NBER WORKING PAPER SERIES

THRIVERS AND DIVERS:
USING NON-ACADEMIC MEASURES TO PREDICT COLLEGE SUCCESS AND FAILURE

Graham Beattie
Jean-William P. Laliberté
Philip Oreopoulos

Working Paper 22629
http://www.nber.org/papers/w22629

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2016

We are very grateful to Aloysius Siow, Nathalie Bau, Yosh Halberstam, Uros Petronijevic for helpful discussions and to seminar participants at the University of Toronto for comments. We also wish to thank Aaron de Mello for spectacular web development and research assistance, and Matthew Hendrickson for helping us compile the administrative data. Financial support for this research was provided by the Ontario Human Capital Research and Innovation Fund, a Social Sciences and Humanities Research Council Insight Grant (#435-2015-0180), and a JPAL Pilot Grant. Any omissions or errors are our own responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Thrivers and Divers: Using Non-Academic Measures to Predict College Success and Failure
Graham Beattie, Jean-William P. Laliberté, and Philip Oreopoulos
NBER Working Paper No. 22629
September 2016
JEL No. I2,I23,J24

ABSTRACT

We collect a comprehensive set of non-academic characteristics for a representative sample of incoming freshman to explore which measures best predict the wide variance in first-year college performance unaccounted for by past grades. We focus our attention on student outliers. Students whose first-year college average is far below expectations (divers) have a high propensity for procrastination – they self-report cramming for exams and wait longer before starting assignments. They are also considerably less conscientious than their peers. Divers are more likely to express superficial goals, hoping to 'get rich' quickly. In contrast, students who exceed expectations (thrivers) express more philanthropic goals, are purpose-driven, and are willing to study more hours per week to obtain the higher GPA they expect. A simple seven-variable average of these key non-academic variables does well in predicting college achievement relative to adding more variables or letting a machine-algorithm choose. Our results, descriptive in nature, warrant further research on the importance of non-linearities for the design and targeting of successful interventions in higher-education.

Graham Beattie
University of Pittsburgh
Department of Economics
4901 Wesley W. Posvar Hall
230 South Bouquet Street
Pittsburgh, PA 15260
grb57@pitt.edu

Jean-William P. Laliberté
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
CANADA
jw.plaliberte@mail.utoronto.ca

Philip Oreopoulos
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
CANADA
and NBER
philip.oreopoulos@utoronto.ca

A online appendix is available at http://www.nber.org/data-appendix/w22629

I. Introduction

In recent decades, college enrollment has increased and both policy makers and parents have continued to emphasize the importance of postsecondary education as a worthy investment for an increasingly competitive labor market, which has directed more attention towards helping entrants actually complete their degrees and exit with valuable experience and skills. But despite efforts to increase college support – additional tutoring, counseling, stress management workshops, time management assistance, and other resources – the fraction of students completing a degree remains alarmingly low. Only about half of students who begin a bachelors' degree in the United States complete it within six years (Symonds, Schwartz and Ferguson, 2011). In Canada, three-quarters complete but many do so with minimum requirements and questionable skill improvement (Arum and Roksa, 2011).[1]

Understanding what factors predict college performance would allow administrators to better target students at risk of struggling and identify incoming skills particularly helpful for academic success. Previous research shows that past performance strongly predicts college achievement,[2] which explains why institutions rely on past grades or standardized tests for admission. But even for students with similar past grades, a high variance exists in subsequent performance. Many arriving with low high school grades perform below average, but some also end up excelling in their new environment and wind up among the top students in their class. Others perform well in high school, but end up struggling in college and eventually drop out. Transitioning from high school to college can be challenging and success in one level of education does not guarantee success in another.[3] When it comes to predicting who among admitted students with similar grades will eventually 'thrive' and who will 'dive', we need to, therefore, look to other measures. College students arrive from an increasing variety of backgrounds with different initial abilities, hopes, goals, and expectations, all of which may influence the degree of ease with which they transition from high-school to college.

Recent research on non-academic abilities suggests that variables aside from past grades may help

[1] Bound and Turner (2011)and Bound, Lovenheim and Turner (2010) discuss recent trends in the US, and Childs, Finnie and Martinello (2016) provide and analysis of Canadian trends.

[2] Bettinger, Evans and Pope (2013); Dooley, Payne and Robb (2012); Cyrenne and Chan (2012); Rothstein (2004).

[3] For example, Scott-Clayton, Crosta and Belfield (2014) discuss how the difficulties associated with identifying at-risk students generate substantial mis-assignment of students to remedial classes.

identify students who are at risk of floundering in college and those who are likely to succeed. There is ample evidence that these skills, particularly personality traits and social background, rival cognitive skills in terms of predictive power for a variety of life outcomes such as educational attainment, earnings, and health.[4] Conscientiousness – a personality trait associated with staying organized, working hard, and persistance – is positively associated with educational achievement independent of intelligence.[5] Gritty students, who persevere towards achieving particular goals, tend to have higher college GPAs than their peers even after conditioning on SAT scores (Duckworth et al., 2007). Also, work by Mischel, Shoda and Rodriguez (1989) and Kirby, Winston and Santiesteban (2005) suggests that the ability to delay gratification also predicts future achievement.

In this paper, we collect a comprehensive set of non-academic characteristics for a large sample of incoming freshman from various backgrounds to explore which measures best predict the large share of the variance in first-year college performance that cannot be accounted for by past grades. We focus our attention on student outliers – those who end up in the bottom and top deciles in our sample, in terms of the difference between actual performance and predicted performance based on high school grades. We call students in the top decile thrivers, and those in the bottom, divers. Thrivers and divers are opposite extremes, making it easier to examine key differences in initial characteristics relative to the rest of the student population. Examining them in isolation helps avoid measuring small linear relationships from the majority of students 'in the middle' and allows for asymmetries between outliers. That a typical B student obtains a GPA of B+, or that a former high school valedictorian receives a GPA of A-, is not out of the ordinary. The search for non-academic predictors of a successful or failed transition to college is magnified by isolating outlier groups that, on their own, are of particular interest. Focusing on them may help administrators better understand how to avoid pitfalls and promote environments for helping incoming college students.

Our data come from partnering with all first-year economics instructors at the three campuses of the University of Toronto and asking students to complete an online 'warm-up exercise' for 2 percent of

[4]Kautz et al. (2014); Almlund et al. (2011); Borghans et al. (2008); Roberts et al. (2007); Heckman, Stixrud and Urzua (2006).

[5]Burks et al. (2015); Almlund et al. (2011); Komarraju, Karau and Schmeck (2009); Poropat (2009); O'Connor and Paunonen (2007); De Fruyt and Mervielde (1996).

their final grade. Over 45 to 90 minutes during the first weeks of school, about 6,000 students – a third of all first year undergraduate students – completed survey questions about procrastination, study habits, social identity, academic expectations, grit, risk aversion, time preference, locus of control, and agreed to link their responses to the university's administrative database of background characteristics and future academic performance.[6] A subset of our data allows us to explore a wide variety of non-academic characteristics at the start of college from asking open-ended questions about aspirations, attitudes, and goals. Our sample is also very large, allowing sufficient statistical power to detect even small differences between performance groups. We explore what variables best predict first-year performance, both unconditionally as well as when conditioning on all other predictors.[7] The exercise does not attempt to uncover causal estimates, but rather document the predictive properties of a large number of characteristics. All relationships between these non-academic variables and college grades are estimated on the same sample, ensuring that the set of controls is consistent and that coefficients' magnitudes are directly comparable.

We find that objective and subjective measures of procrastination and impatience are the best predictors of failing to keep up with grade expectations. Whether conditioning on other traits or characteristics or not, students that self-report tending to cram for exams, wait last minute in general to complete deadlines, or even wait last minute to complete the survey we collected data from them are much more likely to end up in the lowest or second lowest grade decile relative to expectations. Poor performers also tend to work many more hours for pay than their peers and are less conscientious on average. These patterns are not the same for thrivers. The best predictors for far exceeding grade expectations are self-reported intended hours of study and expected grades. Students who expect higher grades tend to get them, and thrivers plan to study more than three hours a week more than divers do, on average.

Another subset of students (one of the experiment's treatment groups) was asked to write freely about their future goals, anticipated setbacks, and mindset. Examining their answers offers the opportunity to vastly expand the set of potential predictors beyond those explicitly measured by

[6] The consent rate was 97 percent.

[7] Access to administrative files will allow us to consider other outcomes such as persistence and academic performance in future work.

questionnaires. We find that thrivers and divers answered these open-ended written questions differently. Thrivers write longer answers and use better spelling than divers, and are also are more likely to identify self-discipline as a trait they admire in themselves. In addition, when asked to identify future goals, thrivers are more likely to discuss the impact they want to make on society, while divers are more likely to emphasize wanting to 'get rich'.

These findings have both theoretical and practical implications. A better understanding of the characteristics of student outliers informs us about the shape of the college education production function. Even among the selected few who make it to college, several noncognitive skills strongly predict how well they do. Accounting for the skills we measure increases the explanatory power in predicting performance compared to using only past performance alone, but high school grades remain the single best predictor of college grades. Also, skills that characterize students who are the most successful in their transition to college are not necessarily the ones that at-risk students lack. On practical grounds, this paper highlights some specific skills that educational policies might target to improve. The abilities to persist, to self-regulate and to set high expectations for oneself all contribute to reducing the risk of struggling in higher education. Our findings also motivate further research on possible policies likely to restrict the scope for the negative effects of behaviors shared by most divers, such as increasing the frequency of deadlines to mitigate procrastination. By helping characterize the profile of students exceptionally poor or great at transition to college, this research may also prove useful for catching students before they run into difficulty, and advising students about how to excel in school.

The paper proceeds as follows. First, we briefly review the existing literature on predictors of college success in Section II. Section III explains the data collection process and the institutional environment and provides an overview of our estimation samples. The methodology is presented in Section IV and results are displayed in section V. Section VI introduces a uni-dimensional "at-risk" factor and benchmarks our results against machine learning results. Section VII concludes with a discussion of the policy implications of this research.

II. Background

Social scientists increasingly stress the importance of noncognitive abilities for a host of socioeconomic outcomes. Both in the labor market and in school, the explanatory power of personality traits and personal preferences is comparable or greater than that of cognitive abilities (Almlund et al., 2011). In a similar vein, successful childhood interventions that have long-term impacts on adult outcomes often show no persistent effect on cognitive skills while significantly improving children's non-academic skills (Kautz et al., 2014; Chetty et al., 2011). Academic performance in high school as well as in college likely reflects both the cognitive and noncognitive abilities of students.

The emphasis on personality traits and other non-academic measures as determinants of educational success has a long tradition in the fields of education and psychology.[8] In recent decades, the emergence of the Big Five dimensions of personality as a broadly accepted general taxonomy (John, Naumann and Soto, 2008), along with an increasing interest in motivational theories (Robbins et al., 2004), generated a substantial amount of research on the incremental effects of personality and individual goals on college success over that of standard predictors such as standardized tests (Conard, 2006). The number of noncognitive measures that have been found to correlate significantly with college GPA is large. Yet, it remains unclear which of them or which set constitute the best predictors of success in college, since few studies consider a broad selection of predictors simultaneously and many distinct measures considerably overlap conceptually and empirically. The lack of a thorough evaluation of how different measures used in separate literatures are related has rendered integration of independent findings difficult.

For example, conscientiousness[9] and grit[10] , which have been the focus of most personality research, are both strong predictors of postsecondary education performance, but recent evidence suggests that the latter might be a facet of the former (Credé, Tynan and Harms, 2016; Dumfart and Neubauer, 2016). In parallel, the literature on motivational theories has emphasized the importance of goals and beliefs about performance. The most comprehensive meta-analytic reviews in psychology and education research generally find that academic self-efficacy – the belief in one's

[8] Willingham (1985) provides an excellent overview of the early work on the topic.
[9] Burks et al. (2015); Komarraju, Karau and Schmeck (2009); Poropat (2009).
[10] Duckworth et al. (2007).

capability to succeed academically – and grade goals – exhibit the strongest correlations with college GPA (Richardson, Abraham and Bond, 2012; Robbins et al., 2004).[11] More recently, researchers in economics of education have emphasized the role of time preferences as important inputs in schooling decision and in the educational production function.[12]

These separate branches of research in education have yet to integrate findings from one another. Our paper casts a wider net by considering multiple predictors from all three fields simultaneously, notably including standard personality constructs, measures of motivational factors previously found to be good predictors of college GPA such as locus of control and grade expectations, as well as economic preference parameters. We further broaden the set of predictors by moving beyond traditional questionnaire-based measures through text analysis, and complement our examination with machine learning techniques.

III. Data

Our data comes from an online exercise completed by first year economics students in all three campuses of the University of Toronto. While more than half of the university's student population attend the main campus, over 25,000 students are registered at two smaller satellite campuses to the West and East of downtown, both about 20 miles away. These campuses receive more commuter students than the main campus and have different admission requirements. As a result, the university's student population comes from a very diverse set of academic backgrounds.

Early in the 2015 Fall semester, all undergraduate students enrolled in an introduction to economics course (approx. 6,000) across all three campuses were asked to participate in an online 'warm-up' exercise. The nature of the exercise varied randomly across students – some were asked to complete a comprehensive personality test while others were assigned a goal-setting program which requires from them to write freely about their future goals. Each group was shown a short video created to introduce the purpose of the program and key take-away points. Beforehand, students were

[11] While these meta-analyses consider many characteristics as predictors, the underlying studies rarely do, plausibly introducing bias. Our setup overcomes this methodological drawback.

[12] Lavecchia, Liu and Oreopoulos (2016); Cadena and Keys (2015); Burks et al. (2015).

required to fill in a brief survey and were asked for consent to work with their administrative data (97 percent agreed). Completion of this one- to two-hour exercise counted for 2 percentage points of their overall grade in the course.[13]

The group of students who took part in the program represents about a third of all first year students enrolled at this university, and almost 10% of the entire undergraduate student population. Linked administrative variables include gender, citizenship, registration status, GPA, all courses taken and grades received at this postsecondary institution and, for the majority of students, the high school performance measure used for admission to Canadian universities (the *admission grade*).[14] In the analyses below, we restrict our estimation sample to students for which we have the high school achievement (77 percent of the sample).

The set of variables that was collected as part of the survey from all students contains detailed background characteristics such as international student status and parental education, as well as a large set of new measures of noncognitive skills, in particular reports of study habits and subjective expectations. Survey questions are presented in the Online Appendix.

For a 30 percent random subsample of students (henceforth the *personality sample*), we collected additional data on a large array of traditional personality traits and economic preference measures as part of the online exercise. These include self-assessed propensity to procrastinate and summary measures of perseverance of effort and consistency of interest, two latent factors loading onto the construct of grit (Duckworth and Quinn, 2009). Two complementary measures of each Big Five trait[15] were also constructed: an absolute measure obtained by implementing the Likert-scale Mini-IPIP questionnaire (Donnellan et al., 2006), and a relative-scored ipsative measure. The ipsative measure indicates the extent to which a given trait is dominant in one's personality profile relative to other traits. This relative-scored method is known to be more resistant to biased responding

[13]The warm-up exercise was setup, in part, to test the effectiveness of new online and text-based approaches for providing student support. For more information about the experimental design, we refer readers to Oreopoulos and Petronijevic (2016).

