The Heterogeneous Effects of Changing SAT Requirements in Admissions: An Equilibrium Evaluation*

Emilio Borghesan

December 24, 2023

Many universities are reducing emphasis on standardized exam scores in admissions out of concern that the exams limit college access for students from disadvantaged backgrounds. This paper analyzes how such a policy change would affect enrollment patterns and graduation rates at four-year colleges in the United States. To do so, I build an equilibrium model in which colleges rebalance their admissions criteria towards other measures of students' human capital in the absence of standardized exam scores. The model allows high school students' application decisions and human capital investments to respond endogenously to the admissions policy, while colleges adjust admissions thresholds to maximize their objectives. I estimate the model using data from the Education Longitudinal Study of 2002. I find that banning the SAT has a negligible effect on the enrollment of under-represented minority (URM) students, despite estimating that many universities have substantial preferences for diversity. Elite colleges are worse off after banning the SAT, as they enroll students with lower skills and see graduation rates drop by 2.7 pp, while completion rates rise at less selective schools. A separate policy requiring all students to take the exam raises college completion for URMs by 2.2 pp relative to the SAT ban by helping colleges to identify stronger students.

JEL Classification: I20, I23, I24.

*Email: eb9090@princeton.edu. I am deeply indebted to Petra Todd, Andrew Shephard, and Francesco Agostinelli for their continued support and mentorship. Special thanks also to Hanming Fang, Michael Gilraine, Margaux Luflade, Aviv Nevo, Holger Sieg, and seminar participants at the University of Pennsylvania, CESifo Junior Workshop on the Economics of Education, KU Leuven Economics of Education Summer Event, Barcelona Summer Forum, and the Economics of Education Workshop at Valle Nevado.

1 Introduction

Standardized tests have been used to determine admission to higher education in the United States since at least the 1930s, but support for them has waned in recent decades as research has documented significant test score disparities by race and family income (Camara and Schmidt 1999, Freedle 2003). Low-income and under-represented minority (URM) students are also less likely to take the SAT and ACT exams, and an extensive literature shows how removing SAT-related barriers, either through increased access to testing facilities or mandatory exam-taking, can raise college attendance (Klasik 2013, Bulman 2015, Pallais 2015, Hyman 2017, Goodman 2016).[1] Many universities are therefore considering banning the SAT to increase access to college for disadvantaged students (del Rio 2021). By inviting applications from non SAT-takers, a ban on the SAT allows for a potentially larger and more diverse applicant pool. But, this comes at the cost of discarding information that may be useful for selecting skilled candidates for admission, and the consequences of this tradeoff for patterns of college attendance and completion are *a priori* ambiguous.

The goal of this paper is to analyze how changing standardized testing policies would affect patterns of sorting to college and rates of college completion, with particular emphasis on the outcomes for URM and low-income students.[2] I compare a policy that fully eliminates the SAT with an SAT-for-All policy that retains exam requirements in admissions while simultaneously requiring all high school students take the exam. To analyze these policies, I specify a model of the objectives of admissions departments and how they use the SAT to achieve them. Within the model, eliminating the SAT generates incentives for application behavior and human capital investment among high school students. These behavioral responses may alter the composition of college applicants and subsequently induce capacity-constrained colleges to modify their admissions criteria. This paper takes seriously the notion that students respond to changes in admissions criteria and that colleges will then have to respond to students' behavior.

Colleges in the model aim to enroll students who are knowledgeable and racially diverse, and they use grades and SAT scores as signals of each student's knowledge at the time of application. Knowledge is treated as a dynamic latent factor that evolves throughout high school in response to inputs that are both exogenous, like family and school characteristics, and endogenous (study time). Eliminating the SAT has two immediate effects. It causes colleges to rely more on the rest of each student's application, their

[1] Throughout this paper, I use SAT to refer jointly to the SAT and ACT exams.

[2] Under-represented minorities include individuals who identify as Black, Hispanic, Native American, or mixed race.

grades and demographic characteristics, when inferring their knowledge. And, it allows students who have not taken the SAT to apply to college.

High school students respond to these two immediate effects by solving a dynamic discrete choice problem with choices of how much to study, whether to take the SAT, and whether and where to apply to college. The direction of incentives for non SAT-takers is clear. Eliminating the SAT removes a barrier to college entry, raising their incentive to study and their probability of applying to college. However, former SAT-takers may face a reduced incentive to study if grades are not a sufficiently precise signal of their knowledge. The overall effect on patterns of college attendance depends on how these endogenous student responses affect the distribution of knowledge among applicants.

I estimate the model using the Education Longitudinal Study of 2002 (ELS 2002), a rich longitudinal survey of a cohort of students as they transition from high school to college. The ELS 2002 contains extensive information on high school grades, SAT scores, college applications, admissions decisions, and college attendance. I combine the ELS 2002 with data on SAT testing locations and dates, first used in Bulman (2015), and a comparable data set I gathered for the ACT. I leverage variation in SAT access and distance to college as exclusion restrictions that affect whether and where to apply to college to aid in identifying college preferences for student characteristics. I estimate the model by maximum likelihood using a nested fixed point algorithm. I then use the estimated model to evaluate several counterfactual admissions policies.

I find that eliminating the SAT from consideration at all schools causes a 0.5 percentage point (pp) increase in URM enrollment and a 2.8 pp increase in low-income student enrollment. These gains are entirely driven by increases in enrollment at less selective universities. The policy reduces sorting by knowledge: The average knowledge of students attending elite private colleges falls by 0.21 sd, while it increases by 0.08 sd at the least selective schools. The reduction in assortative matching causes completion rates at elite private colleges to fall by 2.7 pp and to rise at less selective universities.

These results conflate the effects of the model's four main mechanisms, and an instructive pattern emerges when examining the contribution of each component in isolation. The four mechanisms are a change in admissions criteria when the SAT is eliminated, endogenous applications, endogenous human capital investment, and supply side responses by colleges in equilibrium. Holding fixed the pattern of applications and test scores in the data, eliminating the SAT muddies application signals and reduces assortative matching by knowledge and by household income, which is correlated with knowledge. Allowing for endogenous applications while fixing study time further reduces sorting and raises enrollment of URMs, who are disproportionately unlikely to take the SAT.

However, allowing for endogenous study time increases stratification by income and knowledge. I estimate that grades are noisy measures of knowledge, so eliminating the SAT reduces the incentive to study among former SAT-takers. This reduction in study time causes poorer students with initially lower skills, who are closer to the admissions thresholds, to be narrowly rejected, while richer students remain. In equilibrium, colleges respond to the increase in applications by raising admissions standards and rejecting applicants who would have been marginal admits in partial equilibrium, further reducing college access for low-income and URM students.

Overall, banning the SAT fails to raise URM enrollment because there are too few high school students who fail to take the exam and who could out-compete those already applying to college. The difference in knowledge at the end of high school between SAT-takers and non-takers is 1.15 sd. I show how combining a ban on the SAT with a hypothetical intervention that raises skills for non exam-takers would instead enable many of them to out-compete SAT-takers, causing a large increase in URM college attendance.

I compare banning the SAT with an alternative policy recommended in Dynarski (2018) that mandates all high school students take the SAT. Like the No-SAT policy, SAT-for-All removes a barrier to college application, but it does so without reducing the amount of information available to colleges. Relative to an SAT ban, SAT-for-All causes the fraction of URMs attending a four-year college to rise by 1 pp and completion rates for URMs to increase by 2.2 pp as colleges manage to identify more skilled students for admission. Low-income student enrollment is similar to the No-SAT policy.

This paper's way of modeling admissions departs from standard approaches in the literature. Often, admission to college is modeled as a threshold-crossing model in terms of a continuous index (Kapor 2020). This approach would not be appropriate when some of the measurements comprising the index, say SAT scores, are missing because of the policy. This paper instead microfounds admissions criteria as a search for a dynamically evolving latent factor, leading admissions offices to rebalance their weights towards grades in an optimal way when SAT scores are no longer observed. This method delivers a probability of admission for every student, with or without an SAT score.

I modify the standard dynamic factor model of Cunha, Heckman, and Schennach (2010) and Agostinelli and Wiswall (2020) by including demographic-specific measurement parameters and by letting the initial distribution of knowledge vary by a set of covariates – such as mother's education, income, and race – that are likely correlated with investment prior to high school.[3] Together with the use of threshold rules for admission that vary by demographic, these modifications reproduce a unique feature of the

[3] Saltiel (2023) estimates a static latent factor model with demographic-specific measurement parameters.

college market in the United States, namely that admissions offices interpret grades and test scores relative to each student's background.[4] Within this framework, the same grade or SAT score will cause admissions offices to more aggressively update their prior if the measurement is particularly informative for that student or if observable factors put that student at an initial disadvantage. Colleges in the model use the Kalman Filter to generate these updates.[5]

Estimates of the dynamic factor model reveal that studying is productive (Stinebrickner and Stinebrickner 2008). An increase of ten hours per week causes knowledge to rise by 0.08–0.09 sd each year. I also find that much has been decided by the start of high school. URMs begin high school at a 0.62 sd disadvantage relative to white and Asian students. Ninth grade knowledge is sharply increasing in mother's education. I do not find evidence that the SAT math or verbal exams are more biased than any other measure in the data. If anything, GPAs show more evidence of bias than standardized tests. Moreover, GPAs are noiser than the SAT, and they are noisier for URMs than for white and Asian students. This suggests that colleges will struggle to identify highly skilled URM candidates for admission if they must rely more on grades.

