Table 6: Dropout and Personality traits - wave IV

	Mod.4a	*Mod.4b*
FHC	0.026***	0.010**
	(0.004)	(0.004)
Conscientiousness (-)	0.011	0.003
	(0.007)	(0.007)
Neuroticism	0.058***	0.040***
	(0.006)	(0.006)
Extraversion (-)	0.015***	0.013**
	(0.005)	(0.005)
Openness (-)	0.075***	0.052***
	(0.011)	(0.011)
Agreeableness (-)	0.094***	0.064***
	(0.010)	(0.010)
Individual characteristics	no	yes
Family characteristics	no	yes
constant	0.080	0.196***
	(0.049)	(0.055)
N	19019	18093
F	35.940	35.554

Cohort effects and school fixed effects included in all models.

Missing values dummy variables included in all models.

Table 7: School Characteristics and Non (future) GED-holders

	Mod.5a	*Mod.5b*	*Mod.5c*
FHC	0.035***	0.018***	0.010***
	(0.005)	(0.004)	(0.004)
Size I	-0.059**	-0.045**	-0.021
	(0.024)	(0.021)	(0.017)
Size II	-0.076***	-0.061***	-0.024
	(0.024)	(0.021)	(0.017)
Size III	-0.078***	-0.056***	-0.020
	(0.024)	(0.022)	(0.018)
Catholic	-0.043***	0.009	0.013
	(0.008)	(0.013)	(0.010)
Private	-0.061***	0.012	0.002
	(0.010)	(0.013)	(0.011)
Suburban	-0.014***	0.001	0.001
	(0.005)	(0.005)	(0.004)
Rural	0.008	0.015**	0.008
	(0.008)	(0.007)	(0.006)
Midwest	0.035***	0.035***	0.034***
	(0.007)	(0.007)	(0.006)
South	0.030***	0.028***	0.020***
	(0.006)	(0.007)	(0.005)
Northeast	0.024***	0.010	0.006
	(0.007)	(0.008)	(0.006)
Individual characteristics	no	yes	yes
Family characteristics	no	yes	yes
Personality traits wave I	no	yes	yes
constant	0.097*	0.146**	0.172***
	(0.051)	(0.058)	(0.046)
N	18930	18012	17577
F	16.787	27.290	23.129

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.

Mod.5c excludes future GED-holders.