[14]This corresponds to the student's average grade for a standardized set of high school courses taken by all students in the province of Ontario. Admission to postsecondary education in Ontario is based solely on academic performances. There is no admission criterion, implicit or explicit, based on personal characteristics such as race, ancestry, ethnic origin, sex or age.

[15]The five traits are agreeableness, conscientiousness, extraversion, openness to experience, and emotional stability.

(Hirsh and Peterson, 2008).[16] We also assess students' level of tolerance for risk using both a simple survey question as well as a series of hypothetical choices between a lottery and a certain amount of money (Dohmen et al., 2011, 2010). Finally, we elicit time preferences using lists of hypothetical choices between an amount of money paid at some early point in time and a larger amount received later (Dohmen et al., 2010; Andersen et al., 2008).

The first column of Table 1 shows descriptive statistics for all students included in the personality sample for whom the admission grade is non-missing. The average admission grade is 87 percent with the majority of students scoring above 80.[17] The summary statistics for demographic variables underline the sample's diversity. Roughly half the students have a mother tongue other than English and a citizenship other than Canadian, and a third self-report as international students.[18] Approximately 53 percent are women, and 81 percent started their first year of university in the Fall of 2015. More than 40 percent of our sample intends to major in a field other than economics or business (the two programs for which the introduction to economics course is required). Only 25 percent are first-generation college students (i.e. neither of their parents is college-educated).

There is substantial variation in average first-year college grades. The mean is 66 percent with a standard deviation of 13 percentage points, almost three times larger than the standard deviation of admission grades.[19]

In terms of study habits, students expect to study for approximately 18 hours per week on average and work at a paid job for less than 8 hours per week. Students come in with high expectations: approximately 63 percent intend to eventually pursue graduate studies,[20] and the average expected GPA is 3.6, more than one grade point above the actual first-year mean GPA (2.3) – a difference

[16]The relative-scored measure combines rank-order and forced-choice approaches. The main drawback to this approach is that relative-scored traits are negatively correlated with each other by construction. See the Online Appendix for details.

[17]In terms of high school performance, our sample is reasonably close to the provincial average for those enrolling in university. The most recent application data from the Council of Ontario Universities (2014) indicates that the secondary school average of Full-Time, First Year students at the University of Toronto is 85.9%. The average across Ontario universities is 83.4% with some institutions with entering average grades above 86%.

[18]In practice, domestic student are those with either a Canadian citizenship or a Permanent Resident status.

[19]By construction, the distribution of admission grades we observe is truncated at the bottom. It does not reflect the full distribution of potential applicants as it only includes enrollees. This restriction of range raises methodological issues if one tries to extrapolate the relationship between past grades and college grades to non-enrolled students (Rothstein, 2004). Our objective in this paper is not to inform admission policy and the interpretation of our results is independent of restriction of range issues.

[20]In comparison, only 20% of the university's student population is enrolled in a graduate program.

greater than a full standard deviation. In addition to these subjective expectations, we also consider an objective measure of procrastination, which is the number of days between the first day of class and the time a student started the online survey for this study. Students were encouraged to complete the task early before being burdened with other homework, and given a two-week deadline. On average, four days passed between the beginning of classes and the moment students started the survey, with about half the sample registering within 2 days, but a fifth of students waiting more than a week.

In complementary analyses, we focus on a separate 50% random subsample of students (henceforth the *text sample*) who were asked to answer open-ended questions such as "describe what kind of person you want to become later on in life".[21] The qualitative answers to each of these questions provide sufficient information to analyze whether outliers tend to discuss different topics than other students when they are allowed to choose what to write about. Students were prompted to take their time and take the exercise seriously because it was intended for their benefit. Some questions contained word count and time constraints, with a friendly message of encouragement to students that tried to complete a question before removing these constraints. The large majority of students wrote in detail, with emotion, clarity and personal insight.

IV. Methodology

IV.A. Defining outliers

Admission to college generally relies on standardized tests or high school grades. Yet, substantial variation in freshman performance around past grades remains. High school GPA alone is not sufficient to predict which students are the most likely to struggle and eventually drop out of college. The methodology developed below aims at exploring whether adding more variables is useful for improving predictions. To emphasize the incremental predictive power of non-academic

[21] 50% of first-year students and 70% of upper-year students were randomly assigned to a goal-setting exercise. The proportion of first-year/upper-year students was unknown prior to assignment. Overall, about 53% of students who took part in the warm-up exercise were assigned to the goal-setting exercise. By construction, the personality sample and the text sample are mutually exclusive.

characteristic, we focus on the part of college grades that cannot be expected on the basis of past grades.

To identify students who perform unusually above or below expectations, we first residualize college grades on past performance. More specifically, we extract the portion of college grades that is linearly predicted by past grades and a set of background characteristics by estimating the following equation:

$$CollegeGrade_{ics} = \alpha_0 + \alpha_1 HSGRADE_{ics} + \alpha_2 \kappa_{ics} + \delta_c + \delta_s + \epsilon_{ics} \tag{1}$$

where $CollegeGrade_{ics}$ is the credit-weighted first-year average college grade of student i who started college in semester s and at campus c. Campus fixed-effects are included to take into account differences in admission criteria across campuses, as well as any discrepancy in grading practices. Upper year students included in our sample are more likely to be enrolled in STEM programs and to take introduction to economics as an elective than are first-year students. Therefore, cohort fixed-effects are added to the model. We estimate the model separately for the personality sample and the text sample.

Figure 1 plots residualized college grades against admission grades for the personality sample. In both dimensions, we highlight students who belong to either the top or the bottom decile of the distribution. The vast majority of students who perform significantly above expectations (groups 1, 4 and 7 on the figure) or below expectations (groups 3, 6 and 9) come from the middle of the admission grade distribution. Put differently, students who thrive are not simply students who were already expected to do well and did even better, and students who dive are not merely students who were expected to have relatively low grades and did even worse. Yet, the performance gap between the two outlier groups is colossal: divers' average first-year college grade is 40, and thrivers' is 81. In our main specifications we define the two groups of students who rank in the top and bottom deciles of the distribution of ϵ_{ics} as thrivers and divers, respectively. We explore the robustness of our results with respect to the definition of divers and thrivers in section V.B.

IV.B. Differences in quantitative non-academic measures

The main exercise we undertake compares the distributions of a large set of non-academic measures for the two outlier groups relative to the full sample. Unconditional mean differences for each characteristic $x \in \mathbf{X}$ are obtained from the following regression:

$$x_i = \gamma_1 D_i + \gamma_2 T_i + u_i \tag{2}$$

where x_i is a non-academic measure, D_i is a dummy for diver status and T_i is a dummy for thriver status. To ease the interpretation of the results, each non-binary individual characteristic of interest is standardized with mean zero and unit variance. For continuous predictors, the coefficients of interest, γ_1 and γ_2, indicate the difference in mean for each outlier group relative to the main distribution, in standard deviations units. Correspondingly, binary measures are centered such that their mean is zero and the estimated coefficients reflect the percentage point difference in the fraction of thrivers or divers who exhibit the characteristic of interest relative to the main sample.

As discussed in section II, there is substantial conceptual overlap between different non-academic constructs. To find which of these measures are the best predictors of success and failure in transitioning to college, we assess whether the mean differences remain significant when using only variation in the distribution of a given characteristic that is unexplained by other predictors. These conditional differences are calculated in two-steps. First, we residualize each characteristic x:

$$x_i = a + \mathbf{b}\mathbf{X}_{-\mathbf{x},\mathbf{i}} + v_{x,i} \tag{3}$$

where $\mathbf{X}_{-\mathbf{x}}$ is the subset of $\mathbf{X}$ that excludes characteristic x. Then, differences in means of residualized characteristics are obtained by substituting $v_{x,i}$ for x_i in equation (2). This strategy amounts to comparing the outlier distributions with the main distribution using only the fraction of the variation in a given construct that is unexplained by other non-academic measures.[22]

Figure 2 illustrates the nature of the comparison exercise. In panel A we show the unconditional distributions of (relative-scored) conscientiousness for thrivers, divers and the full personality sample. Divers are considerably less conscientious than average (0.26 standard deviations below the sample mean). This pattern is not symmetric – on average, thrivers are just as conscientious as others. Conditional differences are presented in panel B, where each density plot shows the distribution of residual conscientiousness that is unaccounted for by variation in other non-academic measures. The mass of divers with very low conscientiousness (around -2 s.d.) observed in the unconditional distribution is explained by other predictors, but the conditional mean difference between divers and the full sample remain substantial. Figures A1 and A2 show similar density plots for other non-academic measures.

IV.C. Text analysis

Students in the text sample were asked a series of open-ended questions, such as "Name at least one thing that you admire about yourself". This type of question allows students much more freedom to answer, so individual answers are often very informative. For instance, answers are not restricted to a set of goals pre-selected by the researcher, but rather include any goals students may have. However, aggregating the results over all students in a meaningful way is a challenge. We use two techniques to quantify the writing, one evaluating effort and writing quality, the other analyzing which topics students choose to write about.

There are three measures of effort and writing quality. Firstly, the programming of the survey website allows us to measure how many seconds each student takes to answer a given question. Secondly, we count the number of words each student uses for each answer, where a word is defined as one or more characters separated by one or more spaces. Finally, we run each of these words through the Microsoft Word Canadian English spellchecker, and calculate the proportion of words which are spelled correctly.[23] These variables are taken as measures of conscientiousness and language ability

[22]This is similar but not numerically equivalent to including all other characteristics as controls in equation (2).

[23]Note that this is a noisy measure of spelling quality. If a student's misspelling of a word is a correct spelling of

and analyzed using the method described in Section IV.B.

We also compare the topics that divers and thrivers discuss in their answers using a simplified topic-modelling text analysis approach. In a topic modelling approach to text analysis, it is assumed that an author makes a series of decisions about which topics to discuss. Each topic then maps to a series of words (Blei, Ng and Jordan, 2003; Hofmann, 2000). For example, discussion of procrastination might use words like "procrastinate", "cram", or "all-nighter". From a researcher's perspective, both the topic choice and the mapping from topics to words are latent. The researcher compares the frequency of words across documents to determine, using the assumption that if one document uses more words that are related to a given topic, the author of the document is devoting more space to that topic.

Given the sample size and the fact that students often give brief responses, we adopt a very simple method to apply this approach. Firstly, we clean the students' answers to generate more meaningful results with the following rules: If a word was spelled incorrectly according to the Word spellchecker, we replace it with Word's top suggestion for a replacement. These words are then stemmed, to remove grammatical constructions such as pluralisation and verb tenses. This ensures that words such as "class" and "classes" are treated as identical. Finally, we remove stopwords, which are short, common words such as "and" or "the".

For each word in the cleaned text, we calculate the proportion of students who use the word to answer a given question among divers, thrivers, and in the entire sample. A chi-squared test comparing the share of divers who use a word with the share of the entire sample shows if low performing students are more likely than others to use a given word. If many of the words used more often by divers are related to a given topic, the intuition of the topic modelling approach suggests that divers are more likely to spend more space discussing that topic.

another word – for example, "coarse" for "course" – it will count as a correct spelling. On the other hand, some widely acceptable abbreviations, such as GPA, are not recognized by the spellchecker and counted as incorrect spellings.

V. Results

V.A. Predicting college grades using past academic achievement

Table 2 presents estimates of the relationship between past academic performance and college grades for the personality sample.[24] To ease interpretation, both college and admission grades are normalized with mean zero and a standard deviation of one. Column (1) shows estimates that are only adjusted for campus and cohort fixed effects. Age at entry and non-domestic student status, two demographic characteristics observed by the registrar, are further added as covariates in column (2).[25]

A one standard deviation higher admissions grade is associated with a 0.41-0.43 standard deviation higher first-year average college grade. Older students and non-domestic students receive lower grades in college than do younger and domestic students with equivalent admission grades. While past measures of academic performance do predict success in college, the explanatory power of this model is modest. When no demographics are included, less than 20% of the observed variation in college grades is explained by admission grades, in line with previous findings (Stephan et al., 2015; Bettinger, Evans and Pope, 2013; Richardson, Abraham and Bond, 2012). The inclusion of age at entry and non-domestic student status adds some explanatory power, but more than three quarters of the variation in college grades remain unexplained. We next explore which non-academic characteristics best explain this variance by comparing outliers with the main distribution.

V.B. Predicting Student Outliers with Non-Academic Outcomes

Columns (1) and (3) of Table 3 report average unconditional deviations from the sample mean for divers and thrivers, respectively. For each possible predictor, columns (2) and (4) report deviations from the mean conditional on all other predictors listed in the table. In the last two columns, we test whether the difference between the top and bottom outliers for each non-academic measure is

[24] We focus on this sample because it includes the full set of measured non-academic outcomes. Since group assignment is random, results are the same for the text sample.

[25] We consider a student to be non-domestic if he or she both self-reports as an international student and doesn't have a Canadian citizenship.

significantly different from zero.

Relative to the full distribution, students who perform largely below expectations are much more likely to self report they cram for exam (0.30 s.d. above the mean), much more likely to start the online survey later (0.29 s.d. above the mean) and tend to work much more hours at paid jobs (0.22 s.d. above the mean). They are also significantly less conscientious (0.26 s.d. below the mean) and more impatient than their peers (0.2 s.d. above the mean), consistent with prior evidence (Burks et al., 2015). Even conditional on other predictors, most of these patterns remain strong and statistically significant. Being sure about one's major and intending to pursue graduate studies has little explanatory power, and, if anything, divers are more likely to say they often think about the future. We interpret these results as evidence that students who perform significantly below expectations are neither lacking ambition nor vision, but tend to put themselves in situations that hinder their academic success.

Thrivers are not the mirror image of divers; they are no less likely to cram for exams or to work many hours for pay than the average student. However, they tend to study for relatively more hours (0.22 s.d. above the mean), and expect a higher GPA than divers (difference of 0.23 s.d.). We find that thrivers are more introverted than divers (unconditional difference of -0.27 s.d.), but that the conditional difference is not statistically significant.[26] Relative to the full distribution, thrivers are more risk averse, but this difference is mostly accounted for by variation in other characteristics. Students who excel above expectations do not report finding the transition to university any less challenging than the average student does, and intend to pursue graduate studies in the same proportions as average students and divers do.

We find no statistically significant differences between outliers in terms of agreeableness, openness to experience or emotional stability. Similarly, grit (perseverance of effort and consistency of effort) and locus of control do not help predict extreme outcomes. The point estimates for our subjective measure of procrastination indicate that thrivers are less likely to procrastinate than divers, but we cannot reject the null hypothesis of no difference.