This paper shows that modeling equilibrium in the college market is important when analyzing large changes in admissions policies. Papers that estimate equilibrium models of the market for college admissions in the United States include Epple, Romano, and Sieg (2006, 2008), Fu (2014), and Kapor (2020). My paper shares the three-part application-admission-matriculation equilibrium of Fu (2014) and Kapor (2020), but I add several novel features. I allow the distribution of grades and test scores observed by colleges to be endogenous with respect to the policy, I microfound college preferences for student characteristics, and I analyze the effects of admissions policies on college completion.

The endogenous mechanisms in the model are motivated by a growing literature that shows how pre-college human capital investment responds to changes in admissions policies. Tincani, Kosse, and Miglino (2021) and Cotton, Hickman, and Price (2022) analyze experiments that show how high school students trade off leisure and the probability of admission to college when deciding how much to invest in their skills. Leeds, McFarlin, and Daugherty (2017), Golightly (2019), and Akhtari, Bau, and Laliberté (2020) exploit

[4]Interpreting grades and test scores in the context of each student's educational opportunities and family background has been common since at least the 1990s. Bowen and Bok (2016) quote admissions deans who explain how the same grades and SAT scores would affect admissions probabilities differently for several hypothetical applicants.

[5]The Massachusetts Institute of Technology explains that they use a combination of factor analysis and thresholds to determine whether a student qualifies for admission. A blog on their website explains their recent decision to reinstate the SAT, stating "...we do not consider an applicant's [SAT] scores at all beyond the point where preparedness has been established as part of a multifactor analysis." (Schmill 2022)

changes in admissions policies at Texas public universities to show that effort increases when policies render admission more likely but decreases when admission becomes certain. Grau (2018) and Bodoh-Creed and Hickman (2017) similarly find that effort in high school is shaped by admissions criteria in Chile and the US. Bond et al. (2018) and Goodman, Gurantz, and Smith (2020) show how applications respond endogenously to SAT scores.[6] This paper incorporates these multiple mechanisms in an equilibrium framework, demonstrating their quantitative importance in shaping patterns of college attendance in a world without the SAT.

This paper is organized as follows. The next section describes the data used for the analysis and presents summary statistics. Section 3 describes the model that is taken to the data and discusses some of its properties. Sections 4 and 5 discuss identification and estimation of the model. Section 6 shows the estimated model parameters, while section 7 presents estimates of counterfactual policies and explains the mechanisms behind the results. Section 8 concludes.

2 Data

This study uses data from the Education Longitudinal Study of 2002 (ELS 2002). The ELS 2002 randomly samples a nationally representative cohort of students who were in the tenth grade in 2002 and follows them through high school, college, and into the labor market. Students are surveyed four times, in 2002, 2004, 2006, and 2012. In 2006, students are either in college or participating in the labor market and receive surveys tailored to their status. The 2012 survey wave, eight years after graduation from high school, records educational attainment.

The ELS 2002 contains multiple measurements of cognitive skills throughout high school. Grade-point averages (GPAs) for each year of high school have been converted to a common scale and are weighted by Carnegie units.[7] The ELS 2002 also contains SAT scores in math and verbal skills obtained from the College Board, and ACT scores in math, English, reading, and science obtained from ACT, Inc. In addition, the National Center for Education Statistics (NCES) administers exams in math and reading to all students in the ELS 2002 in grades ten and twelve. I use the criterion-referenced (as opposed to

[6] Caucutt and Lochner (2020) analyze how credit constraints affect college enrollment and completion. Otero, Barahona, and Dobbin (2021), Arcidiacono et al. (2011), and Arcidiacono (2005) analyze affirmative action's effects on labor market outcomes and potential mismatch between students and colleges. Dillon and Smith (2017) consider whether uncertainty in the admissions process leads to mismatch. Arcidiacono, Kinsler and Ransom (2023, 2022a, 2022b) show how preferences at Harvard University vary by race.

[7] A Carnegie unit corresponds to one course taken every day, for one period per day, for a full school year.

norm-referenced) math scores in the 10th and 12th grades, which do not recenter scores each year and make it possible to quantify changes in the acquisition of skills over time.

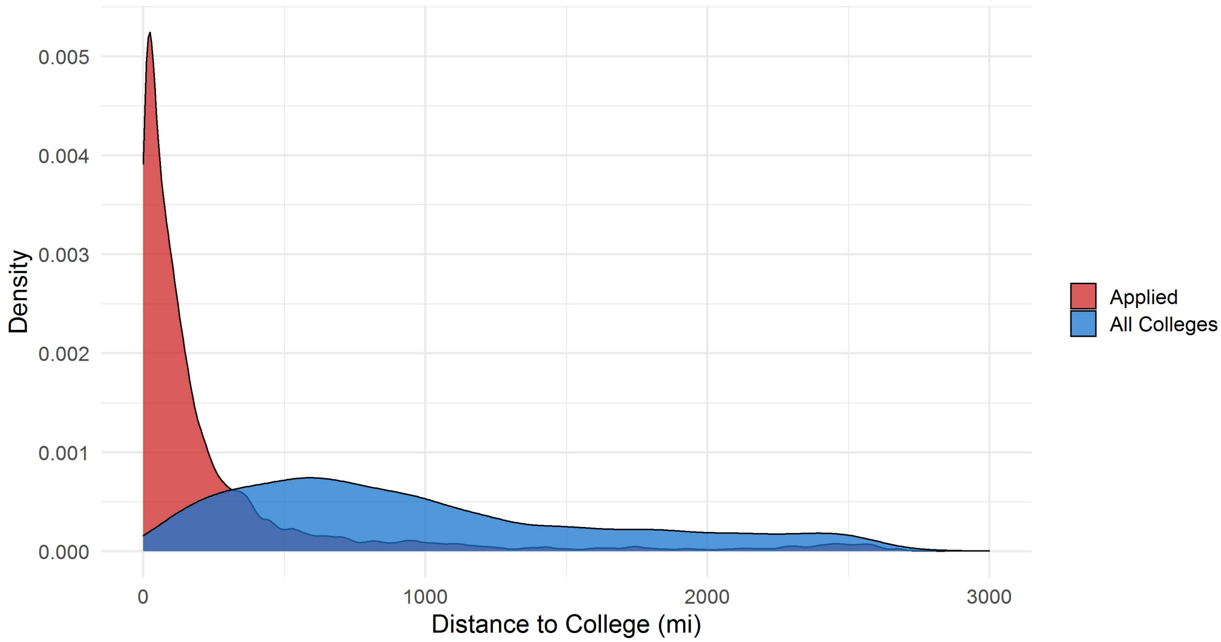
In 2002 and 2004, respondents report how much time they typically spend studying for classes. I average the two responses to create a single measure of average study time in high school. Stinebrickner and Stinebrickner (2004) compare student answers to the sort of time-use questions posed in the ELS 2002 with study time in time diaries, which they consider to be more accurate, and find a correlation of 0.72. If self-reported study time in the ELS 2002 is only an approximate measure of true study time, then averaging the two measures should reduce noise. Still, it is possible that estimates of the marginal productivity of study time in this paper represent a lower bound on its true effect.

I combine the ELS 2002 with a database of SAT testing center dates first used in Bulman (2015) and information from yearly ACT test registration booklets I obtained from ACT, Inc. These data allow me to construct a measure of access to the SAT. I define access to be the number of testing dates at one's own high school during spring of the junior year, when students typically take the SAT. Figure A-1 in Appendix A plots the distribution of exam access for students in the sample. The modal number of testing dates per school is zero, the mean is 1.09 days, and some schools that host both the SAT and ACT exams have up to seven testing dates during the course of the semester.

Test centers open after an employee at a particular school, typically a teacher or guidance counselor, volunteers to act as a test coordinator and applies to the College Board or ACT, Inc. to host the exam on a specific day. Testing sites must satisfy certain criteria, like having a quiet examination room and a secure location to store materials, but most applications are approved. The key factor in a school becoming a testing center is therefore whether someone at that school takes the initiative to apply. Bulman (2015) surveys fifty test coordinators to understand their motivations. Many expressed concern that nearby testing centers were at capacity and a desire to offer their students the exam in a familiar environment. In this paper, I control for school type (private, public, Catholic), geography (urban, rural, or suburban), total enrollment, and poverty rates within the school district (obtained from the Census Small Area Income and Poverty Estimates), all of which are likely to influence demand for testing facilities. I argue that residual variation in SAT access is quasi-random, stemming in part from differences across schools in whether an employee decides to apply.

This paper also exploits variation in distance to college to shift college applications. Figure 1 shows that, relative to distance to all colleges, the distribution of distance to colleges where students apply is weighted towards zero, suggesting that distance may shape application decisions. Carneiro and Heckman (2002) and Cameron and Taber (2004) ex-

Figure 1: Distance to College[8]



The figure shows the density of distance to college for students in the ELS 2002. Distance is computed between the centroid of each student's home census block and the latitude and longitude of each college in IPEDs. The figure plots both the unconditional density of distance in blue and the density of distance to schools applied to in red. Distances above 3,000 miles (relevant only for Alaska and Hawaii) have been truncated. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

press concern that distance to school may be correlated with student ability and thus endogenous. The model that I describe in the next section addresses this concern by letting distance shift demand for college conditional on a precise measure of ability, a student's posterior knowledge after grades and test scores in each year of high school are revealed.

After removing observations with missing data, the sample has $9,910$ observations.[9] Tables 1 and 2 present summary statistics for this sample. The measurements indicate that URMs have lower standardized test scores and GPAs. URMs have less educated mothers, are more likely to grow up in a household headed by a single parent, and are more likely to have been retained prior to high school. They also attend schools with higher class sizes, where more students qualify for free or reduced-price lunch, and they grow up in families with an average income that is about $23,000 less than white and Asian families. URMs actually attend schools with more SAT and ACT testing dates in the spring of their junior year, but this is largely due to white and Asian students attending smaller schools

[8] All figures and tables in this paper rely on data from the Educational Longitudinal Study of 2002, henceforth cited as (ELS 2002).

[9] Appendix B provides details regarding the construction of the sample.

Table 1: Summary Statistics, I

	URM		White & Asian	
	Mean	SD	Mean	SD
Measurements				
GPA, 9th grade	-0.31	0.98	0.22	0.91
GPA, 10th grade	-0.31	0.97	0.22	0.91
NCES Reading, 10th grade	-0.34	0.95	0.23	0.94
NCES Math, 10th grade	-0.41	0.92	0.26	0.94
GPA, 11th grade	-0.32	0.98	0.20	0.92
SAT Math	-0.57	0.90	0.08	0.93
SAT Verbal	-0.51	0.92	0.09	0.95
GPA, 12th grade	-0.35	1.02	0.19	0.92
NCES Math, 12th grade	-0.05	1.00	0.64	1.03
Controls				
Student-teacher Ratio	17.40	4.46	16.19	4.13
Free Lunch	0.28	0.23	0.15	0.16
Num Testing Dates	1.06	1.53	0.96	1.44
Residualized Num Testing Dates	-0.04	1.46	0.06	1.34
Yearly Household Income	50,000	40,500	72,800	49,000

The table shows means and standard deviations (SD) for knowledge measurements and controls in the ELS 2002. Apart from the 12th grade NCES Math exam, all knowledge measurements have been standardized by their sample mean and standard deviation. The 12th grade NCES math exam has been standardized by the mean and standard deviation of the 10th grade exam to permit longitudinal analysis of knowledge gains. I construct a verbal score for students who take the ACT by summing the English and reading subscales before standardizing. Free Lunch refers to the fraction of students at the student's school who qualify for a free or reduced-price lunch. Num Testing Dates refers to the number of SAT or ACT testing dates held at a student's school during the spring of their junior year of high school. This number is then residualized on controls for school type, geography, enrollment, and district poverty rates. SOURCE: (ELS 2002)

and Catholic or private schools where the exams are rarely held. After controlling for school size, geography, type, and the school district poverty rate, URM students have lower exam access, as indicated by the variable "Residualized Num Testing Dates."

Table 2 demonstrates that URM students are less likely to take the SAT, less likely to attend college, less likely to complete college, and less likely to complete conditional on attending college. They also study fewer hours while in high school. When URMs attend college, they attend very different colleges than white and Asian students. Figure A-2 in Appendix A indicates that URMs are under-represented at highly selective colleges and

Table 2: Summary Statistics, II

	URM	White & Asian
Choices		
Study Hours, per week	5.87	6.49
Take SAT	0.63	0.79
Attend 4-yr College	0.33	0.49
Complete 4-yr College	0.19	0.34
Complete 4-yr College Given Attendance	0.59	0.70
Initial Conditions		
Female	0.52	0.50
Retained before High School	0.10	0.06
Single Parent	0.32	0.16
Mother : High School	0.25	0.27
Mother : Some College	0.35	0.34
Mother : 4-year Degree	0.14	0.22
Mother : Postgraduate	0.07	0.11
Observations	2860	7050

SOURCE: (ELS 2002)

state flagships but are over-represented at less selective private universities and public satellite colleges. The model I describe in the next section will assess the sources of the college completion gap and evaluate whether alternative admissions mechanisms can raise completion rates.

3 Model

This paper uses an equilibrium model of the college market to analyze how eliminating the SAT affects patterns of college attendance and completion. In this section, I describe high school students' endogenous application and human capital investment behavior and colleges' optimization problem, which is affected by whether they observe the SAT.

3.1 Timing

The model has three time periods: high school, college transition, and college completion. I treat ninth grade GPA as an initial condition, and at the beginning of tenth grade students choose how much time to allocate to studying and whether they will take the

SAT. This choice of study time will take effect for three years. At the end of high school, grades and SAT scores are realized for each student. These measurements depend on a student's ninth grade knowledge, as well as time devoted to study, educational inputs, and idiosyncratic shocks. After these measurements are realized, the second period of the model, the transition to college, begins.

Similar to Kapor (2020) and Fu (2014), the transition to college consists of three parts. First students apply to college, then colleges decide which applicants to admit, and students with multiple admissions offers matriculate to their preferred feasible alternative.

College completion, up to eight years later, occurs in the final period of the model. Completion depends on students' knowledge when matriculating to college, the type of college they attend, and a set of controls.

3.2 Skill Technology and Measurement System

Knowledge is a latent variable that evolves deterministically as a result of prior knowledge, study time, and educational inputs, $\mathbf{I}_{i,t}$.[10] The technology of skill formation is allowed to differ by whether an individual belongs to an under-represented minority to account for whether differences in the marginal productivity of study time and schooling inputs may influence study decisions. Knowledge evolves according to the following value-added equation:

$$\log K_{i,t} = \gamma^{K,R} \log K_{i,t-1} + \beta^{H,R} h_{i,t} + \mathbf{I}'_{i,t} \boldsymbol{\beta}^{I,R} \ , \tag{1}$$

where $R \in \{URM, WA\}$ denotes parameters that are specific to either URM or white and Asian students.

I allow the distribution of ninth grade knowledge to vary by a set of predetermined covariates, $\mathbf{W}_i$, as follows:

$$\log K_{i,9} \sim N\left(\mathbf{W}'_i \mathbf{a}, \sigma^2_k(\mathbf{W}_i)\right) \ , \tag{2}$$

where the variance of ninth grade knowledge is given by $\sigma^2_k(\mathbf{W}_i) = \exp(\mathbf{W}'_i \mathbf{b})$. The vector $\mathbf{W}_i$ includes demographic and family controls to reflect how a history of unequal investment prior to high school can lead to skill differences by the ninth grade.

Grade point averages (GPAs) and standardized tests are noisy measures of knowl-

[10] Throughout the paper, I use bold font to denote vectors and matrices.

edge. In each grade, this mapping is

$$\mathbf{y}_{i,t}^{R} = \boldsymbol{\mu}_{t}^{R} + \boldsymbol{\alpha}_{t}^{R} \log K_{i,t} + \boldsymbol{\varepsilon}_{i,t}^{R} \,, \tag{3}$$

where $\boldsymbol{\mu}_t^R$ is a vector of intercepts, $\boldsymbol{\alpha}_t^R$ is a vector of factor loadings, and $\boldsymbol{\varepsilon}_{i,t}^R$ is a vector of normally-distributed disturbances with mean zero and a diagonal covariance matrix. As with the technology of skill formation, the measurement system is allowed to vary by URM status in an unrestricted way. This allows me to conduct inference on whether the SAT is biased against URMs and whether the informativeness of grades varies by demographic groups (section 6). I define bias and differential signal informativeness for measurement j in year t as follows:

$$\text{Bias} := \mu_{t,j}^{URM} - \mu_{t,j}^{WA} \,, \tag{4}$$

$$\text{Differential Signal Informativeness} := \frac{\alpha_{t,j}^{URM}}{\sigma_{t,j}^{URM}} - \frac{\alpha_{t,j}^{WA}}{\sigma_{t,j}^{WA}} \,. \tag{5}$$

Let $URM_i \in \{0,1\}$ indicate whether a student belongs to an under-represented minority, define the initial information set by

$$\Omega_{i,9} := \{URM_i, \mathbf{W}_i, \mathbf{y}_{i,9}, \{\mathbf{I}_{i,k}\}_{k=10}^{12}\} \,,$$

and subsequent updates by $\Omega_{i,t} := \{\Omega_{i,t-1}, \mathbf{y}_{i,t}, h_{i,t}\}$.[11] Admissions decisions, described in the next section, will depend on

$$\log K_{i,12} | \Omega_{i,12} \sim N(m_{i,12}, P_{i,12}) \,,$$

where $m_{i,12} := \mathbb{E}[\log K_{i,12}|\Omega_{i,12}]$ and $P_{i,12} := Var[\log K_{i,12}|\Omega_{i,12}]$ are obtained by the Kalman Filter (details provided in Appendix C).

The model treats grades and test scores as noisy measures of a dynamically evolving latent state. Colleges observe these measurements and form expectations over each student's knowledge at the time of application by using the Kalman Filter. Eliminating the SAT affects admissions decisions through changing the set of measurements available to filter this latent state. The informativeness of grades is therefore particularly relevant for college's ability to select skilled candidates for admission without the SAT.