[26]While less common in the literature, this result is not entirely new (O'Connor and Paunonen, 2007; Noftle and Robins, 2007). Chamorro Premuziz and Furnham (2005) discusses how introverts may have a greater ability to consolidate learning and have better study habits (e.g. spend more time studying than socializing).

Men are overrepresented in both tails of the distribution of college grade residuals: the proportion of women is approximately 10 percentage point lower among divers and thrivers than in the full sample. Previous research has also found that boys exhibit higher variance in test scores than girls (Machin and Pekkarinen, 2008; Hedges and Nowell, 1995). Other demographic characteristics have little or no predictive power.

To evaluate the robustness of these findings, we consider two alternative definitions of outliers. In table A1, divers (thrivers) are defined as students who fall in the bottom (top) 20% of the distribution of college grade residuals. Broadening the groups' composition improves precision, but may dilute results by including students with less extreme outcomes in the outlier groups. We continue to find that divers are less conscientious and more impatient, more likely to cram for exam and to start the exercise later, and that thrivers study more hours on average. Under this specification, the difference between divers and the full sample in terms of procrastination does reach statistical significance at conventional levels, but differences in gender composition do not. In table A2, we verify that our results are not driven by students who came in with extraordinarily high or low admission grades by restricting the sample to those with past grades in the middle 80% of the distribution. [27]

Overall, students who perform markedly lower than expected given their past academic achievement are less prone to procrastination, more impatient and less conscientious than the average. In contrast, students who perform significantly better than expected exhibit very few differences with the full sample, with the exception of the number of hours spent studying. This suggests that most measured characteristics conductive to success in college are already reflected through high school grades, but that non-academic measures do help predict negative outcomes that were unexpected on the basis of past performance.

For completeness, table A3 shows differences between students in the top and bottom distribution of college grades that are not adjusted for high school grades, so that students in the right tail also include students who do very well and were expected to do so.[28] While results for students

[27] We restrict the sample to students in groups 4, 5 and 6 on Figure 1.

[28] We here use the distribution of grades adjusted only for cohort and campus fixed effect, as well as age at entry and non-domestic status.

in the bottom 10% are essentially the same as in our main analyses, the picture for top students is noticeably different: relative to the main distribution, students with the best college grades are more conscientious and less extroverted, less likely to cram for exams, expect higher GPAs and are significantly less tolerant of risk. Our interpretation is that these characteristics contribute to success both in college and in high school, but cannot explain why some students thrive beyond expectations. Overall, our findings suggest that effort (study hours), rather than conscientiousness or patience, is the key predictor to an exceptionally successful transition to college.

V.C. Text Analysis of Student Outliers

The analysis of open ended questions yields results that are consistent with the main results. Table 4 replicates the methodology in Table 3, and shows that divers use fewer words when answering questions and thrivers use more, which suggests that thrivers are providing more detailed and careful answers. Thrivers spend more time answering these questions, although this difference is not significant. Finally, in the unconditional comparisons, thrivers have stronger spelling than the average and divers have weaker spelling. The significance of this result dissipates in the conditional comparisons, which suggests that most of this difference is a function of other non-academic variables.

Topic analysis results are shown in Table 5. For a selected set of questions, the table shows a word if the difference between the share of thrivers (or divers) who use it and the share of the whole sample is significant at 5%, and if at least 5 thrivers (or divers) use it. These results reinforce the point that conscientiousness is a crucial trait. When asked to identify traits they admired about themselves, thrivers were more likely to use words such as "discipline", "practice", or "responsibility", which are indicative of conscientiousness. Sample phrases using these words include "I admire the fact that I have discipline", and "One of the qualities I admire most about myself is responsibility".[29]

Thrivers and divers can also be differentiated when they are asked to list their goals or hopes for the future. Divers are significantly likely to use words which highlight wealth. Examples include "rich", in contexts such as "be a rich man" and "business" in contexts such as "being successful, having so

[29]Appendix tables A5-A8 show a list of sample phrases for each of the words in Table 6.

many successful businesses." Thrivers, on the other hand, are more likely to highlight how they plan to contribute to society, using words such as "human" and "people". Previous work has emphasized the importance for educational success of pursuing long-term goals. Our text analysis stresses the importance of the nature and content of these goals.

VI. Summary Measures of Non-Academic Characteristics

We combine our key non-academic predictors of college success and failure into an overall predictor, to examine how well it performs in forecasting outcomes not anticipated using previous grades. Our seven robust predictors of outliers are: propensity to cram for exams, number of hours studying, number of hours of paid work, expected GPA, time started the exercise, conscientiousness and impatience. We remain agnostic about the exact relative importance of each of these seven constructs and take the unweighted average of these standardized variables for each student. We later explore whether we can improve our predictions by using weights obtained from machine learning techniques.

Figure 3 shows the distributions of unadjusted college grades for students in the top and bottom 10% of the distribution of our relatively simple summary measure of non-academic characteristics. Students deemed the most at-risk under this metric have first-year grades on average more than a full standard deviation below students considered least at risk of struggling during the transition to college.

The one-dimensional measure performs relatively well in terms of predicting freshman performance, but its incremental explanatory power after accounting for past performance is modest, as shown in Table 6. The table displays estimates of a modified version of equation (1) in which one-dimensional at-risk measures are substituted for admission grades in panel A, and added as regressors in panel B. Our preferred metric, the simple unweighted average of 7 large predictors of extreme unexpected performance, correlates strongly with college grades (column (3)). In terms of adjusted R-squared, the explanatory power of this measure alone (0.163) is not as high as that of high school grades (0.218), but adding the at-risk factor to past grades increases the model's fit by almost 4 percentage points.[30]

We then benchmark the predictive abilities of our summary at-risk factor against measures computed with more sophisticated but less transparent approaches to constructing indices. In column (4), the seven best predictors are summarized by their principal component. The model's fit is actually lower with this method than under our preferred approach, suggesting that only using the variance common to all 7 variables is too restrictive.

In columns (5) to (9) we use least angle regressions (LARS) and let the algorithm pick the best predictors and put optimal weights on these (Efron et al., 2004). A summary measure of characteristics can then be defined as the fitted values associated with the LARS estimates. The dependent variable used in the process is average college grade adjusted for our conditioning variables, but not for high school grades. Since we chose our 7 best predictors by examining outliers, we first run the LARS algorithm on the subsample of divers and thrivers, as they are defined in section IV.A. Comparing the adjusted R-squared in columns (5), (6) and (7) with column (3) indicates that the weights put on the seven selected predictors that maximize the share of the variance in grades *among* outliers do not necessarily generalize to the full distribution since the unweighted average has better predictive power over the full personality sample. This observation underscores the importance of non-linearities in the education production function. Column (8) demonstrates that the fit of the model is minimally improved by letting the algorithm pick more predictors than the ones we selected. The summary measure used in column (9) is obtained using LARS on the full distribution of students in the personality sample and therefore puts an upper bound on the joint predictive power of all the non-academic characteristics we observe on that sample. We find that our simple summary measure raises the adjusted R-squared almost as high as this upper-bound measure does, but without compromising on transparency.

Results on outliers highlight important asymmetries in the distribution of non-academic characteristics across the grade distribution, suggesting that the at-risk factor may have more predictive power for extreme outcomes than over the entire distribution of grades. Table A4 shows the proportion of students considered most or least at-risk under different criteria who fall in the bottom

[30]The adjusted R-square is 0.096 when only conditioning variables (cohort and campus fixed effects, age, non-domestic status) are included.

of the distribution of first-year college grades. About a quarter of all students below the 10th percentile of admission grades end up below the 10th percentile of college grades. The proportion of students deemed 'at-risk' by our simple measure that ends up with such dramatic outcomes is very similar (23%).[31] Yet, there is little overlap in the tails of the distributions of admission grades and of our at-risk measure – only 2.2% of students in our sample fall below the 10th percentile in both distributions. But among students in this situation, 40% will fall in the bottom decile of college grades, and all of them will end up below the median. Importantly, falling below the 10th percentile of college grades may have serious consequences: In our sample, no student in the bottom decile of college grades has an average grade above 50 and all are therefore put on probation at the end of their first year in college, which substantially reduce the probability of graduating (Lindo, Sanders and Oreopoulos, 2010). When used jointly with high school grades, the at-risk factor can substantially improve the prediction of extreme outcomes, with potentially important benefits for school administrators and students alike.

VII. Conclusion

A vast array of personality traits and other non-cognitive constructs are used in education research in order to predict performance in college, with substantial overlap across distinct measures. In this paper, we gathered a comprehensive set of non-academic measures for a large and diverse sample of incoming freshman. We investigated which of these variables, unconditional and conditional on other predictors, best explain the variation in college grades that could not have been expected on the basis of variables known upon admission, notably past academic performance. Our results suggest that a few non-academic measures have reasonable predictive power and that linear assumptions often implicit in prior research mask important asymmetries.

Students whose first-year college average is far below expectations (divers) have a high propensity for procrastination – they self-report cramming for exams and wait longer before starting a short exercise worth 2 percent of their overall grade in a first-year economics course. They are also

[31] In contrast, the proportion of least 'at-risk' students falling below the 10th percentile of college grades is only 2%.

considerably less conscientious than their peers. Divers are generally more impatient for positive experiences. For instance, qualitative analyses of short texts written by students suggest divers are more likely to express superficial goals, hoping to 'get rich' quickly. In contrast, students who exceed expectations (thrivers) express more philanthropic goals, are purpose-driven, and are willing to study more hours per week to obtain the higher GPA they expect. The only background characteristic that help predict outlier status is gender, with men being more likely to both thrive and dive.

Consistent with the extensive literature on the correlates of college GPA, we found that high school grades remain the best predictor of college grades. However, non-academic constructs are especially useful for predicting extreme outcomes that cannot be explained by prior educational achievement. Our findings speak directly to initiatives targeting at-risk students. Importantly, the characteristics that best predict successful transitions to college are not necessarily the ones that struggling students lack. Our results, descriptive in nature, warrant further research on the importance of non-linearities for the design and targeting of successful interventions in higher-education.

References

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Chapter 1 - Personality Psychology and Economics." In *Handbook of The Economics of Education.* Vol. 4 of *Handbook of the Economics of Education*, , ed. Stephen Machin Eric A. Hanushek and Ludger Woessmann, 1 – 181. Elsevier.

Andersen, Steffen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutström. 2008. "Eliciting risk and time preferences." *Econometrica*, 76(3): 583–618.

Arum, Richard, and Josipa Roksa. 2011. *Academically adrift: Limited learning on college campuses.* University of Chicago Press.

Bettinger, Eric P, Brent J Evans, and Devin G Pope. 2013. "Improving college performance and retention the easy way: Unpacking the ACT exam." *American Economic Journal: Economic Policy*, 5(2): 26–52.

Blei, David M, Andrew Y Ng, and Michael I Jordan. 2003. "Latent dirichlet allocation." *Journal of Machine Learning Research*, 3(Jan): 993–1022.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel. 2008. "The economics and psychology of personality traits." *Journal of human Resources*, 43(4): 972–1059.

Bound, John, and Sarah Turner. 2011. "Chapter 8 - Dropouts and Diplomas: The Divergence in Collegiate Outcomes." In *Handbook of The Economics of Education.* Vol. 4 of *Handbook of the Economics of Education*, , ed. Stephen Machin Eric A. Hanushek and Ludger Woessmann, 573 – 613. Elsevier.

Bound, John, Michael F Lovenheim, and Sarah Turner. 2010. "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources." *American Economic Journal: Applied Economics*, 2(3): 129–157.

Burks, Stephen V, Connor Lewis, Paul A Kivi, Amanda Wiener, Jon E Anderson, Lorenz Götte, Colin G DeYoung, and Aldo Rustichini. 2015. "Cognitive skills, personality, and economic preferences in collegiate success." *Journal of Economic Behavior & Organization*, 115: 30–44.

Cadena, Brian C, and Benjamin J Keys. 2015. "Human capital and the lifetime costs of impatience." *American Economic Journal: Economic Policy*, 7(3): 126–153.

Chamorro Premuziz, Tomas, and Adrian Furnham. 2005. "Personality and intellectual competence."

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *The Quarterly Journal of Economics*, 126(4): 1593–1660.

Childs, Stephen E., Ross Finnie, and Felice Martinello. 2016. "Postsecondary Student Persistence and Pathways: Evidence From the YITS-A in Canada." *Research in Higher Education*, 1–25.

Conard, Maureen A. 2006. "Aptitude is not enough: How personality and behavior predict academic performance." *Journal of Research in Personality*, 40(3): 339–346.

Credé, Marcus, Michael C Tynan, and Peter D Harms. 2016. "Much Ado About Grit: A Meta-Analytic Synthesis of the Grit Literature." *Journal of Personality and Social Psychology.*

Cyrenne, Philippe, and Alan Chan. 2012. "High school grades and university performance: A case study." *Economics of Education Review*, 31(5): 524–542.

De Fruyt, Filip, and Ivan Mervielde. 1996. "Personality and interests as predictors of educational streaming and achievement." *European journal of personality*, 10(5): 405–425.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are risk aversion and impatience related to cognitive ability?" *The American Economic Review*, 100(3): 1238–1260.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. 2011. "Individual risk attitudes: Measurement, determinants, and behavioral consequences." *Journal of the European Economic Association*, 9(3): 522–550.

Donnellan, M Brent, Frederick L Oswald, Brendan M Baird, and Richard E Lucas. 2006. "The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality." *Psychological assessment*, 18(2): 192.

Dooley, Martin D, A Abigail Payne, and A Leslie Robb. 2012. "Persistence and academic success in university." *Canadian Public Policy*, 38(3): 315–339.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly. 2007. "Grit: perseverance and passion for long-term goals." *Journal of personality and social psychology*, 92(6): 1087.

Duckworth, Angela Lee, and Patrick D Quinn. 2009. "Development and validation of the Short Grit Scale (GRIT–S)." *Journal of personality assessment*, 91(2): 166–174.

Dumfart, Barbara, and Aljoscha C Neubauer. 2016. "Conscientiousness Is the Most Powerful Noncognitive Predictor of School Achievement in Adolescents." *Journal of Individual Differences*.

Efron, Bradley, Trevor Hastie, Iain Johnstone, Robert Tibshirani, et al. 2004. "Least angle regression." *The Annals of statistics*, 32(2): 407–499.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

Hedges, Larry V, and Amy Nowell. 1995. "Sex differences in mental test scores, variability, and numbers of high-scoring individuals." *Science*, 269(5220): 41.

Hirsh, Jacob B, and Jordan B Peterson. 2008. "Predicting creativity and academic success with a "fake-proof" measure of the Big Five." *Journal of Research in Personality*, 42(5): 1323–1333.

Hofmann, Thomas. 2000. "Learning the Similarity of Documents: An Information-Geometric Approach to Document Retrieval and Categorization." 914–920.

John, Oliver P, Laura P Naumann, and Christopher J Soto. 2008. "Paradigm shift to the integrative big five trait taxonomy." *Handbook of personality: Theory and research*, 3: 114–158.

Kautz, Tim, James J. Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. 2014. "Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success." OECD Publishing OECD Education Working Papers 110.