[11] Individuals have full information over the realization of future educational inputs, $\mathbf{I}_{i,t}$. The sources of incomplete information in the model are over future actions, realizations of test scores, admissions, and college completion.

3.3 Preferences

3.3.1 Colleges

Colleges are grouped into tiers, $c = 1, \ldots, C$, each comprising a continuum of capacity-constrained colleges that have preferences over the knowledge and diversity of matriculating students.

An application to college c is defined to be a pair, (S_i^c, URM_i), where S_i^c is a scalar signal of student i's knowledge at the time of application and URM_i is the student's demographic. The signal is drawn from the distribution of twelfth grade knowledge conditional on $\Omega_{i,12}$: $S_i^c \sim f(\log K_{i,12} \mid \Omega_{i,12}) = N(m_{i,12}, P_{i,12})$. Hence, the signal is both college-specific and unbiased for $\log K_{i,12}$.[12]

Denote the set of applications to school c by $\mathcal{A}_c$. I define an admissions policy to be a mapping from the space of applications to acceptance probabilities:

$$\text{Policy} : (\mathbb{R} \times \{0, 1\})^{\mathcal{A}_c} \rightarrow [0, 1]^{\mathcal{A}_c} . \tag{6}$$

Define λ^c_{URM} to be the fraction of students matriculating to college c who belong to an under-represented minority: $\lambda^c_{URM} := \mathbb{P}(URM_i = 1 \mid \mathit{Attend}_{i,c} = 1)$. Colleges in tier c have a production function that is increasing in knowledge and diversity, and they choose an acceptance policy to solve

$$\begin{aligned} \max_{\text{Policy} \in [0,1]^{\mathcal{A}_c}} \quad & \kappa_c \mathbb{E}[\log K_{i,12} | \mathit{Attend}_{i,c} = 1] + (1 - \kappa_c) \log \lambda^c_{URM} \\ \text{s.t.} \quad & \sum_{i=1}^{N} \mathbb{P}(\mathit{Attend}_{i,c} = 1 \mid \Omega_{i,12}) = N^c , \end{aligned} \tag{7}$$

where N^c is college c's capacity, which I take to equal the number of students in the ELS 2002 who attend college c. The expectation is taken over the probability that a student matriculates conditional on the admissions policy. This specification assumes that admissions offices maximize a weighted sum of knowledge and diversity and that they satisfy their capacity constraints in expectation.[13] The weights are allowed to vary by college tier. While an admissions policy is defined to be a mapping from the space of applications to

[12] Drawing signals from $f(\log K_{i,12} \mid \Omega_{i,12})$ presumes that colleges observe study effort. A more realistic approach would have colleges integrate over the distribution of study effort, rather than observing it, when deciding whom to admit. It is, however, unlikely that after observing measurements in each year of high school, $y_{i,9}, \ldots, y_{i,12}$, integrating over study time would generate markedly different predictions of $\log K_{i,12}$. For computational reasons, this approach was not adopted.

[13] I do not directly model tuition and assume that both tuition setting and financial aid formulas are invariant to the counterfactuals explored in this paper.

acceptance probabilities, the structure of the problem leads to the following proposition, which states that admissions policies can be characterized by a pair of threshold rules.

Proposition 1. *The optimal policy for college c is a pair of demographic-specific threshold rules: (S_0^{c*}, S_1^{c*}).*

Proof. Let the policy rule for students with $URM_i = 1$ be arbitrary. Suppose the policy rule for students with $URM_i = 0$ is not a threshold rule. Then there exist two students, l and m, with probabilities of matriculation conditional on admission given by p_l and p_m, such that $S_l^c > S_m^c$ but $\mathbb{P}(Accept_l) < 1$ and $\mathbb{P}(Accept_m) > 0$. Consider the following modified admission policy: $\tilde{\mathbb{P}}(Accept_l) = \mathbb{P}(Accept_l) + \varepsilon$, $\tilde{\mathbb{P}}(Accept_m) = \mathbb{P}(Accept_m) - \varepsilon \frac{p_l}{p_m}$. The modified acceptance rule satisfies the constraint and leaves λ_{URM}^c unchanged, but increases $\mathbb{E}[\log K_{i,12} \mid Attend_{i,c} = 1]$. Hence, the optimal policy for $URM_i = 0$ students is a threshold rule in S_i^c.

Let the policy rule for students with $URM_i = 0$ be arbitrary. By similar argument, the optimal admissions policy for $URM_i = 1$ students is a threshold rule. Hence, the optimal admissions policy is a pair of threshold rules, (S_0^{c*}, S_1^{c*}), that (potentially) differ by demographic. □

Formulating college preferences in this way is equivalent to a model where colleges observe $\Omega_{i,12}$ and admit students with probability $\mathbb{P}(\log K_{i,12} > S_{URM_i}^{c*} \mid \Omega_{i,12})$. The use of threshold rules combined with a preference for URMs implies that colleges would choose to accept only URM students if they had greater average knowledge than white and Asian students and existed in the population in sufficient proportions. This is, however, not an empirically relevant scenario. The distribution of skills in the ELS 2002 together with a value of $\kappa_c < 1$ will cause the threshold to be lower for URMs than for white and Asian students, $S_1^{c*} < S_0^{c*}$, but for students from both groups to attend each tier.

3.3.2 Students

Students in the model choose how much to study, whether to take the SAT while in high school, and whether and where to apply to college upon graduation. I first describe the three parts of the college transition phase – application, admission, and matriculation – in reverse chronological order before characterizing the problem of a high school student.

Matriculation: The indirect utility function for a student who attends college c is expressed as the following linear function of a fixed effect for that school, net tuition, distance to college, the probability of completing college, and a Type 1 extreme value shock:

$$U_{i,c}(\Omega_{i,12}) = \underbrace{\bar{U}_c + \beta_T NetTuition_{i,c} + \beta_{D,c} Dist_{i,c} + \beta_P P(Complete_{i,c} = 1 \mid \Omega_{i,12}, c)}_{V_{i,c}} + \varepsilon_{i,c} \ . \tag{8}$$

Net tuition is the difference between posted tuition at school c, which may vary depending on whether the student is in-state or out-of-state, and the financial aid student i would receive at school c, $NetTuition_{i,c} = Tuition_{i,c} - Aid_{i,c}$.[14] The probability of completing college depends on $\Omega_{i,12}$, meaning that students form their expectation based on measurements that are observed by both students and colleges. Distaste for distance, $\beta_{D,c}$, varies by whether the college is public or private. $V_{i,c}$ represents the deterministic component of utility. The value of not attending college varies with local labor market conditions and geographic controls as follows:

$$U_{i,0} = \beta_W \log\left(\frac{Earn_i^{HS}}{Earn_i^{COLL}}\right) + \beta_S Suburb_i + \beta_R Rural_i + \varepsilon_{i,0} \ , \tag{9}$$

where the first term is the gender-specific average earnings differential between high school and college graduates in student $i's$ county. Earnings differentials are computed using the Quarterly Workforce Indicators. A positive value for β_W would indicate that students are less likely to attend college where earnings for non-college workers are high relative to the earnings of college graduates. Students choose from among their admissions portfolio, B, the option that maximizes their utility. The chosen option, C_i, satisfies

$$C_i = \arg\max_{c \in B} \{U_{i,c}\} \ , \tag{10}$$

and the probability of making choice C_i given admissions portfolio B is

$$P(C_i = c \mid B, \Omega_{i,12}) = \frac{\exp(V_{i,c})}{\sum_{k \in B} \exp(V_{i,k})} \ , \tag{11}$$

where B always contains the option of not attending college and obtaining utility $U_{i,0}$.

The value of being admitted to portfolio B is given by the following log-sum term:

$$U_{i,B} := \mathbb{E}[\max_{c \in B} U_{i,c}] = \log\left(\sum_{c \in B} \exp V_{i,c}\right) \ . \tag{12}$$

Admissions: Each student's application portfolio is transformed into an admissions

[14] Appendix D provides details on the computation of financial aid at each school.

portfolio depending on whether their application signals exceed the thresholds at the schools where they apply. I assume that application signals are iid draws from f($\log K_{i,12} \mid \Omega_{i,12}$), so the probability that student i obtains admissions set B given application set A is

$$P(B \mid A, \Omega_{i,12}) = \prod_{c \in B} \mathbb{P}(S_i^c > S_{URM_i}^{c*} \mid \Omega_{i,12}) \prod_{d \in A \setminus B} \mathbb{P}(S_i^d < S_{URM_i}^{d*} \mid \Omega_{i,12}) \ . \tag{13}$$

The distribution $f(\log K_{i,12} \mid \Omega_{i,12})$ is fully characterized by its mean and variance, $(m_{i,12}, P_{i,12})$, so I replace $P(B \mid A, \Omega_{i,12})$ with $P(B \mid A, m_{i,12}, P_{i,12})$. This means that, regardless of the number of measurements in $\Omega_{i,12}$, the state space for each individual is two-dimensional. High school students who are deciding how much to study and whether to take the SAT form expectations over $(m_{i,12}, P_{i,12})$ rather than over the realization of each individual GPA and exam score.