Kirby, Kris N, Gordon C Winston, and Mariana Santiesteban. 2005. "Impatience and grades: Delay-discount rates correlate negatively with college GPA." *Learning and individual Differences*, 15(3): 213–222.

Komarraju, Meera, Steven J Karau, and Ronald R Schmeck. 2009. "Role of the Big Five personality traits in predicting college students' academic motivation and achievement." *Learning and individual differences*, 19(1): 47–52.

Lavecchia, Adam M, Heidi Liu, and Philip Oreopoulos. 2016. "Chapter 1 - Behavioral Economics of Education: Progress and Possibilities." In . Vol. 5 of *Handbook of the Economics of Education*, , ed. Stephen Machin Eric A. Hanushek and Ludger Woessmann, 1 – 74. Elsevier.

Lindo, Jason M, Nicholas J Sanders, and Philip Oreopoulos. 2010. "Ability, gender, and performance standards: Evidence from academic probation." *American Economic Journal: Applied Economics*, 2(2): 95–117.

Machin, Stephen, and Tuomas Pekkarinen. 2008. "Global sex differences in test score variability." *Science*.

Mischel, Walter, Yuichi Shoda, and Monica L Rodriguez. 1989. "Delay of gratification in children." *Science*, 244(4907): 933–938.

Noftle, Erik E, and Richard W Robins. 2007. "Personality predictors of academic outcomes: big five correlates of GPA and SAT scores." *Journal of personality and social psychology*, 93(1): 116.

O'Connor, Melissa C, and Sampo V Paunonen. 2007. "Big Five personality predictors of post-secondary academic performance." *Personality and Individual differences*, 43(5): 971–990.

Oreopoulos, Philip, and Uros Petronijevic. 2016. "Leveraging Technology and Behavioral Insights to Help College Students."

Poropat, Arthur E. 2009. "A meta-analysis of the five-factor model of personality and academic performance." *Psychological bulletin*, 135(2): 322.

Richardson, Michelle, Charles Abraham, and Rod Bond. 2012. "Psychological correlates of university students' academic performance: a systematic review and meta-analysis." *Psychological bulletin*, 138(2): 353.

Robbins, Steven B, Kristy Lauver, Huy Le, Daniel Davis, Ronelle Langley, and Aaron Carlstrom. 2004. "Do psychosocial and study skill factors predict college outcomes? A meta-analysis." *Psychological bulletin*, 130(2): 261.

Roberts, Brent W, Nathan R Kuncel, Rebecca Shiner, Avshalom Caspi, and Lewis R Goldberg. 2007. "The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes." *Perspectives on Psychological Science*, 2(4): 313–345.

Rothstein, Jesse M. 2004. "College performance predictions and the SAT." *Journal of Econometrics*, 121(1): 297–317.

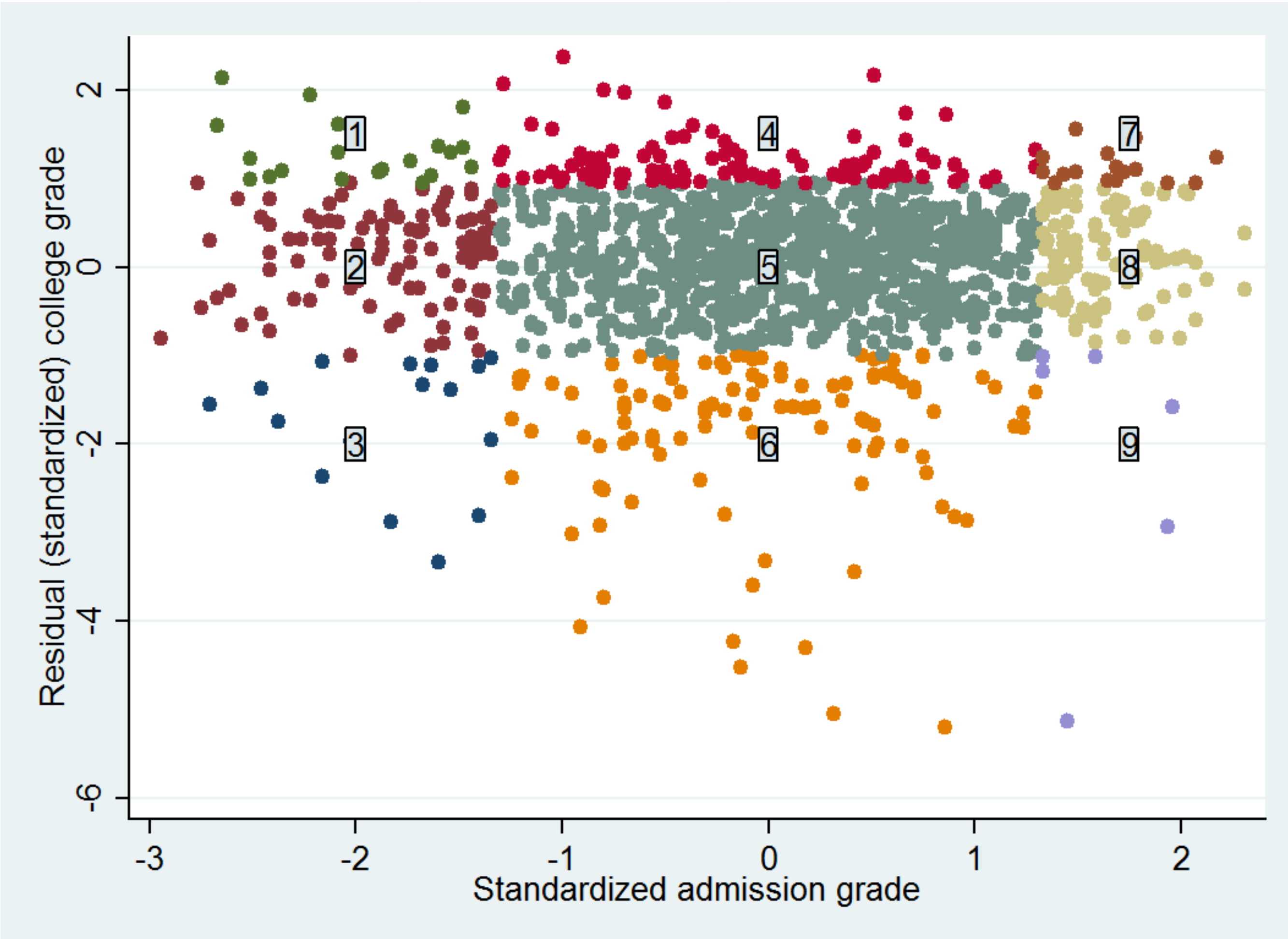
Scott-Clayton, Judith, Peter M Crosta, and Clive R Belfield. 2014. "Improving the Targeting of Treatment Evidence From College Remediation." *Educational Evaluation and Policy Analysis*, 36(3): 371–393.

Stephan, Jennifer L, Elisabeth Davis, Jim Lindsay, and Shazia Miller. 2015. "Who will succeed and who will struggle? Predicting early college success with Indiana's Student Information System." U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest REL2015-078.

Symonds, William C, Robert Schwartz, and Ronald F Ferguson. 2011. "Pathways to prosperity: Meeting the challenge of preparing young Americans." *Cambridge, MA: Pathways to Prosperity Project at Harvard Graduate School of Education.*

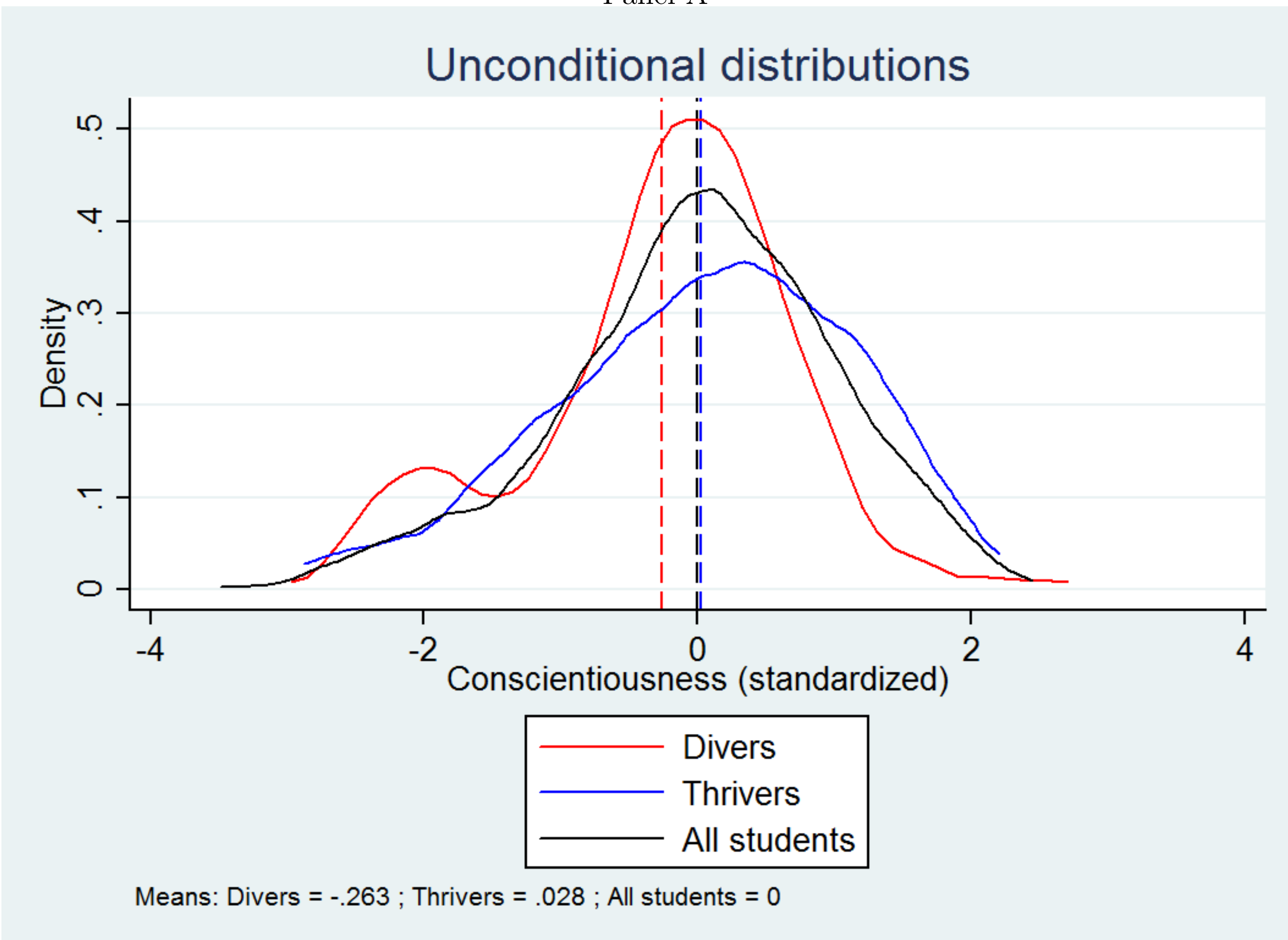
Willingham, Warren W. 1985. "Success in college." *New York: College Entrance Examination Board.*

Figure 1: Distribution of grade residuals



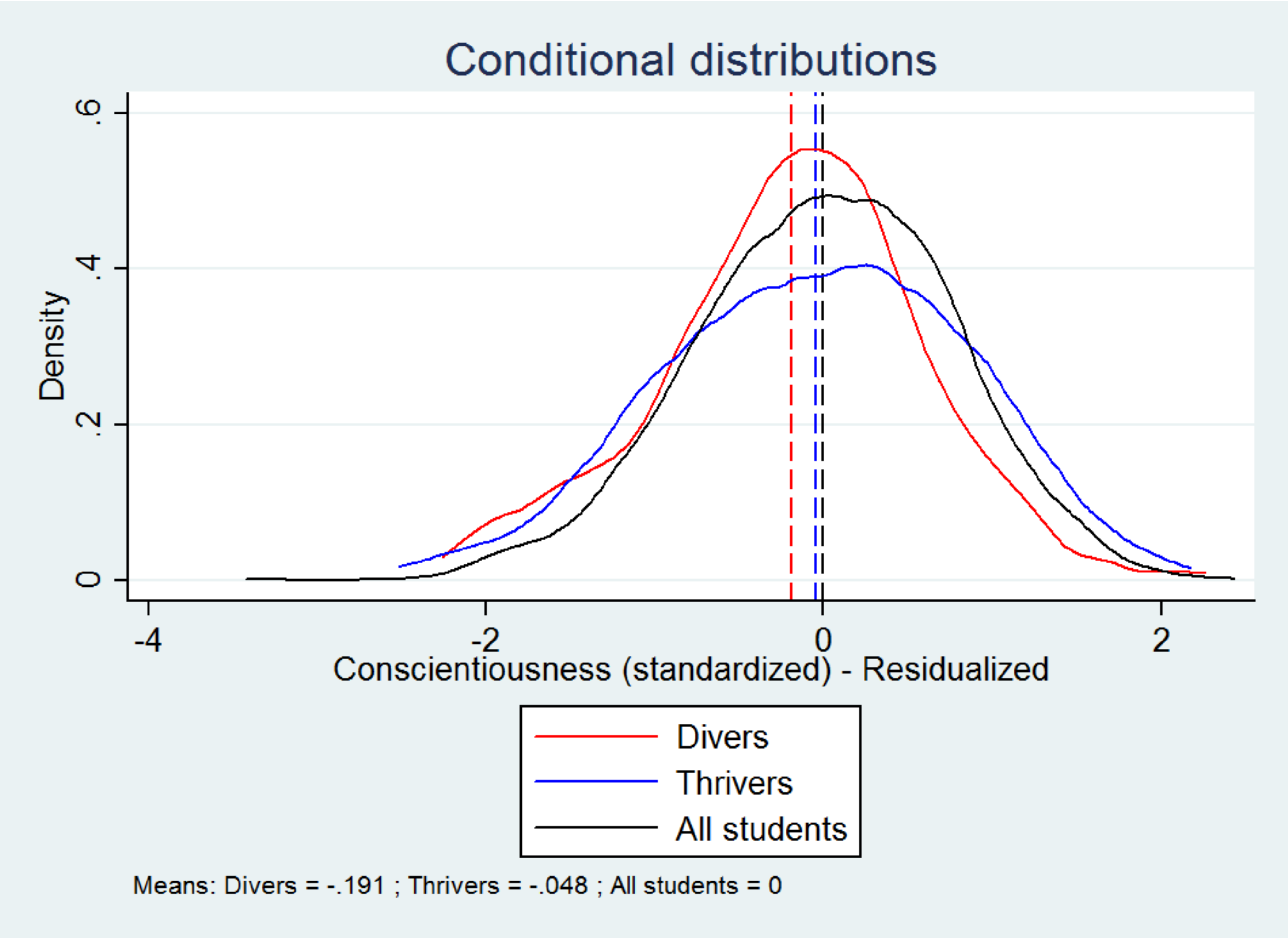
Colors indicate whether students are in the top or bottom decile of the distribution on each dimension. College grade residuals are obtained from specification (2), table 2. The sample is restricted to the personality sample.