Application: Applicants to portfolio A pay a fixed cost of applying to each school and then a marginal cost of applying to additional schools within the same tier. Fixed and marginal costs may vary by school so that the total application cost can be written as

$$cost_i(A) = \sum_{c=1}^{C} FC_{i,c}(A) + MC_{i,c}(A) + \varepsilon_{i,A} \ , \tag{14}$$

where $\varepsilon_{i,A}$ represents unobserved factors that shift application costs and is modeled as a Type 1 Extreme Value shock with scale parameter λ_A. Letting $n_c(A)$ denote the number of applications to school c in portfolio A, I specify the fixed and marginal costs as follows:

$$\begin{aligned} FC_{i,c}(A) &= \mathbb{1}_{c \in A} \left(\delta_c^{(1)} + \delta_c^{(2)} Inc_i + \delta_c^{(3)} MomCollege_i + \delta_c^{(4)} Dist_{i,c} \right) \ , \\ MC_{i,c}(A) &= \max\{n_c(A) - 1, 0\} \left(\delta_c^{(5)} + \delta_c^{(6)} Inc_i \right) \ , \end{aligned}$$

where $MomCollege_i = 1$ if individual i's mother has a college degree. Fixed costs therefore vary by household income, mother's education, and the distance between student i's home and college c, while marginal costs vary by household income.

Students who apply to a portfolio, A, obtain a benefit that integrates over the expected utility of every possible admissions subset, B, that could be obtained from A. Formally, the utility of submitting application portfolio A is:

$$V_i^{Coll}(m_{i,12}, P_{i,12}, A) = \sum_{B \in A} P(B \mid A, m_{i,12}, P_{i,12}) U_{i,B}(m_{i,12}, P_{i,12}) - cost_i(A) \ , \tag{15}$$

where $U_{i,B}$ is written as a function of the state variables $(m_{i,12}, P_{i,12})$ to denote that the value of admissions set B depends on the probability of completing college at each school within B, which in turn varies with the mean and variance of the student's knowledge at the end of high school.

The student's portfolio choice problem is

$$A_i(m_{i,12}, P_{i,12}, SAT_i) = \underset{A \in \mathcal{A}(SAT_i)}{\arg\max} \{V_i^{Coll}(m_{i,12}, P_{i,12}, A)\} \ . \tag{16}$$

The set of possible application portfolios, $\mathcal{A}(SAT_i)$, depends on whether the student took the SAT while in high school. $\mathcal{A}(1)$ is the universe of all possible application portfolios, while $\mathcal{A}(0) = \{0\}$, because all four-year colleges required the SAT during this time.[15] I vary $\mathcal{A}(0)$ under counterfactual policy regimes that allow students to apply to college even without an SAT score. The probability of applying to application set A is

$$P(A \mid m_{i,12}, P_{i,12}, SAT_i) = \frac{\exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A)}{\lambda_A}\right)}{\sum_{A' \in \mathcal{A}(SAT_i)} \exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A')}{\lambda_A}\right)} \ . \tag{17}$$

The value of beginning the college application phase with state variables $(m_{i,12}, P_{i,12}, SAT_i)$ is given by the log-sum term

$$\begin{aligned} \overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i) &:= \mathbb{E}[max_{A \in \mathcal{A}(SAT_i)} V_i^{Coll}(m_{i,12}, P_{i,12}, A)] \\ &= \lambda^A \log \left(\sum_{A \in \mathcal{A}(SAT_i)} \exp\left(\frac{V_i^{Coll}(m_{i,12}, P_{i,12}, A)}{\lambda_A}\right)\right) \end{aligned}$$

High School: Students in high school have preferences over hours spent studying, $H_i(a)$, taking the SAT, $SAT_i(a)$, and their expectation of admission to college as follows:

$$\begin{aligned} U_i^{HS}(a) = (Z_i^{H'}\gamma_H)H_i(a) + (\gamma_S + \gamma_S^Z Z_i^{SAT} + \gamma_S^{Inc} Inc_i)SAT_i(a) + \varepsilon_i(a) + \\ \int \overline{V}_i^{coll}(m_{i,12}, P_{i,12}, SAT_i(a)) dF(m_{i,12}, P_{i,12} \mid \Omega_{i,9}, a) \ , \end{aligned} \tag{18}$$

where a refers to the action chosen by the student. The disutility of studying varies by covariates, Z_i^H, that include family income, gender, private school attendance, and 9th grade skills, $m_{i,9} = \mathbb{E}[K_{i,9} \mid \Omega_{i,9}]$. The preference for taking the SAT is allowed to vary by both income and exam access, Z_i^{SAT}, as defined in section 2. The SAT cost shifters permit

[15] Over 85% of schools within each tier required the SAT in 2004. A small number of liberal arts colleges were SAT-optional. The model assumes an SAT score is necessary to submit an application to each college tier.

the model to capture logistical challenges that limit students' ability to take the exam and thus apply to college, while the study cost shifters capture differences in cognitive skills and environmental factors that influence students' ability or inclination to study. $H_i(a)$ denotes the average amount of time spent studying each week while in high school. Students make this decision at the beginning of 10th grade, and study time is assumed to take effect for three consecutive years, i.e. $h_{i,t}(a) = H_i(a)$ for $t = 10, 11, 12$ in equation (1). $\varepsilon_i(a)$ is a Type 1 Extreme Value shock with scale normalized to one.

High school students choose an action to maximize their utility subject to the technology of skill formation and the measurement system. Their problem is written as follows:

$$\max_a U_i^{HS}(a) \tag{19}$$

subject to

$$\log K_{i,t} = \gamma^{K,R} \log K_{i,t-1} + \beta^{H,R} H_i(a) + \mathbf{I}'_{i,t}\beta^{I,R} \ ,$$

$$\mathbf{y}_{i,t}^R = \boldsymbol{\mu}_t^R + \boldsymbol{\alpha}_t^R \log K_{i,t} + \boldsymbol{\varepsilon}_{i,t}^R \qquad \text{for t = 10, 11, 12 and } R \in \{URM, WA\} \ .$$

Note that students do not observe $K_{i,t}$. The model assumes that students use the Kalman Filter to forecast the distribution of $(m_{i,12}, P_{i,12})$ given their initial conditions, $\Omega_{i,9}$, and their choice, a. The student cares about how their actions in high school influence their probability of admission to college and their chance of completing college. The true value of $K_{i,t}$ is irrelevant for admissions decisions, because colleges base their decisions on signals drawn from $f(\log K_{i,12} \mid \Omega_{i,12})$ and $K_{i,12} \notin \Omega_{i,12}$. Knowing the true value of $K_{i,12}$ might help students predict college completion, but, for the sake of simplicity, I do not add $K_{i,t}$ as an additional continuous state variable, and instead let completion depend on the same state variables, $(m_{i,12}, P_{i,12})$, as admission. This implies that students do not precisely know their own cognitive skill but learn about it from grades and SAT scores.[16]

3.4 College Market Equilibrium

A *College Market Equilibrium* is defined as a set of policy functions for students $\{a_i, A_i, C_i\}_{i=1}^N$, a set of threshold rules $\{S_{URM}^{c*}\}_{URM=0}^1$ for colleges $c = 1, \ldots, C$, and a distribution of grades and SAT scores, $\{\Omega_{i,12}\}_{i=1}^N$, such that

1. Given the admissions thresholds $\{S_{URM}^{c*}\}_{URM=0}^1$ for $c = 1, \ldots, C$, and the state variables in each period, the policy functions, $\{a_i, A_i, C_i\}_{i=1}^N$, solve students' maximization problems in (19), (16), and (10); and

[16] Stinebrickner and Stinebrickner (2012) and Arcidiacono et al. (2016) study how college students learn about themselves through the realization of course grades.

2. The admissions thresholds, $\{S^{c*}_{URM}\}^1_{URM=0}$ for $c = 1, \ldots, C$, maximize colleges' objective function in (7) subject to their capacity constraints, taking as given the realized distribution of student test scores, $\{\Omega_{i,12}\}^N_{i=1}$, the applications that have been submitted, $\{A_i\}^N_{i=1}$, the matriculation rules of the students $\{C_i\}^N_{i=1}$, and the actions of other colleges; and

3. The distribution of realized scores that colleges take as given, $\{\Omega_{i,12}\}^N_{i=1}$, is consistent with the initial conditions, $\{\Omega_{i,9}\}^N_{i=1}$, and student decision rules, $\{a_i\}^N_{i=1}$.