Figure 2: Differences in distributions of conscientiousness
Panel A



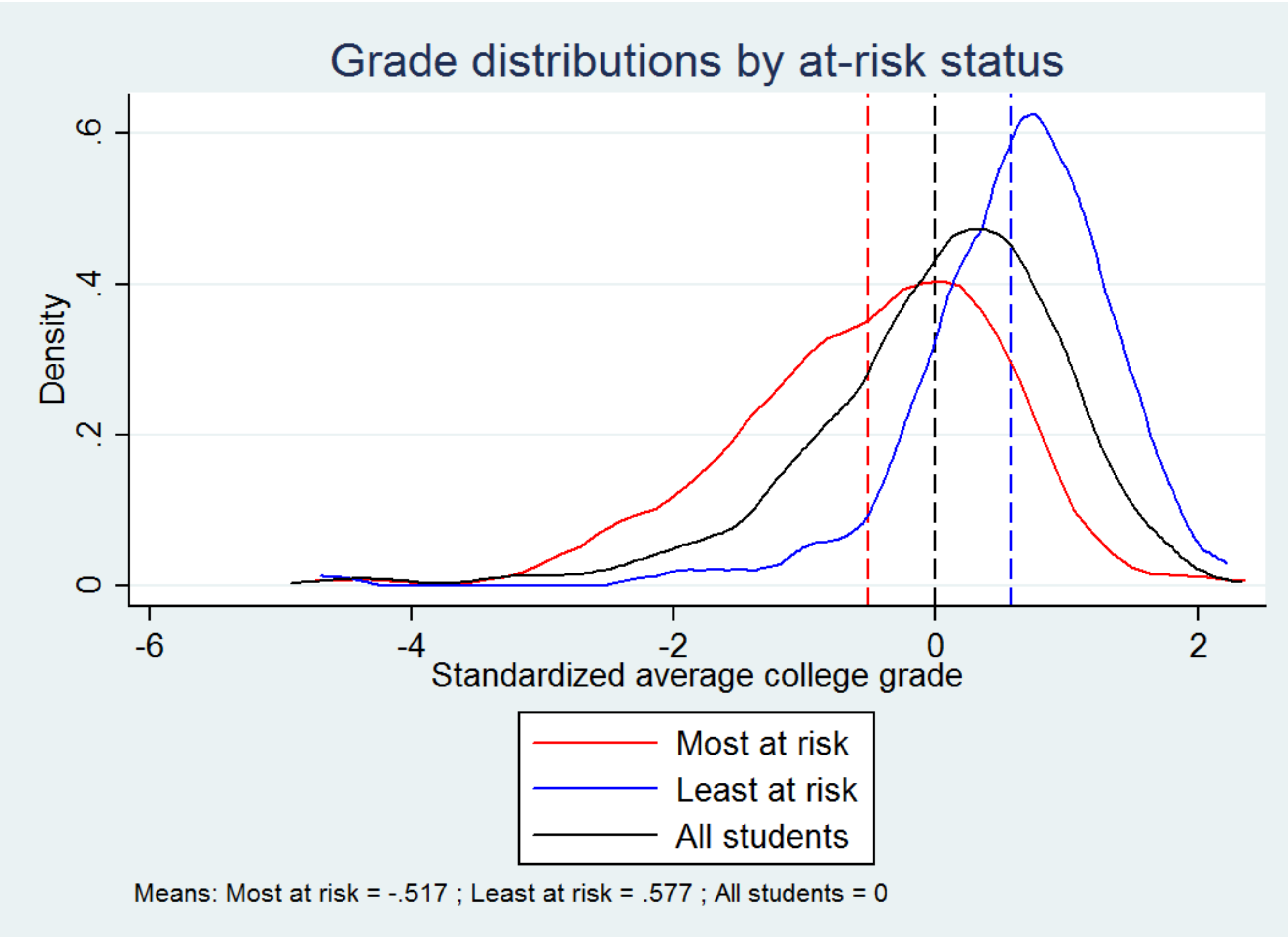
Conscientiousness is relative-scored and unadjusted. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Panel B



Conscientiousness is relative-scored and residualized from a regression on all other non-academic characteristics. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Figure 3: College grades by at-risk status



College grades are unadjusted. Most at-risk students are defined as students below the 10th percentile in the distribution of the seven-variable unweighted average of key characteristics. Least at-risk students rank above the 90th percentile. The full distribution corresponds to the personality sample.

Table 1: Summary statistics

	Mean	Standard deviation
Age at entry	18.07	[0.959]
Mother tongue: English	0.48	[0.500]
Citizenship: Canadian	0.52	[0.500]
Women	0.52	[0.500]
First-year student in 2015	0.82	[0.388]
International student	0.34	[0.473]
Economics is a required course	0.59	[0.491]
Living in Residence	0.30	[0.459]
Mother has BA or more	0.50	[0.500]
Father has BA or more	0.59	[0.491]
First-generation student	0.25	[0.430]
Hours expected to study	18.18	[10.816]
Hours expected to work for pay	7.46	[9.807]
Expects to get more than undergraduate degree	0.63	[0.482]
Expected college GPA	3.61	[0.434]
Day started the survey (relative to first day of class)	3.84	[5.257]
Admission grade	87.38	[5.121]
Average college grade	66.33	[13.467]
Observations	1,317	

Note: Sample is restricted to students in the personality sample whose admission grade is not missing, and who finished at least one university course in their first year. First-year and international student status, gender, parental education, study habits and expectations are self-reported. We infer that economics is a required course if a student intends to major in either Economics or Business. Age, mother tongue, citizenship and grades are from administrative records. The average college grade is calculated over all courses for which a valid grade is reported in the administrative file and weighted by number of credits.

Table 2: First-stage - Predicting performance using admission grades

	Dependent variable: Average first year college grade	
	(1)	(2)
Admission Grade	0.409***	0.433***
	[0.030]	[0.030]
Non-Domestic status		-0.273***
		[0.051]
Age at entry		-0.049*
		[0.027]
Observations	1,317	1,317
Adjusted R-squared	0.195	0.218

Notes: Both average college and admission grades are standardized to have mean zero and a standard deviation of one. All regressions include admission unit and cohort fixed-effects. Non-domestic status is one if a student either self-declared as international or has a citizenship other than Canadian. Standard errors are in brackets.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Differences between outliers and full distribution - Personality sample

	Bottom decile		Top decile		Difference between outliers	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
	Mean diff.	Mean diff.	Mean diff.	Mean diff.	(3) - (1)	(4) - (2)
	[s.e.]	[s.e.]	[s.e.]	[s.e.]	[p-val test (3)=(1)]	[p-val test (4)=(2)]
	(1)	(2)	(3)	(4)	(5)	(6)
Study hours per week (z-score)	-0.079	-0.019	0.224**	0.226***	0.303**	0.245**
	[0.087]	[0.083]	[0.087]	[0.083]	[0.014]	[0.037]
Sure about program of study (z-score)	-0.074	-0.085	-0.067	-0.08	0.007	0.005
	[0.087]	[0.082]	[0.087]	[0.082]	[0.954]	[0.966]
Think about future goals (z-score)	0.134	0.150**	-0.119	-0.095	-0.253**	-0.244**
	[0.087]	[0.075]	[0.087]	[0.075]	[0.040]	[0.022]
Identify with university (z-score)	0.067	0.04	-0.011	0.04	-0.078	0.001
	[0.087]	[0.080]	[0.087]	[0.080]	[0.529]	[0.995]
Transition has been challenging (z-score)	0.061	-0.024	-0.059	-0.091	-0.12	-0.067
	[0.087]	[0.078]	[0.087]	[0.079]	[0.330]	[0.548]
Cram for exams (z-score)	0.297***	0.209***	-0.043	-0.058	-0.34***	-0.268**
	[0.087]	[0.076]	[0.087]	[0.076]	[0.006]	[0.013]
Work hours per week (z-score)	0.216**	0.140*	0.049	0.076	-0.168	-0.064
	[0.087]	[0.083]	[0.087]	[0.084]	[0.173]	[0.588]
Expected GPA (z-score)	-0.097	-0.106	0.137	0.126	0.233*	0.232**
	[0.087]	[0.080]	[0.087]	[0.080]	[0.058]	[0.040]
Day started exercise (z-score)	0.288***	0.199**	-0.041	-0.03	-0.329***	-0.229*
	[0.087]	[0.083]	[0.087]	[0.083]	[0.007]	[0.051]
Expects more than undergraduate	0.003	-0.009	0.016	0.033	0.012	0.043
	[0.042]	[0.040]	[0.042]	[0.040]	[0.834]	[0.448]
Aggreableness (z-score)	-0.023	0.045	-0.03	-0.056	-0.007	-0.101
	[0.087]	[0.079]	[0.087]	[0.080]	[0.956]	[0.368]
Conscientiousness (z-score)	-0.263***	-0.191***	0.028	-0.048	0.292**	0.143
	[0.087]	[0.066]	[0.087]	[0.067]	[0.018]	[0.128]
Extraversion (z-score)	0.170*	0.1	-0.102	-0.038	-0.272**	-0.138
	[0.087]	[0.078]	[0.087]	[0.078]	[0.027]	[0.212]
Openness (z-score)	0.087	0.035	0.094	0.011	0.007	-0.024
	[0.087]	[0.077]	[0.087]	[0.077]	[0.953]	[0.829]
Emotional stability (z-score)	0.05	-0.014	0.024	-0.102	-0.026	-0.088
	[0.087]	[0.076]	[0.087]	[0.077]	[0.834]	[0.414]
Risk tolerance (z-score)	0.098	0.042	-0.231***	-0.109	-0.328***	-0.151
	[0.087]	[0.078]	[0.087]	[0.078]	[0.008]	[0.171]
Impatience (z-score)	0.199**	0.180**	-0.134	-0.112	-0.333***	-0.292**
	[0.087]	[0.085]	[0.087]	[0.085]	[0.007]	[0.016]
Procrastination (z-score)	0.102	0.067	0.043	-0.039	-0.059	-0.106
	[0.087]	[0.078]	[0.087]	[0.078]	[0.633]	[0.335]
Locus of Control (z-score)	0.112	0.094	-0.034	0.02	-0.146	-0.074
	[0.087]	[0.081]	[0.087]	[0.081]	[0.238]	[0.519]
Perseverance of effort (z-score)	-0.135	-0.089	-0.04	0.038	0.095	0.127
	[0.087]	[0.081]	[0.087]	[0.081]	[0.439]	[0.265]
Consisency of interest (z-score)	0.053	0.051	-0.132	-0.086	-0.185	-0.137
	[0.087]	[0.078]	[0.087]	[0.078]	[0.133]	[0.214]
Women	-0.105**	-0.082**	-0.094**	-0.090**	0.011	-0.008
	[0.043]	[0.041]	[0.044]	[0.041]	[0.860]	[0.888]
English mother tongue	0.013	0.017	0.009	0.041	-0.004	0.024
	[0.044]	[0.031]	[0.044]	[0.031]	[0.950]	[0.594]
Canadian citizenship	-0.017	-0.029	-0.036	-0.034	-0.019	-0.005
	[0.044]	[0.026]	[0.044]	[0.026]	[0.757]	[0.899]
International student	-0.005	-0.014	-0.01	-0.015	-0.005	-0.001
	[0.041]	[0.028]	[0.041]	[0.028]	[0.931]	[0.976]
Economics is required	0.021	0.011	-0.059	-0.022	-0.079	-0.033
	[0.043]	[0.041]	[0.043]	[0.041]	[0.191]	[0.570]
Mother has at least bachelor degree	-0.022	-0.049	-0.003	0.004	0.019	0.053
	[0.044]	[0.034]	[0.044]	[0.034]	[0.759]	[0.270]
Father has at least bachelor degree	0.043	0.033	0.018	0.034	-0.026	0.001
	[0.043]	[0.029]	[0.043]	[0.029]	[0.672]	[0.983]
First-generation student	-0.026	-0.011	0.03	0.041	0.055	0.052
	[0.037]	[0.026]	[0.038]	[0.026]	[0.299]	[0.151]

Notes: All non-z-score predictors are binary. In columns (1) through (4), coefficients represent the difference in means between outlier groups and the full sample. For conditional differences (columns (2) and (4)), each characteristic is first regressed on the set of other characteristics reported in this table. Big Five traits are relative-scored. Likert-scale Big Five traits are used as controls in lieu of relative-scored traits in the residualization process for columns conditional differences.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Differences between outliers and full distribution - Text sample

	Bottom decile		Top decile		Difference between outliers	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
	Mean diff.	Mean diff.	Mean diff.	Mean diff.	(3) - (1)	(4) - (2)
	[s.e.]	[s.e.]	[s.e.]	[s.e.]	[p-val test (3)=(1)]	[p-val test (4)=(2)]
	(1)	(2)	(3)	(4)	(5)	(6)
Total number of words used (z-score)	-0.263***	-0.134**	0.046	0.130**	0.309***	0.264***
	[0.065]	[0.061]	[0.065]	[0.061]	[0.001]	[0.002]
Proportion spelled correctly (z-score)	-0.135**	-0.058	0.046	0.077	0.170*	0.135
	[0.066]	[0.064]	[0.065]	[0.064]	[0.066]	[0.136]
Time taken on written questions (z-score)	-0.043	-0.066	0.098	0.066	0.141	0.132
	[0.065]	[0.062]	[0.065]	[0.062]	[0.126]	[0.135]

Notes: In columns (1) through (4), coefficients represent the difference in means between outlier groups and the full sample. For conditional differences (columns (2) and (4)), each characteristic is first regressed on the set of controls (variables from survey and administrative data.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Words used more frequently by outliers

Question	Top decile words	Bottom decile words
Name two goals	build	rich business own actuary
Qualities admire in self	discipline specific word practice responsibility smart confident game	cause communicate receive friendly trust
Your future self	human meet people deal god trustworthy computer whole provide famous love mature determine helpful wise	tough man book father moment rich
Qualities admire in others	weakness avoid challenge overcome read mistake creativity word people Steve area general initiative understand	power concentration waste

Table 6: Predictive properties of summary measures of non-academic characteristics

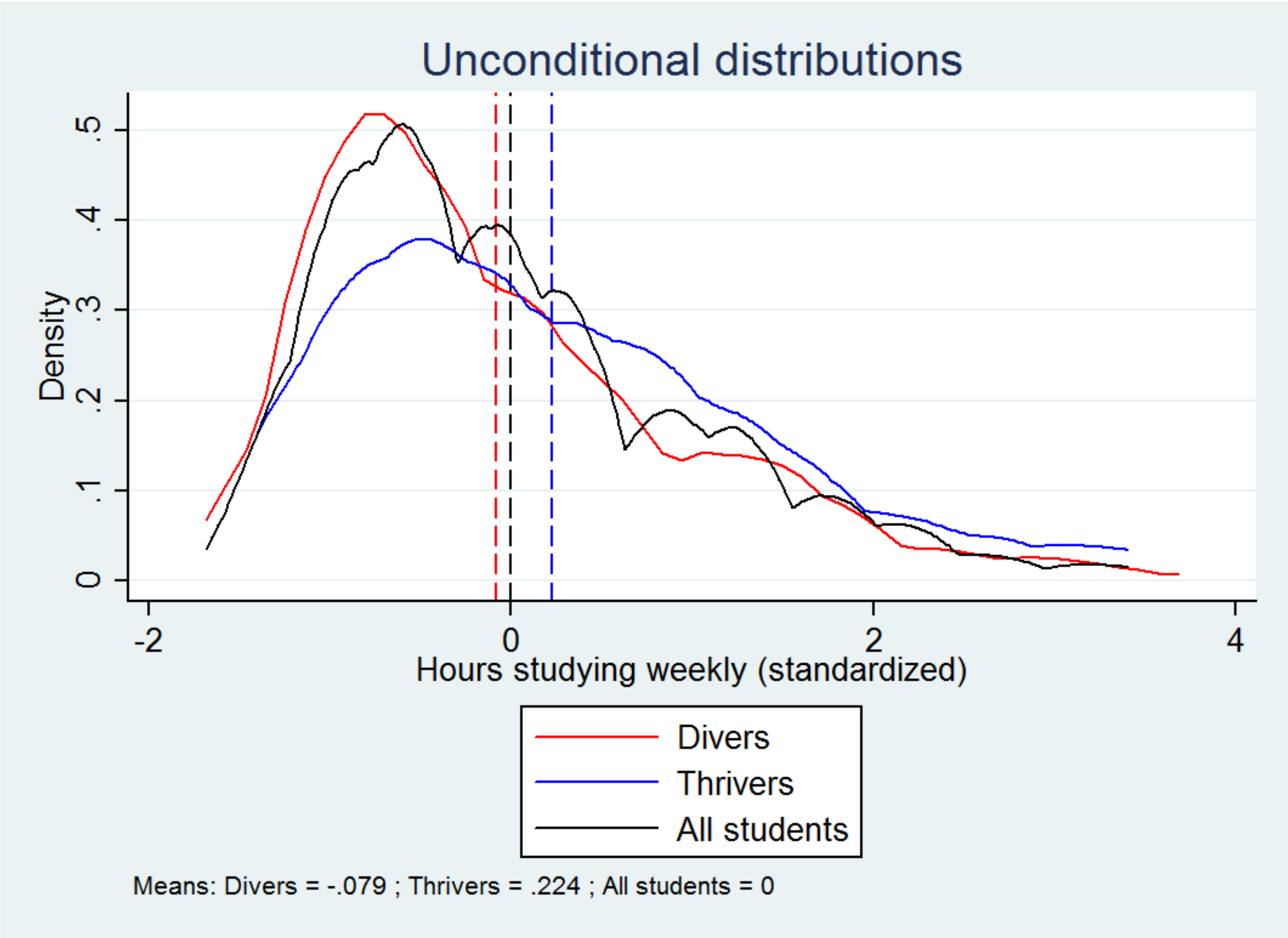
	Dependent variable : Standardized first-year college grades								
Summary measure	-	-	Unweighted average	Principal component	LARS on outliers	LARS on outliers	LARS on outliers	LARS on outliers	LARS on full sample
Predictors included	-	-	7 best	7 best	7 best	7 best + polynomials	7 best + polynomials + interactions	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Separate predictive power									
Admission grade	0.433***	0.446***							
	[0.030]	[0.030]							
Admission grade^2		0.046**							
		[0.019]							
Summary measure			0.653***	0.249***	0.560***	0.410***	0.701***	0.499***	1.355***
			[0.061]	[0.027]	[0.054]	[0.045]	[0.076]	[0.045]	[0.110]
Observations	1,317	1,317	1,317	1,317	1,317	1,317	1,317	1,317	1,317
Adjusted R^2	0.218	0.221	0.163	0.145	0.160	0.145	0.146	0.169	0.184
Panel B: Incremental predictive power									
Admission grade	0.433***	0.446***	0.387***	0.401***	0.392***	0.408***	0.405***	0.390***	0.370***
	[0.030]	[0.030]	[0.030]	[0.030]	[0.030]	[0.030]	[0.030]	[0.030]	[0.030]
Admission grade^2		0.046**	0.037**	0.038**	0.038**	0.042**	0.041**	0.037**	0.035*
		[0.019]	[0.018]	[0.019]	[0.018]	[0.019]	[0.019]	[0.018]	[0.018]
Summary measure			0.466***	0.171***	0.406***	0.308***	0.515***	0.381***	0.992***
			[0.059]	[0.026]	[0.052]	[0.043]	[0.072]	[0.043]	[0.109]
Observations	1,317	1,317	1,317	1,317	1,317	1,317	1,317	1,317	1,317
Adjusted R^2	0.218	0.221	0.255	0.245	0.255	0.250	0.249	0.265	0.267

Note: All regressions include campus and cohort fixed-effects, as well as non-domestic statusand age at entry. Standard errors are in brackets. In column (6), quadratic and cubic terms for each of the 7 best predictors are used in the LARS algorithm. In column (7) all paiwise interactions between the 7 best predictors are further added. In columns (8) and (9), the set of potential predictors used in the algorithm is all variables encompass all variables listed in table 3.

*** p<0.01, ** p<0.05, * p<0.1

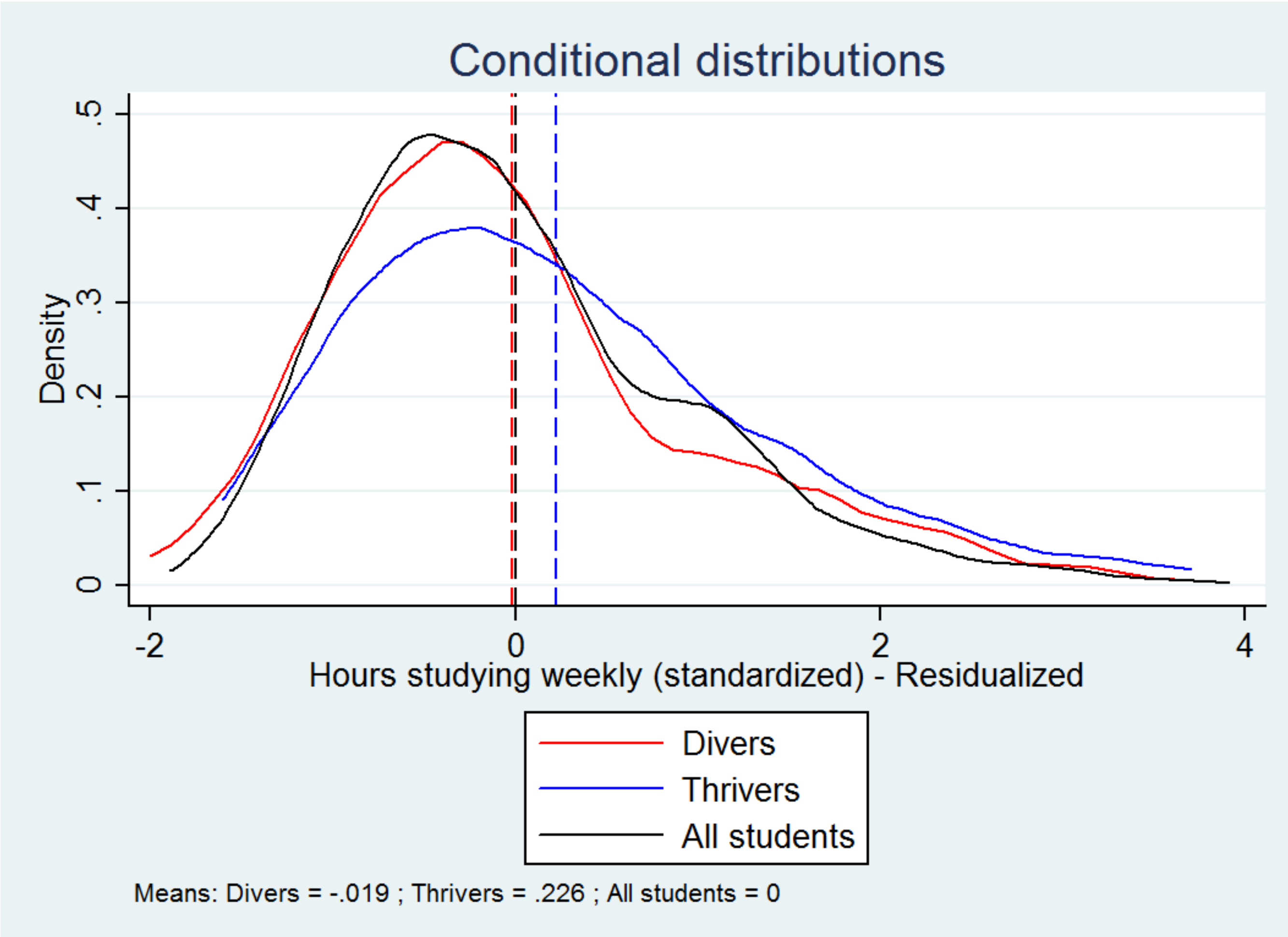
Appendix

Figure A1: Differences in distributions of study hours
Panel A



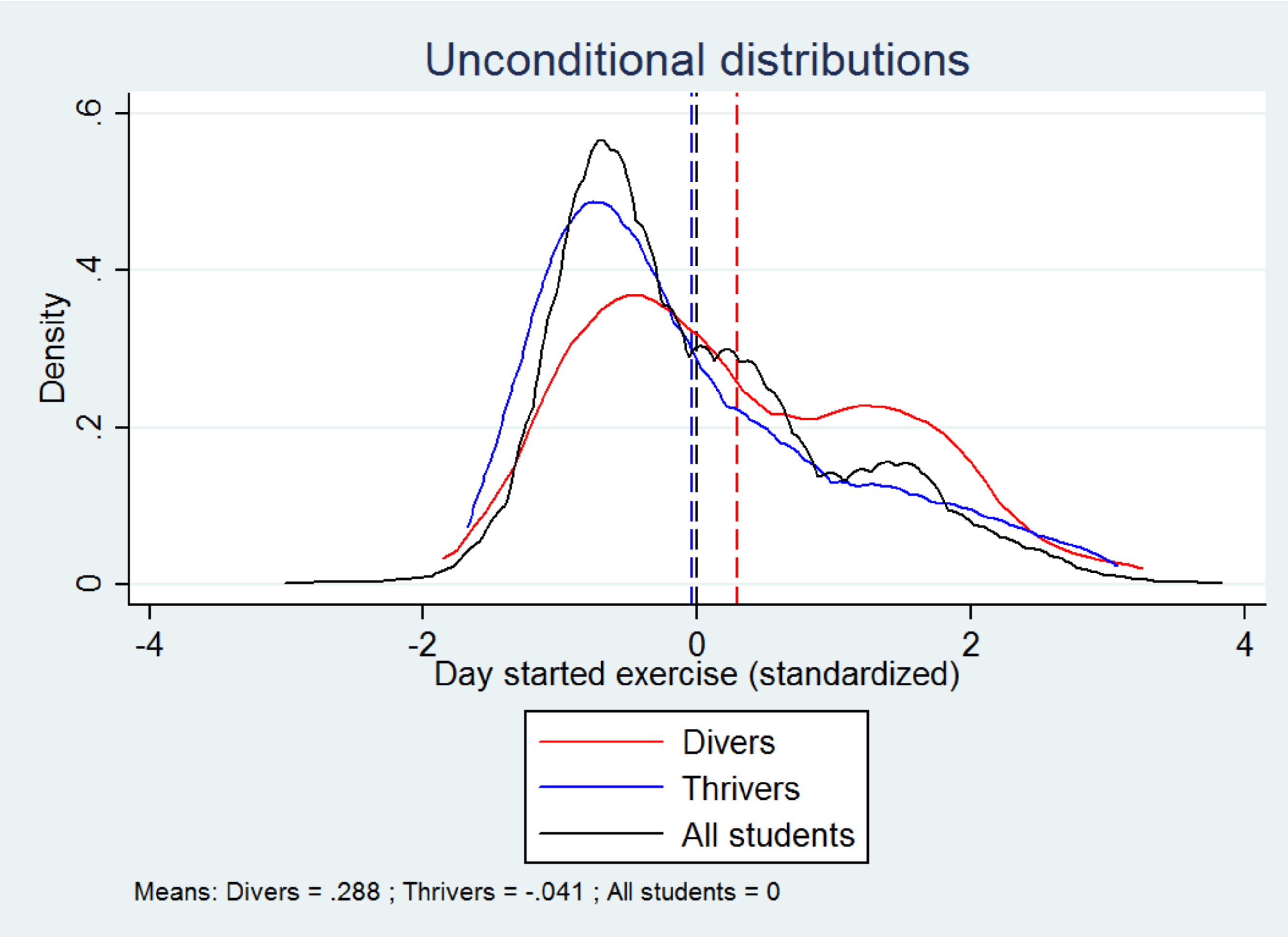
The number of hours studying weekly is standardized with mean zero and unit variance. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Panel B