The equilibrium notion is a standard Nash Equilibrium with a consistency condition, and it is characterized by a system of equations in terms of the best response functions for all C colleges. These best response functions are derived from the first-order conditions of the college optimization problem in equation (7) and form a system of $2 \times C$ equations in $2 \times C$ unknowns. The two equations for college c are:

$$\frac{\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial S^{c*}_1}}{\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial S^{c*}_0}} = \frac{\kappa_c \frac{\partial \mathbb{E}[S^c_i|Attend_{i,c}=1]}{\partial S^{c*}_1} + (1-\kappa_c)\frac{\partial \mathbb{P}(i=URM|Attend_{i,c}=1)}{\partial S^{c*}_1}/\mathbb{P}(i=URM|Attend_{i,c}=1)}{\kappa_c \frac{\partial \mathbb{E}[S^c_i|Attend_{i,c}=1]}{\partial S^{c*}_0} + (1-\kappa_c)\frac{\partial \mathbb{P}(i=URM|Attend_{i,c}=1)}{\partial S^{c*}_0}/\mathbb{P}(i=URM|Attend_{i,c}=1)},$$

$$\sum_{i=1}^{N} \mathbb{P}(Attend_{i,c} = 1 \mid \Omega_{i,12}) = N^c \, .$$

I do not prove existence or uniqueness of the equilibrium. For certain extreme parameter values, colleges will be unable to satisfy their capacity constraints and an equilibrium will not exist. However, such a scenario is not empirically relevant. I have always been able to solve for an equilibrium, and I have never found multiple equilibria for a fixed set of parameter values. Different starting guesses for the admission thresholds converge to the same equilibrium, and small perturbations of the parameters produce equilibria with nearby thresholds. This suggests that the optimizer is not jumping between equilibria as it searches over the parameter space.[17]

3.5 College Completion

The ELS 2002 records whether each individual obtains a bachelors degree within eight years of graduating from high school. I model college completion as a production function that depends on student inputs, $K_{i,12}$, the tier of school the student attends, and

[17] Appendix G explains how I solve the model.

controls $\mathbf{X}_i \subseteq \mathbf{W}_i$ as follows:

$$Complete_{i,c} = \mathbb{1}(\omega_c^{(1)} + \omega_c^{(2)} \log K_{i,12} + \mathbf{X}_i'\omega^{(3)} + \eta_i > 0) \ , \tag{20}$$

where $\eta_i \sim N(0,1)$.[18] I let the constants in equation (20) vary by college, thereby capturing both observed and unobserved factors that influence rates of completion at each college. Students, who do not observe $K_{i,12}$, compute their completion probability by integrating over it using the distribution $f(\log K_{i,12} \mid \Omega_{i,12})$ when deciding where to matriculate. Hence, $P(Complete_{i,c} = 1 \mid \Omega_{i,12})$ in equation (8) is given by

$$P(Complete_{i,c} = 1 \mid \Omega_{i,12}) = \int \mathbb{P}(Complete_{i,c} = 1 \mid \log K_{i,12}) dF(\log K_{i,12} \mid \Omega_{i,12}) \ . \tag{21}$$

3.6 Discussion

The model assumes that each college prioritizes the knowledge and racial diversity of its student body. Specifying college preferences in this way is consistent with their mission statements, nearly all of which express a desire to enroll diverse and academically prepared students.[19] Papers that model college admissions all assume that colleges value cognitive skill. Epple, Romano, and Sieg (2006) additionally give colleges preferences for socioeconomic diversity, while Kapor (2020) and Epple, Romano, and Sieg (2008) add preferences for racial diversity. My choice of giving preferences for racial diversity is consistent with a principle goal of many who advocate for eliminating the SAT, namely to increase access to college for URMs (Soares 2020). Even without giving schools direct preferences for socioeconomic diversity, the model closely matches sorting to college by household income (section 6.4).

There is some concern that highly selective schools may place substantial emphasis on non-academic factors like athletic or musical skill. This is most plausible at tier one schools, where athletes constitute a significant share of enrollment. A correlation between preferences for these skills and the demographics of students possessing them may bias estimates of university preferences for diversity. To alleviate this concern, Appendix E shows that admissions chances for URM and white and Asians students at all schools are monotonically increasing in $\log K_{i,12}$. This suggests that the admissions model in the paper predicts who is admitted to these schools quite well. I also find that tier one colleges

[18] $\mathbf{X}_i$ includes URM status, family income, and an indicator for whether the mother has a college degree.

[19] Of the top 50 universities ranked by US News and World Report in 2022, 47 have clearly defined mission statements, of which 42 mention knowledge directly in the statement and 41 mention diversity or have a separate statement affirming a commitment to diversity. Schools that do not directly mention these words use related words like intellectual, discovery, and inclusion.

place a relatively low weight on diversity (Table 7), which puts an upper bound of the extent to which preferences for other factors contribute to diversity at these schools.

Eliminating the SAT in the model sets in motion a range of behavioral responses by high school students. First, it allows a new pool of students to apply to college. The model predicts that the new applicants will have a low application cost, as determined by income, distance to college, and mother's education, and either a high probability of admission or a high value for college attendance. Second, banning the SAT raises the incentive to study for non SAT-takers, for whom college may now be in reach, but weakens it for former SAT-takers, by inducing more noise in the relationship between study effort and college admission. The overall effect of removing the SAT boils down to how these endogenous responses affect the distribution of knowledge among college applicants in the new equilibrium.

The model allows for Roy-style sorting into college (Roy 1951). Students have preferences over college completion in equation (8), which is a reduced-form way of capturing both the pecuniary and nonpecuniary benefits of a college degree, while not attending college (equation 9) depends on the relative earnings of high school and college graduates in the county where the student attends high school. Students with a comparative advantage in college completion will therefore be more likely to attend college, while those with better labor market prospects will more likely enter the labor market after high school.

The model also allows for the possibility of mismatch in college completion (Bleemer 2022, Arcidiacono et al. 2014, Arcidiacono et al. 2011). A low-skilled student would be mismatched at a tier one school, for example, if $\omega_1^{(1)} < \omega_2^{(1)}$ and $\omega_1^{(2)} > \omega_2^{(2)}$ in equation (20). In this case, students with low $\log K_{i,12}$ would be more likely to complete at tier two schools, but students with high $\log K_{i,12}$ would be more likely to complete at tier one schools because of the greater return to knowledge at these schools.

There is some debate over whether rational expectations (RE) or some other form of expectations best characterize student perceptions of the admissions process.[20] In this paper, I give students RE over their probability of admission to college and over the effect

[20] Cotton, Hickman, and Price (2022) show in a field experiment that investment in human capital in the presence of affirmative action (AA) is consistent with rational expectations. Arcidiacono et al. (2020) collect data on subjective earnings expectations and occupational choice probabilities and find that they are highly predictive of future earnings and occupational choices. On the side of biased expectations, Hastings, Neilson, and Zimmerman (2015) demonstrate that students who choose unprofitable college degree programs considerably overestimate the earnings of past graduates while high-ability students have relatively accurate beliefs. Wiswall and Zafar (2015) and Delavande and Zafar (2019) find that providing accurate information about wages causes students to update their *beliefs* but has little effect on their *choices*, suggesting the presence of large nonpecuniary preferences for the type of college and field of study.

Table 3: Aggregation of Four-year Colleges into Groups

Type	Tier	Barron's Rank	Description	Examples
Private	1	1	Elite	Harvard, Swarthmore, Univ. of Chicago, USC
	2	2/3	Highly selective	Univ. of Miami, DePaul, Pepperdine
	3	4/5/6	Less selective	Univ. of Mobile, Concordia University-St. Paul, Monmouth University
Public	4	1/2	Elite	Univ. of Michigan, UCLA, UNC-Chapel Hill
	5	3	Most state flagships	Univ. of Wisconsin-Madison, Univ. of Arizona, most SUNY campuses
	6	4/5/6	Satellite campus, some flagships	Alabama A&M, Boise State, Northern Kentucky

of study time on academic performance. The ELS 2002 does not provide sufficiently rich data on subjective expectations to permit a major departure from RE.

3.7 Aggregation

This paper analyzes attendance and completion at four-year colleges. I group colleges according to a combination of their Barron's selectivity ranking and type (public vs private non-profit). The exact groupings are depicted in Table 3. These groupings have been chosen so that the analysis can speak to admissions practices at an identifiable set of schools – like elite public and private universities and state flagships – while retaining a sufficiently large sample size in each group to estimate preferences. Tiers one through three correspond to private colleges and universities, ranked in descending order of selectivity, while tiers four through six are public universities, ranked in descending order of selectivity. Community colleges are grouped together with no college as part of the outside option. Colleges in the same tier are assumed to have the same preferences and admissions thresholds. Classifying colleges in this manner is consistent with the purpose of the Barron's selectivity rankings, which aim to group schools together that have a common admissions standard.

Although it reduces the computational burden, aggregation creates challenges. Preferences for college in equation (8) depend on tuition and the distance student i would need to travel to attend school c. Which of the many schools within tier c should should determine the values of tuition and distance, $(Dist_{i,c}, Tuition_{i,c})$? In the analysis that follows, I choose the reference school for individual i to be the closest school within tier c.[21] Tuition for this reference school is in-state tuition when the student and school are located in the same state and out-of-state tuition otherwise.

To limit the size of the choice set, I allow students to send up to two applications to each college tier. I do not allow students to apply to all possible permutations of colleges, but instead limit them to the set of unique application portfolios in the data. Hence, while there are $3^6 = 729$ potential portfolios with up to two applications per tier, students in the model can choose from among the 584 unique portfolios observed in the data.

4 Identification

4.1 Dynamic Factor Model

When students submit an application to college, the admissions office observes a signal drawn from the distribution of latent knowledge conditional on $\Omega_{i,12}$. Estimating the model requires identifying the parameters that determine this distribution, which consist of the skill technology (equation 1) and the measurement system (equations 2 and 3).

Identification of the dynamic factor model follows from arguments in the literature (Cunha, Heckman, and Schennach 2010; Agostinelli and Wiswall 2020; Williams 2020). It is possible to write the entire vector of measurements throughout high school as a function of the initial knowledge draw, $\log K_{i,9}$.[22] To reduce notational clutter, the following equations condition on $\boldsymbol{\mu}_{i,t}$, H_i, $\mathbf{W}_i$, and $\mathbf{I}_{i,t}$:

$$\begin{pmatrix} \mathbf{y}_{i,9} \\ \mathbf{y}_{i,10} \\ \mathbf{y}_{i,11} \\ \mathbf{y}_{i,12} \end{pmatrix} = \underbrace{\begin{pmatrix} \boldsymbol{\alpha}_9^R \\ \gamma^{K,R}\boldsymbol{\alpha}_{10}^R \\ \gamma^{K,R^2}\boldsymbol{\alpha}_{11}^R \\ \gamma^{K,R^3}\boldsymbol{\alpha}_{12}^R \end{pmatrix}}_{\mathbf{A}} \log K_{i,9} + \begin{pmatrix} \boldsymbol{\varepsilon}_{i,9}^R \\ \boldsymbol{\varepsilon}_{i,10}^R \\ \boldsymbol{\varepsilon}_{i,11}^R \\ \boldsymbol{\varepsilon}_{i,12}^R \end{pmatrix},$$

[21] In many cases, this is the very school to which students apply. When students in the ELS 2002 apply to a school within in particular tier, the school they apply to is the closest one to their home between 20% and 50% of the time, depending on the tier. When the school that students apply to is not the closest within a given tier, it will often be the second closest, and distance to this school will be correlated with distance to the closest school, limiting the severity of measurement error.

[22] Appendix F explores whether adding a stochastic shock to equation (1) affects the inferences drawn from the dynamic factor model. It does not, and since identifying a model with this shock requires additional normalizations beyond those discussed in this section, the main analysis uses a deterministic skill technology.

where $\log K_{i,9}$ is one-dimensional, and $\mathbf{y}_{i,t}$, $\boldsymbol{\alpha}_t^R$, and $\boldsymbol{\varepsilon}_{i,t}^R$ are vectors with lengths that vary by t. As long as $\mathbf{A}$ contains at least three measurements, $\mathbf{A}$ satisfies the row deletion property and it is possible to separately identify $\mathbf{A}\Phi\mathbf{A}'$ from $\mathbf{\Sigma}_\varepsilon$, where $\Phi := var(\log K_{i,9})$.[23] A further normalization is needed to separately identify $\mathbf{A}$ and Φ. This is achieved by excluding a constant from $\Phi(\mathbf{W}_i) = \exp(\mathbf{W}_i'\mathbf{b})$, so that $\mathbf{W}_i = \mathbf{0}$ implies that $\Phi(\mathbf{W}_i) = 1$. With this normalization, $(\mathbf{A}\Phi\mathbf{A}')_{1,1}$ identifies $\boldsymbol{\alpha}_9$.

$\mathbf{A}$ is now separately identified from Φ, but it is still necessary to identify $\gamma^{K,R}$ separately from the other factor loadings, $\boldsymbol{\alpha}_{10}^R$, $\boldsymbol{\alpha}_{11}^R$, and $\boldsymbol{\alpha}_{12}^R$. It would, in principle, be possible to scale up $\gamma^{K,R}$ by c and scale down $\boldsymbol{\alpha}_{10}^R$, $\boldsymbol{\alpha}_{11}^R$, and $\boldsymbol{\alpha}_{12}^R$ by c, c^2, and c^3, respectively. I am able to rule out this observational equivalence, because the criterion-referenced NCES math exams in grades 10 and 12 are scored on the same vertical scale, which Agostinelli and Wiswall (2020) show implies that $\alpha_{10,j}^R = \alpha_{12,j}^R$ and $\mu_{10,j}^R = \mu_{12,j}^R$ for j equal to the NCES math exam and $R = URM, WA$.[24]

The mean of the latent factor is not separately identified from the mean of the measurements and is typically normalized to zero. This paper instead lets the mean of $\log K_{i,9}$ depend on a vector of initial conditions: $\mathbb{E}[\log K_{i,9} \mid \mathbf{W}_i] = \mathbf{W}_i'\mathbf{a}$. Note that it is not possible to identify a level shift in the constant for ninth grade measurements for underrepresented minorities, $\boldsymbol{\mu}_9^{URM}$, from a shift in the initial mean of knowledge for underrepresented minorities, $\mathbb{E}[\log K_{i,9} \mid URM_i = 1]$. A normalization is necessary, and I constrain the NCES math exams to have the same constants regardless of demographic: $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ and $\mu_{12,j}^{URM} = \mu_{12,j}^{WA}$ for j equal to the NCES math exam. This assumption is termed scalar invariance, and it means that there cannot be differences across demographics in how students interpret and answer the questions on the exam. Under this scalar invariance assumption, $\mathbb{E}[\log K_{i,9} \mid URM_i = 1]$ can be identified separately from $\mathbb{E}[\log K_{i,9} \mid URM_i = 0]$.[25]

4.2 Identification of College Completion and Preference Parameters

College preferences for diversity, $1 - \kappa_c$, are identified by the measurements of marginally admitted URM and white and Asian applicants. Lower GPAs and SAT scores for marginal URM admits relative to marginal white and Asian admits would identify a positive preference for diversity. This paper exploits variation in exam access and distance to col-

[23] This is proven by Theorem 5.1 in Anderson and Rubin (1956).

[24] I still allow for $\sigma_{10,j}^R$ to differ from $\sigma_{12,j}^R$ for the NCES math exams, so that the signal-to-noise ratios of the two exams may differ.

[25] Appendix F shows that the inferences drawn from the dynamic factor model are robust to alternative normalizations.

lege, which are excluded from college preferences but shift the probability of applying to college. Appendix A presents estimates from first-stage regressions of applications to college on exam access (Table A-1) and distance to college (Table A-2). The regressions show that exam access increases applications to college, while distance to college affects where students apply. The exclusion restrictions provide reassurance that the model is not identified solely on the basis of functional form assumptions.

This paper analyzes how counterfactual admissions policies affect sorting to college and rates of college completion. The model generates estimates of treatment effects for college completion at school j relative to school k, $Complete_{ij}(K_{i,12}) - Complete_{ik}(K_{i,12})$. These treatment effects are identified by randomness in admissions signals that causes students with the same $K_{i,12}$ to have different admissions sets and thus attend different colleges. Distributional plots of $\log K_{i,12}$ by school in section 6 reveal that there is considerable overlap in the knowledge distribution across colleges and thus sufficient support to analyze these treatment effects.

5 Estimation

I estimate the model by deriving the likelihood function and optimizing it using a Nested Fixed Point algorithm (NFXP). In estimation, I make use of the three NCES exams present in the ELS 2002, which are not observed by colleges, to aid in identification of the dynamic factor model. For clarity I distinguish between $\mathbf{y}_{i,10}$ and $\tilde{\mathbf{y}}_{i,10} = (\mathbf{y}_{i,10}, y^{(j)}_{i,10}, y^{(k)}_{i,10})$ for j and k equal to the NCES math and reading exams, and between $\mathbf{y}_{i,12}$ and $\tilde{\mathbf{y}}_{i,12} = (\mathbf{y}_{i,12}, y^{(j)}_{i,12})$ for j again equal to the NCES math exam. Colleges observe $\Omega_{i,12}$ while the econometrician observes

$$\tilde{\Omega}_{i,12} := \{\Omega_{i,9}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,12}, h_{i,10}, h_{i,11}, h_{i,12}\}\ .$$

For each student, I observe the college attended, C_i, the admissions set, B_i, the application set A_i, their observed measurements, $(\mathbf{y}_{i,9}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,12})$, their actions while in High School, a_i, and their initial conditions, $\Omega_{i,9}$. I also observe whether an individual graduates from college, $Complete_{i,c}$. Letting θ denote the entire set of model parameters, the likelihood contribution for individual i is

$$\begin{aligned} l_i(Complete_{i,c}, C_i, B_i, A_i, \tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,9}, a_i \mid \Omega_{i,9}, \theta) = P(Complete_{i,c} \mid C_i, \tilde{\Omega}_{i,12}, \theta)\times \\ P(C_i \mid B_i, \Omega_{i,12}, \theta) \times P(B_i \mid A_i, \Omega_{i,12}, \theta) \times P(A_i \mid \Omega_{i,12}, a_i, \theta)\times \\ f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid a_i, \Omega_{i,9}, \theta) \times P(a_i \mid \Omega_{i,9}, \theta) \times f(\Omega_{i,9}; \theta)\ , \end{aligned} \tag{22}$$

where $P(Complete_{i,c} \mid C_i, \tilde{\Omega}_{i,12}, \theta)$ comes from equation (21); $P(C_i \mid B_i, \Omega_{i,12}, \theta)$ comes from equation (11); $P(B_i \mid A_i, \Omega_{i,12}, \theta)$ comes from equation (13); $P(A_i \mid \Omega_{i,12}, a_i, \theta)$ comes from equation (17); $f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid \Omega_{i,9}, a_i, \theta)$ comes directly from the technology and measurement system in equations (1) and (3); $P(a_i \mid \Omega_{i,9}, \theta)$ is the solution to the problem of a high school student in (19); and $f(\Omega_{i,9}; \theta)$ are the initial conditions that vary with $\mathbf{W}_i$ in equation (2).[26]

I choose θ to minimize the log-likelihood function:

$$L(\theta) = \sum_{i}^{N} \log l_i(Complete_{i,c}, C_i, B_i, A_i, \tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10}, \mathbf{y}_{i,9}, a_i \mid \Omega_{i,9}, \theta) \; .$$

Optimization proceeds in two steps. The first step searches over the parameters that govern the dynamic factor model and college completion. It is possible to obtain consistent estimates of these measurement and technology parameters by optimizing over the following partial likelihood,

$$\sum_{i=1}^{N} \log P(Complete_{i,c} \mid C_i, \tilde{\Omega}_{i,12}, \theta) + \log f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid a_i, \Omega_{i,9}, \theta) + \log f(\Omega_{i,9}; \theta) \; , \quad (23)$$

which does not require solving for equilibrium in the college market. The second step searches over the preference parameters for both students and colleges by maximizing the remainder of the log likelihood function:

$$\sum_{i=1}^{N} \log P(C_i \mid B_i, \Omega_{i,12}, \theta) + \log P(B_i \mid A_i, \Omega_{i,12}, \theta) + \log P(A_i \mid \Omega_{i,12}, a_i, \theta) + \log P(a_i \mid \Omega_{i,9}, \theta) \quad (24)$$

I use the delta method to obtain standard errors for the parameters that are optimized through the second step.[27] .