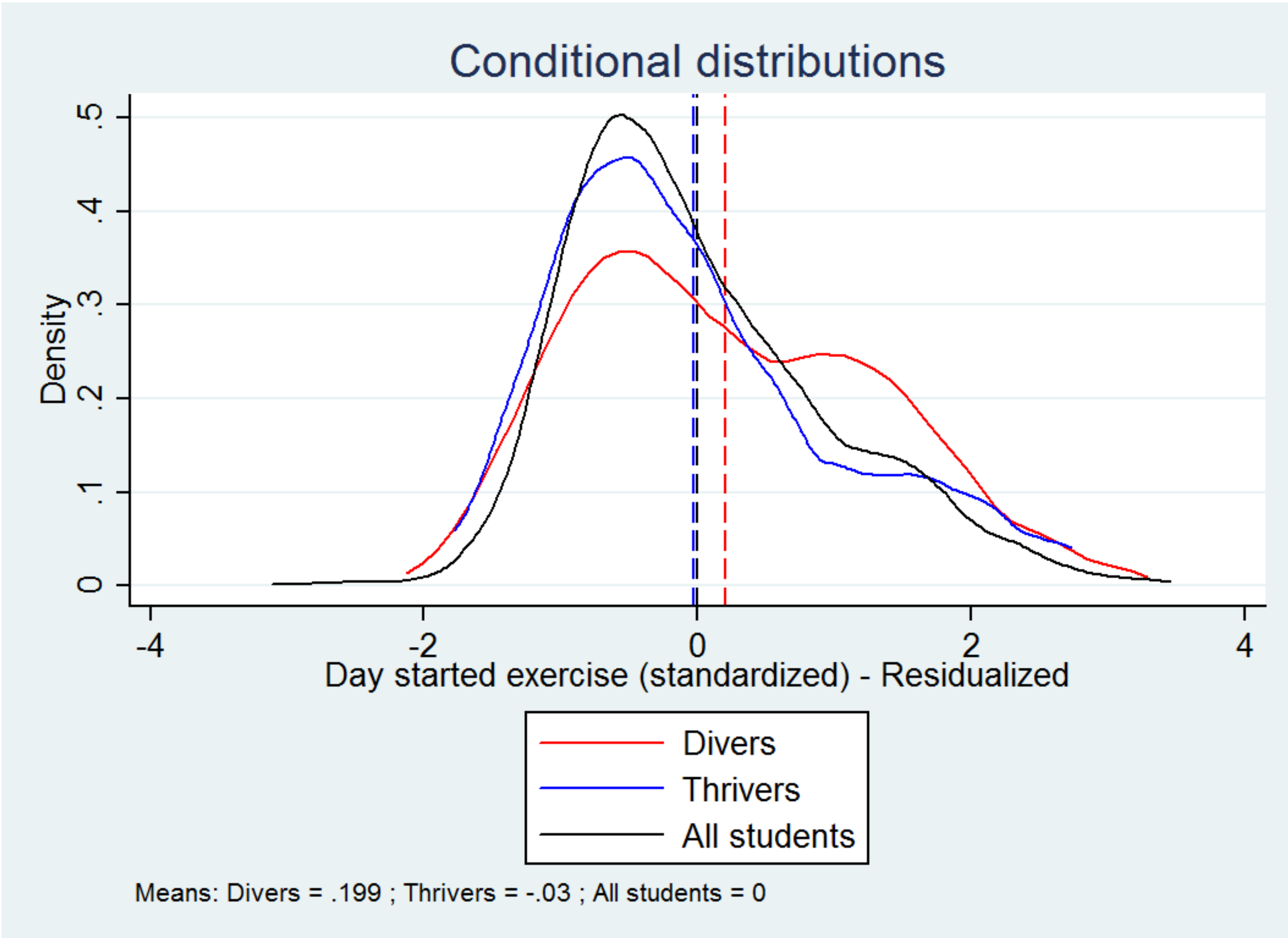
The number of hours studying weekly is standardized with mean zero and unit variance and residualized from a regression on all other non-academic characteristics. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Figure A2: Differences in distributions of time started exercise
Panel A



The number of days relative to first day of class is standardized with mean zero and unit variance. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Panel B



The number of days relative to first day of class is standardized with mean zero and unit variance and residualized from a regression on all other non-academic characteristics. Divers are defined as students with residual college grade below the 10th percentile. Thrivers have residual college grades above the 90th percentile. The full distribution corresponds to the personality sample.

Table A1: Differences between outliers and full distribution - Quintiles

	Bottom quintile		Top quintile		Difference between outliers	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
	Mean diff.	Mean diff.	Mean diff.	Mean diff.	(3) - (1)	(4) - (2)
	[s.e.]	[s.e.]	[s.e.]	[s.e.]	[p-val test (3)=(1)]	[p-val test (4)=(2)]
	(1)	(2)	(3)	(4)	(5)	(6)
Study hours per week (z-score)	-0.003	0.043	0.172***	0.144**	0.175**	0.101
	[0.061]	[0.059]	[0.062]	[0.059]	[0.044]	[0.225]
Sure about program of study (z-score)	-0.059	-0.033	-0.002	-0.021	0.057	0.012
	[0.062]	[0.058]	[0.062]	[0.058]	[0.512]	[0.888]
Think about future goals (z-score)	0.000	0.036	-0.092	-0.081	-0.092	-0.117
	[0.062]	[0.053]	[0.062]	[0.053]	[0.290]	[0.120]
Identify with university (z-score)	-0.001	-0.013	-0.016	0.002	-0.014	0.015
	[0.062]	[0.056]	[0.062]	[0.057]	[0.868]	[0.854]
Transition has been challenging (z-score)	0.129**	0.070	-0.067	-0.084	-0.197**	-0.154**
	[0.061]	[0.055]	[0.062]	[0.055]	[0.024]	[0.049]
Cram for exams (z-score)	0.211***	0.119**	-0.150**	-0.122**	-0.361****	-0.242***
	[0.061]	[0.054]	[0.061]	[0.054]	[0.000]	[0.002]
Work hours per week (z-score)	0.139**	0.084	0.005	0.024	-0.134	-0.06
	[0.061]	[0.059]	[0.062]	[0.059]	[0.123]	[0.475]
Expected GPA (z-score)	-0.124**	-0.115**	0.152**	0.115**	0.276***	0.23***
	[0.061]	[0.056]	[0.061]	[0.056]	[0.002]	[0.004]
Day started exercise (z-score)	0.230***	0.165***	-0.084	-0.032	-0.314***	-0.198**
	[0.061]	[0.058]	[0.061]	[0.059]	[0.000]	[0.017]
Expects more than undergraduate (binary)	-0.001	-0.012	0.002	0.021	0.002	0.033
	[0.030]	[0.028]	[0.030]	[0.028]	[0.954]	[0.403]
Aggreableness (z-score)	0.042	0.062	-0.000	-0.048	-0.042	-0.11
	[0.062]	[0.056]	[0.062]	[0.056]	[0.626]	[0.166]
Conscientiousness (z-score)	-0.206***	-0.128***	0.081	-0.015	0.287***	0.114*
	[0.061]	[0.047]	[0.061]	[0.047]	[0.001]	[0.087]
Extraversion (z-score)	0.127**	0.101*	-0.115*	-0.071	-0.241***	-0.172**
	[0.061]	[0.055]	[0.062]	[0.055]	[0.006]	[0.028]
Openness (z-score)	-0.009	-0.014	-0.007	-0.031	0.002	-0.017
	[0.062]	[0.055]	[0.062]	[0.055]	[0.977]	[0.822]
Emotional stability (z-score)	0.059	0.046	0.043	-0.034	-0.016	-0.08
	[0.062]	[0.054]	[0.062]	[0.054]	[0.852]	[0.294]
Risk tolerance (z-score)	-0.001	-0.046	-0.133**	-0.020	-0.131	0.026
	[0.061]	[0.055]	[0.062]	[0.055]	[0.131]	[0.736]
Impatience (z-score)	0.155**	0.157***	-0.124**	-0.110*	-0.279***	-0.267***
	[0.061]	[0.060]	[0.061]	[0.060]	[0.001]	[0.002]
Procrastination (z-score)	0.165***	0.158***	-0.034	-0.069	-0.199**	-0.227***
	[0.061]	[0.055]	[0.062]	[0.055]	[0.022]	[0.003]
Locus of Control (z-score)	0.010	0.011	-0.057	-0.001	-0.067	-0.012
	[0.062]	[0.057]	[0.062]	[0.057]	[0.443]	[0.879]
Perseverance of effort (z-score)	-0.082	-0.038	0.026	0.075	0.108	0.113
	[0.062]	[0.057]	[0.062]	[0.057]	[0.215]	[0.163]
Consisency of interest (z-score)	0.069	0.109**	-0.091	-0.087	-0.16*	-0.197**
	[0.061]	[0.055]	[0.062]	[0.055]	[0.066]	[0.011]
Women	-0.037	-0.021	-0.012	-0.024	0.025	-0.003
	[0.031]	[0.029]	[0.031]	[0.029]	[0.571]	[0.942]
English mother tongue	-0.009	0.002	-0.034	0.018	-0.025	0.016
	[0.031]	[0.022]	[0.031]	[0.022]	[0.569]	[0.616]
Canadian citizenship	-0.025	-0.016	-0.068**	-0.030	-0.044	-0.015
	[0.031]	[0.018]	[0.031]	[0.018]	[0.314]	[0.575]
International student (survey)	0.014	-0.003	0.031	-0.006	0.017	-0.003
	[0.029]	[0.020]	[0.029]	[0.020]	[0.688]	[0.917]
Economics is required	0.009	0.010	-0.023	-0.003	-0.032	-0.013
	[0.030]	[0.029]	[0.030]	[0.029]	[0.456]	[0.748]
Mother has at least bachelor degree	0.008	-0.027	0.006	0.009	-0.002	0.037
	[0.031]	[0.024]	[0.031]	[0.024]	[0.966]	[0.283]
Father has at least bachelor degree	0.051*	0.013	0.015	0.022	-0.036	0.009
	[0.030]	[0.021]	[0.030]	[0.021]	[0.406]	[0.762]
First-generation student	-0.052**	-0.029	0.021	0.028	0.073*	0.057**
	[0.026]	[0.018]	[0.027]	[0.018]	[0.052]	[0.026]

Notes: All non-z-score predictors are binary. In columns (1) through (4), coefficients represent the difference in means between outlier groups and the full sample. For conditional differences (columns (2) and (4)), each characteristic is first regressed on the set of other characteristics reported in this table. Big Five traits are relative-scored. Likert-scale Big Five traits are used as controls in lieu of relative-scored traits in the residualization process for columns conditional differences.
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Differences between outliers and full distribution - Groups 4,5,6 only

	Bottom decile		Top decile		Difference between outliers	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
	Mean diff.	Mean diff.	Mean diff.	Mean diff.	(3) - (1)	(4) - (2)
	[s.e.]	[s.e.]	[s.e.]	[s.e.]	[p-val test (3)=(1)]	[p-val test (4)=(2)]
	(1)	(2)	(3)	(4)	(5)	(6)
Study hours per week (z-score)	-0.069	-0.034	0.265**	0.241**	0.334**	0.275**
	[0.096]	[0.092]	[0.103]	[0.099]	[0.018]	[0.041]
Sure about program of study (z-score)	-0.06	-0.045	-0.074	-0.105	-0.014	-0.06
	[0.095]	[0.088]	[0.102]	[0.095]	[0.919]	[0.640]
Think about future goals (z-score)	0.098	0.152*	-0.15	-0.13	-0.248*	-0.282**
	[0.095]	[0.081]	[0.102]	[0.088]	[0.075]	[0.018]
Identify with university (z-score)	0.025	-0.007	-0.061	-0.012	-0.086	-0.005
	[0.095]	[0.088]	[0.103]	[0.094]	[0.537]	[0.970]
Transition has been challenging (z-score)	0.102	-0.002	-0.057	-0.093	-0.159	-0.091
	[0.095]	[0.086]	[0.102]	[0.092]	[0.257]	[0.470]
Cram for exams (z-score)	0.302***	0.195**	-0.083	-0.094	-0.385***	-0.289***
	[0.094]	[0.083]	[0.101]	[0.089]	[0.005]	[0.018]
Work hours per week (z-score)	0.249***	0.169*	0.009	0.063	-0.24*	-0.106
	[0.096]	[0.091]	[0.103]	[0.098]	[0.088]	[0.426]
Expected GPA (z-score)	-0.099	-0.106	0.179*	0.193**	0.278**	0.299**
	[0.096]	[0.089]	[0.103]	[0.096]	[0.050]	[0.022]
Day started exercise (z-score)	0.307***	0.196**	-0.084	-0.091	-0.391***	-0.288**
	[0.093]	[0.089]	[0.100]	[0.096]	[0.004]	[0.028]
Expects more than undergraduate (binary)	-0.024	-0.018	0.019	0.051	0.044	0.069
	[0.046]	[0.044]	[0.050]	[0.047]	[0.521]	[0.279]
Aggreableness (z-score)	-0.05	0.015	-0.021	-0.037	0.029	-0.052
	[0.097]	[0.088]	[0.104]	[0.095]	[0.836]	[0.688]
Conscientiousness (z-score)	-0.254***	-0.153**	-0.004	-0.116	0.251*	0.037
	[0.096]	[0.073]	[0.103]	[0.079]	[0.076]	[0.731]
Extraversion (z-score)	0.185*	0.119	-0.111	-0.051	-0.297**	-0.17
	[0.096]	[0.086]	[0.103]	[0.092]	[0.035]	[0.175]
Openness (z-score)	0.031	-0.032	0.112	0.021	0.081	0.052
	[0.096]	[0.085]	[0.103]	[0.092]	[0.567]	[0.676]
Emotional stability (z-score)	0.099	0.01	0.046	-0.118	-0.053	-0.127
	[0.097]	[0.085]	[0.105]	[0.091]	[0.712]	[0.308]
Risk tolerance (z-score)	0.051	0.002	-0.251**	-0.088	-0.303**	-0.09
	[0.094]	[0.084]	[0.101]	[0.090]	[0.029]	[0.465]
Impatience (z-score)	0.168*	0.168*	-0.200**	-0.161	-0.368***	-0.329**
	[0.094]	[0.092]	[0.101]	[0.099]	[0.008]	[0.015]
Procrastination (z-score)	0.081	0.042	0.025	-0.041	-0.056	-0.083
	[0.095]	[0.085]	[0.102]	[0.091]	[0.688]	[0.504]
Locus of Control (z-score)	0.107	0.086	0.015	0.063	-0.092	-0.023
	[0.096]	[0.089]	[0.103]	[0.095]	[0.513]	[0.858]
Perseverance of effort (z-score)	-0.210**	-0.168*	0.016	0.056	0.226*	0.224*
	[0.093]	[0.086]	[0.101]	[0.093]	[0.100]	[0.077]
Consisency of interest (z-score)	0.053	0.072	-0.125	-0.061	-0.178	-0.132
	[0.094]	[0.083]	[0.101]	[0.089]	[0.198]	[0.277]
Women	-0.113**	-0.086*	-0.111**	-0.110**	0.001	-0.024
	[0.047]	[0.045]	[0.051]	[0.049]	[0.984]	[0.713]
English mother tongue	0.003	0.008	0.026	0.036	0.023	0.028
	[0.048]	[0.034]	[0.051]	[0.037]	[0.738]	[0.579]
Canadian citizenship	-0.026	-0.021	-0.033	-0.036	-0.007	-0.015
	[0.048]	[0.029]	[0.051]	[0.031]	[0.924]	[0.717]
International student (survey)	0.008	-0.007	-0.033	-0.022	-0.04	-0.015
	[0.045]	[0.030]	[0.048]	[0.033]	[0.541]	[0.741]
Economics is required	0.007	-0.01	-0.067	-0.036	-0.074	-0.027
	[0.047]	[0.045]	[0.050]	[0.048]	[0.282]	[0.687]
Mother has at least bachelor degree	-0.036	-0.080**	-0.015	0.009	0.021	0.089
	[0.048]	[0.037]	[0.051]	[0.040]	[0.769]	[0.105]
Father has at least bachelor degree	0.080*	0.045	-0.004	0.029	-0.083	-0.016
	[0.047]	[0.032]	[0.050]	[0.034]	[0.224]	[0.728]
First-generation student	-0.054	-0.024	0.06	0.057*	0.114*	0.081*
	[0.040]	[0.028]	[0.044]	[0.030]	[0.055]	[0.045]

Notes: Sample restricted to students with admission grades between the 10th and 90th percentile. All non-z-score predictors are binary. In columns (1) through (4), coefficients represent the difference in means between outlier groups and the full sample. For conditional differences (columns (2) and (4)), each characteristic is first regressed on the set of other characteristics reported in this table. Big Five traits are relative-scored. Likert-scale Big Five traits are used as controls in lieu of relative-scored traits in the residualization process for columns conditional differences.
*** p<0.01, ** p<0.05, * p<0.1

Table A3: Differences between outliers and full distribution - Grades not adjusted for high school performance

	Bottom decile		Top decile		Difference between outliers	
	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
	Mean diff.	Mean diff.	Mean diff.	Mean diff.	(3) - (1)	(4) - (2)
	[s.e.]	[s.e.]	[s.e.]	[s.e.]	[p-val test (3)=(1)]	[p-val test (4)=(2)]
	(1)	(2)	(3)	(4)	(5)	(6)
Study hours per week (z-score)	-0.123	-0.035	0.169*	0.124	0.291**	0.159
	[0.087]	[0.083]	[0.087]	[0.083]	[0.018]	[0.176]
Sure about program of study (z-score)	-0.145*	-0.125	0.01	0.019	0.155	0.145
	[0.087]	[0.082]	[0.087]	[0.082]	[0.209]	[0.211]
Think about future goals (z-score)	0.091	0.124*	-0.175**	-0.158**	-0.266**	-0.282***
	[0.087]	[0.075]	[0.087]	[0.075]	[0.031]	[0.008]
Identify with university (z-score)	-0.009	-0.022	-0.169*	-0.109	-0.16	-0.088
	[0.087]	[0.080]	[0.087]	[0.080]	[0.194]	[0.437]
Transition has been challenging (z-score)	0.097	0.032	-0.142	-0.133*	-0.239*	-0.165
	[0.087]	[0.078]	[0.087]	[0.079]	[0.052]	[0.137]
Cram for exams (z-score)	0.248***	0.127*	-0.227***	-0.188**	-0.475***	-0.315***
	[0.087]	[0.076]	[0.087]	[0.076]	[0.000]	[0.004]
Work hours per week (z-score)	0.172**	0.091	-0.011	0.043	-0.183	-0.048
	[0.087]	[0.083]	[0.087]	[0.084]	[0.138]	[0.685]
Expected GPA (z-score)	-0.160*	-0.129	0.281***	0.266***	0.441***	0.394***
	[0.087]	[0.079]	[0.087]	[0.080]	[0.000]	[0.000]
Day started exercise (z-score)	0.412***	0.298***	-0.118	-0.08	-0.53***	-0.378***
	[0.086]	[0.082]	[0.087]	[0.083]	[0.000]	[0.001]
Expects more than undergraduate (binary)	-0.012	-0.026	0.031	0.054	0.043	0.08
	[0.042]	[0.040]	[0.042]	[0.040]	[0.471]	[0.154]
Aggreableness (z-score)	-0.07	0.02	-0.089	-0.155*	-0.019	-0.175
	[0.087]	[0.079]	[0.087]	[0.079]	[0.879]	[0.118]
Conscientiousness (z-score)	-0.275***	-0.132**	0.327***	0.147**	0.602***	0.279***
	[0.086]	[0.066]	[0.087]	[0.067]	[0.000]	[0.003]
Extraversion (z-score)	0.222**	0.134*	-0.322***	-0.226***	-0.544***	-0.36***
	[0.086]	[0.078]	[0.087]	[0.078]	[0.000]	[0.001]
Openness (z-score)	0.116	0.077	0.192**	0.140*	0.076	0.064
	[0.087]	[0.077]	[0.087]	[0.077]	[0.536]	[0.560]
Emotional stability (z-score)	0.023	-0.001	-0.1	-0.189**	-0.123	-0.188*
	[0.087]	[0.076]	[0.087]	[0.076]	[0.318]	[0.082]
Risk tolerance (z-score)	0.145*	0.034	-0.375***	-0.230***	-0.52***	-0.264**
	[0.086]	[0.077]	[0.087]	[0.078]	[0.000]	[0.016]
Impatience (z-score)	0.251***	0.223***	-0.150*	-0.118	-0.401***	-0.342***
	[0.087]	[0.085]	[0.087]	[0.085]	[0.001]	[0.005]
Procrastination (z-score)	0.163*	0.142*	-0.018	-0.045	-0.181	-0.186*
	[0.087]	[0.078]	[0.087]	[0.078]	[0.143]	[0.091]
Locus of Control (z-score)	0.07	0.036	0.023	0.094	-0.047	0.057
	[0.087]	[0.081]	[0.087]	[0.081]	[0.705]	[0.619]
Perseverance of effort (z-score)	-0.093	-0.056	0.01	0.071	0.103	0.127
	[0.087]	[0.081]	[0.087]	[0.081]	[0.404]	[0.268]
Consisency of interest (z-score)	0.068	0.081	-0.037	-0.012	-0.105	-0.093
	[0.087]	[0.078]	[0.087]	[0.078]	[0.394]	[0.397]
Women	-0.128***	-0.104**	-0.025	-0.045	0.102*	0.059
	[0.043]	[0.041]	[0.044]	[0.041]	[0.096]	[0.310]
English mother tongue	0.044	0.022	-0.044	-0.019	-0.088	-0.041
	[0.043]	[0.031]	[0.044]	[0.031]	[0.155]	[0.359]
Canadian citizenship	0.028	0	-0.021	0.029	-0.049	0.029
	[0.044]	[0.026]	[0.044]	[0.026]	[0.424]	[0.433]
International student (survey)	-0.005	0.018	0.036	0.025	0.041	0.007
	[0.041]	[0.028]	[0.041]	[0.028]	[0.486]	[0.868]
Economics is required	0.013	0.001	-0.051	-0.022	-0.064	-0.024
	[0.043]	[0.041]	[0.043]	[0.041]	[0.291]	[0.684]
Mother has at least bachelor degree	-0.037	-0.043	-0.026	0	0.011	0.043
	[0.044]	[0.034]	[0.044]	[0.034]	[0.856]	[0.366]
Father has at least bachelor degree	0.021	0.032	-0.013	0.019	-0.033	-0.014
	[0.043]	[0.029]	[0.043]	[0.029]	[0.581]	[0.744]
First-generation student	-0.003	0	0.045	0.036	0.048	0.036
	[0.037]	[0.026]	[0.038]	[0.026]	[0.369]	[0.313]

Notes: All non-z-score predictors are binary. Outliers are top and bottom deciles of the adjusted college grades distribution. College grades are adjusted for cohort and campus fixed effects, as well as age and non-domestic status. In columns (1) through (4), coefficients represent the difference in means between outlier groups and the full sample. For conditional differences (columns (2) and (4)), each characteristic is first regressed on the set of other characteristics reported in this table. Likert-scale Big Five traits are used as controls in lieu of relative-scored traits in the residualization process.
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Proportion of 'at-risk' students in bottom tail of distribution of college grades

	Distribution of unadjusted first-year college grades				
	Bottom 10%	Bottom 20%	Bottom 30%	Bottom 40%	Bottom 50%
Most at-risk:					
(a) Bottom decile of summary measure	23%	40%	52%	61%	71%
(b) Bottom decile of admission grades	25%	43%	60%	76%	87%
(c) Bottom decile in both metrics	38%	55%	76%	86%	100%
Least at-risk:					
(d) Top decile of summary measure	3%	6%	8%	14%	24%
(e) Top decile of admission grades	2%	2%	5%	9%	15%
(f) Top decile in both metrics	0%	0%	0%	0%	7%

Note: By construction 10% of our sample is defined as most/least at-risk in rows (a), (b), (d) and (e). Only 2.2% of students in our sample satisfy the 'at-risk' criterion in row (c), and 2.3% satisfy the criterion in (f).

Table A5: List two goals

With these thoughts in mind, what would you say are your 2 most inspiring goals?		
	Top decile words	**Bottom decile words**
Significant at 99%	build	rich business own
Significant at 95%		actuary
Sample Phrases		
build	To build a good network of friends around me Build a strong foundation to suceed that's why I choose U of T Build my network name and career	
rich	being rich be a rich man Become rich where I do not have to worry about running out of money	
business	start my own business Being successful having so many successful businesses have a international business	
own	Have my own company have my own house and car k Start my own business	
actuary	Be an actuary achieve my future career goal as an actuary Becoming an actuary	

Table A6: Qualities you admire in yourself

What do you admire about yourself? Why?		
	Top decile words	**Bottom decile words**
Significant at 99%	discipline	cause communicate receive
Significant at 95%	specific word practice responsibility smart confident game	friendly trust
Sample Phrases		
discipline	I admire that I've managed to attain discipline and spirituality I admire the fact that I have discipline enables me to have some discipline in my life	
specific	my persistence in one specific matter if i am interested in a specific thing I will go for it I feel most alive when I am giving everything I've got to a specific goal	
word	I admire my ability to be a clever speaker and use words to my advantage I would be careful about my words and behaviours in public places I try to keep my word at all times	
practice	get as much practice as I possibly can during my time at university to understand the lecture and practice a lot I am not so clever so that I need to practice a lot	
responsibility	im taking responsibility for my own health and fitness One of the qualities I admire most about myself is responsibility I have sober personality and have high sense of responsibility	
smart	Smartness I do smarter things than my peers People just tell me that i am smart we enjoy working hard and working smart	
confident	I can become competitive and confident studying at UofT feel more confident regarding being on track of all my obligations Grew more confident	

game	other students around my age who are keen on playing video games Being good at gaming if i need to play the game that people want from me i can do that
cause	previous has though in turn blinded me to the root cause of my problems causes me to be an extremely appreciative After I found the cause of my problems
communicate	Have patient to communicate with strangers It was easy for me to communicate with other people around me It allows me to make acquaintances easily and communicate comfortably
receive	receive a high level of education not expecting to receive anything back from them do not receive any running training
friendly	I am a friendly person who barely angry at others The best thing I admire myself is that I am very friendly I think i am friendly to my friends
trust	shows I could be trusted and reliable my first thought is to trust them i can make other to trust me

Table A7: Your future self

Describe what kind of person you want to become later on in life.		
	Top decile words	**Bottom decile words**
Significant at 99%	human meet people	tough
Significant at 95%	deal god trustworthy computer whole provide famous love mature determine helpful wise	man book father moment rich
	Sample Phrases	
human	I wish I can contribute to the human advancement I want to become an human I myself would be proud of We are all humans and we make mistakes	
meet	I would love to meet tons of people I wish I could be on the stage where I meet my satisfaction I want to try something different and meet new people	
people	Go all over and fix people's problems I want to be qualities that are inspiring to other people put themselves in other people's shoes and feel exactly what they feel	
deal	He is tackling space research to deal with overpopulation always has the confidence to deal with any problems independent person who can deal with problems	
God	importantly i want to find my path with God balance between worldly pleasures and devotion to God An avid worshipper of the Lord God	
trustworthy	i want them to know i very trustworthy I want to be known as someone who is trustworthy and truthful A fellow who is trustworthy honest and understanding	

computer	i see myself working in the field of Computer Science the reason why I want to choose Computer Science as a major business man or ceo or some computer company
whole	A person who changes the whole goddamn world and also can contribute to the whole society I just imagine that I am a real and famous musician in the whole world
provide	can provide assistances to others in need of help and support I should be able to provide our children with what they want be in tune with other people's needs and provide them with assistance
famous	don't have to be powerful or famous its not necessary to be famous I want to be famous and let everyone know about me although I don't expect that
love	Have a decent job a loving family a bunch of lovely friends Kind and loving parent with loads of patience I want to be loving and caring
mature	grown so much from the child that she was to a fully matured adult I want to become a more mature and socialized person I always want to be enterprising decisive mature thoughtful and careful
determine	I hope to be someone who is determined ambitious yet balanced I want them to be able to describe me as uplifting and determined I would like to be a persistent and determined person
helpful	That I am intelligent and hardworking as well as a very helpful person I want to be a kind and helpful person later on in life Helpful for the company
wise	I also want to be a wise person who talk and think deeply I want to become a extravert confident and wise person I'd like to be described as patient curious and wise

tough	I want to become a really tough woman Firstly I want to become a tough person I want to become a tough brave and caring person in the future
man	a happy and balanced man with a close group of friends and family I would like to become a perfect man Successful business man
book	I not only want to be book smart but I'd also like to be street smart make school a place where learning isn't just don't through a text book If somebody were to write a book on me
father	i want to be a good father I want to be a person as my father did When I was young my father told me to be a successful person
moment	ruthless and stern in the appropriate moments This is one of the qualities that I lack in at present moment As of this moment I can only hope to have the same character
rich	as a very rich guy and owns a big company I want become person like Bill Gates rich rich rich He is not rich but he lived as a real human

Table A8: Qualities you admire in others

What are some qualities you admire in others? Why? What qualities do you wish you had or could improve?		
	Top decile words	**Bottom decile words**
Significant at 99%	weakness avoid	
Significant at 95%	challenge overcome read mistake creativity word people Steve area general initiative understand	power concentration waste
Sample Phrases		
weakness	class in time and handing in my homework on time are my weakness I admire my father for everything from his strengths and weaknesses an individual who knows his her weaknesses	
avoid	I want to be as efficient as the others in order to avoid wasting time I tend to avoid problems until the very last minute I am a timid person and want to avoid stirring up animosity in others	
challenge	they are calm when they meet some challenges or something is difficult I know that university will be a challenge for me at first She has faced many challenges but managed to overcome them	
overcome	they know how to handle it and overcome any adversity they have overcome it and they accomplished certain things He use his own power to overcome many problems	
read	I know he likes reading It is his tips to understand the world I wish I can be interested in reading One of qualities I lack which I feel like I should have is reading	

mistake	I often crack under pressure when I feel like I made a mistake really amazing when people do not give up on me when I make mistakes people who are always willing to forgive people's mistakes
creativity	Sometimes I'm admire people who have quick minds and creativity His creativity impressed me which helps develop one's creativity
word	I wish I had the quality to say a lot of things with little words cannot understand what i am saying or can get the point of my words instructions I also have a hard time putting my thoughts into words
people	People that do well in presentations and don't procrastinate some people will always work as best and hard as they can I like to admire and praise on the good qualities that people possess
Steve	Steven Jobs he used his creative idea creative people like the former president of Apple Steve Jobs One is Steve Jobs because he pursues perfection
area	usually involved in most of the areas in school or at work I hope I can achieve in multiple areas general confidence in most areas of my life
general	Another person that I admire is UN Secretary General Ban Ki moon We live in a very secular society with a general emphasis on ME they in general believe themselves regardless of making mistakes
initiative	I admire people who have initiative like my dad I admire dedication professionalism communication initiative focus and initiative and hard-workmenship
understand	have a good understanding of equity and anti oppression They are very caring and understanding an ability of understanding the real situation
power	will power self determination confidence and drive she had a strong will and power of rapid decision a lot of my friends have a great creative power

concentration	That is what I'm lacking of concentration Concentration is the desirable quality that I wish I could have That ability to focus on the everyday tasks with concentration
waste	I did not work very hard and my tuition money was all a waste but time passes dramatically fast I do not want to waste it I feel like that have wasted so much time in my life