[26] SAT math and verbal scores are not observed for all students, but I model them as missing at random conditional on the latent factor, $\log K_{i,t}$. The likelihood contributions of the measurement system, $f(\tilde{\mathbf{y}}_{i,12}, \mathbf{y}_{i,11}, \tilde{\mathbf{y}}_{i,10} \mid a_i, \Omega_{i,9}, \theta)$ in equation (22), therefore do not require a Tobit-style selection correction. Selection into SAT-taking occurs on the basis of $\log K_{i,t}$, but not on SAT-specific shocks ($\varepsilon_{i,t}$ in equation 3), which is consistent with the literature that typically treats standardized test scores as loading only onto cognitive skills in multidimensional factor models of human capital (Carneiro, Hansen, and Heckman 2003, Heckman, Stixrud, and Urzua 2006, Cunha, Heckman, and Schennach 2010).

[27] Appendix H describes how the standard errors are computed.

6 Results

6.1 Dynamic Factor Model

The parameters of the initial skills distribution, in Table 4, reveal dramatic differences in ninth grade knowledge across individuals. The normalizations discussed in section 4.1 mean that the coefficients in the column labeled Mean can be interpreted in terms of standard deviations. The table shows that average ninth grade knowledge is 0.62 sd lower for URMs relative to white and Asian students. Students who were retained prior to high school have nearly three-quarters of a sd lower knowledge, students who grew up with a single parent lag behind by 0.12 sd, and initial knowledge is sharply increasing in mother's education. The coefficient on household income indicates that initial knowledge is higher by 0.09 sd for each additional $100 spent on the child per week.[28] The variance of initial knowledge is lower for girls and students growing up in richer households. Retained students have lower variance by 0.30 log points, consistent with them typically being selected from the left tail of the skill distribution.

Table 5 displays estimates of $\boldsymbol{\mu}_t^R$ for $R = URM, WA$. The table can be used to assess whether grades and exams are biased against URMs, as $\boldsymbol{\mu}_t^R$ governs level shifts in the measurements across demographic groups after controlling for knowledge. Recall that the identifying normalization discussed in section 4.1, $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ for j equal to the NCES math exam, rules out bias in this exam. The numbers in Table 5 should therefore be interpreted as bias relative to this exam.[29] There does not appear to be evidence that the SAT is biased against URMs. In fact, URMs score marginally higher than might be expected conditional on their knowledge. The estimated parameters in Table 5 suggest that, if anything, GPAs are more biased against URM students than standardized exams.[30]

Even if the SAT is not biased against URMs, its informativeness may still vary across demographic groups. The right panel of Table 5 presents estimates of signal-to-noise ratios for the same measurements. For a given measurement, $y_{t,j}$, the signal-to-noise ratio is computed as $\alpha_{t,j}/\sigma_{t,j}$. The table indicates that GPAs become worse signals in later years of high school. It also shows that the standardized exams convey significantly greater information than GPAs, with the math portion of the SAT and the math exams administered

[28] I assume, consistent with a range of estimates reviewed in Donni (2015), that families spend one quarter of their household income on the child. The median value of this variable in the data, 3, represents $300 per week and corresponds to a yearly income of $62,400 ($3 \times 100 \times \frac{1}{0.25} \times 52 = 62400$).

[29] Appendix F shows that the inferences in the table are robust to alternative normalizations.

[30] Implicit bias among teachers, as measured by the Implicit Association Test, has been shown by Carlana (2019) and Van den Bergh et al. (2010) to predict both gender and racial test scores gaps and could be a source of the GPA biases seen here.

Table 4: Parameters Governing Initial Distribution of Knowledge

	Mean	Log Variance
URM	-0.62	0.03
	(0.03)	(0.05)
Female	-0.05	-0.11
	(0.02)	(0.03)
Retain	-0.73	-0.30
	(0.04)	(0.06)
Single Parent	-0.12	0.00
	(0.03)	(0.04)
Mother: High School	0.20	0.01
	(0.04)	(0.06)
Mother: Some College	0.36	-0.03
	(0.04)	(0.06)
Mother: Bachelors	0.66	0.04
	(0.05)	(0.06)
Mother: Postgraduate	0.85	0.04
	(0.06)	(0.07)
HH Income	0.09	-0.02
	(0.01)	(0.01)

The table presents estimates of parameters governing the initial distribution of knowledge in the ninth grade. The mean and variance have been normalized to 0 and 1, respectively, for individuals whose covariates are all equal to 0. HH Income is measured in hundreds of dollars per week. High school dropout is the omitted education category. Details regarding the distribution of knowledge are provided in section 3. SOURCE: (ELS 2002)

by the NCES being particularly informative. GPAs in every grade are less informative for URMs than for white and Asian students. The math portion of the SAT is also a worse signal for URMs, while the NCES math exams are more informative for URMs.[31]

The parameters of the measurement system therefore indicate that the SAT does not appear to be biased against URMs. But, there is merit to the concern that the math portion of the SAT may not be as informative for a URM student as it is for a white or Asian student. The same can also be said for grades in high school. Much of the literature on grading practices, for example Botelho, Madeira, and Rangel (2015) and Rauschenberg (2014), has focused on the first moments of grades. The results presented here suggest that second moments may also vary across demographic groups.

[31]This conclusion rests on an assumption of configural invariance, namely that the latent factor identified by the measurement system reflects the same underlying traits for both URMs and white and Asian students (Putnick and Bornstein 2016).

Table 5: Bias and Signal-to-Noise Ratios

	$\boldsymbol{\mu}_t^R$			$\frac{\alpha_{t,j}}{\sigma_{t,j}}$		
	URM	WA	Difference	URM	WA	Difference
GPA, 9th grade	-0.20	-0.16	-0.04	0.66	0.90	-0.24
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
GPA, 10th grade	-0.31	-0.27	-0.04	0.63	0.86	-0.23
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
GPA, 11th grade	-0.42	-0.37	-0.05	0.54	0.73	-0.19
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.10)
GPA, 12th grade	-0.53	-0.38	-0.15	0.41	0.53	-0.12
	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.05)
SAT Math	-0.98	-1.04	0.06	1.55	1.78	-0.22
	(0.03)	(0.03)	(0.02)	(0.06)	(0.06)	(0.02)
SAT Verbal	-0.89	-0.91	0.01	1.15	1.15	0.00
	(0.03)	(0.03)	(0.02)	(0.06)	(0.04)	(0.05)
NCES Reading, 10th grade	-0.35	-0.32	-0.02	1.04	0.99	0.05
	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)
NCES Math, 10th grade	-0.41	-0.41	0	2.22	1.79	0.43
	(0.03)	(0.03)	(-)	(0.08)	(0.06)	(0.05)
NCES Math, 12th grade	-0.41	-0.41	0	2.11	1.89	0.23
	(0.03)	(0.03)	(-)	(0.08)	(0.06)	(0.05)

The table displays estimates of bias (in the left panel) and signal-to-noise ratios (in the right panel). These correspond to $\boldsymbol{\mu}_t^R$ and $\frac{\alpha_{t,j}}{\sigma_{t,j}}$ in equation (3). The NCES exams are used to identify the technology of skill formation, but are not available to colleges when determining whom to admit. Estimates of the proportion of variance in each measurement due to the latent factor are provided in Table A-3 in Appendix A. WA refers to the population of white and Asian students. SOURCE: (ELS 2002)

Table 6 presents estimates of the technology of skill formation. The parameter estimates are similar for URMs and white and Asian students. Both have an autoregressive parameter for knowledge approximately equal to one, indicating that knowledge does not depreciate throughout high school. This suggests that it may be difficult for students who enter high school with a low level of knowledge to catch up to their peers by the time they apply to college. I also find that an additional ten hours of study time per week increases knowledge by 0.08–0.09 sd. A URM student who studies 10 hours per week will, all else equal, improve her skills by 0.27 sd between the end of 9th grade and the end of high school. Studying may therefore deliver significant marginal returns for students whose initial conditions place them on the cusp of gaining admission to college. Students

Table 6: Technology of Skill Formation

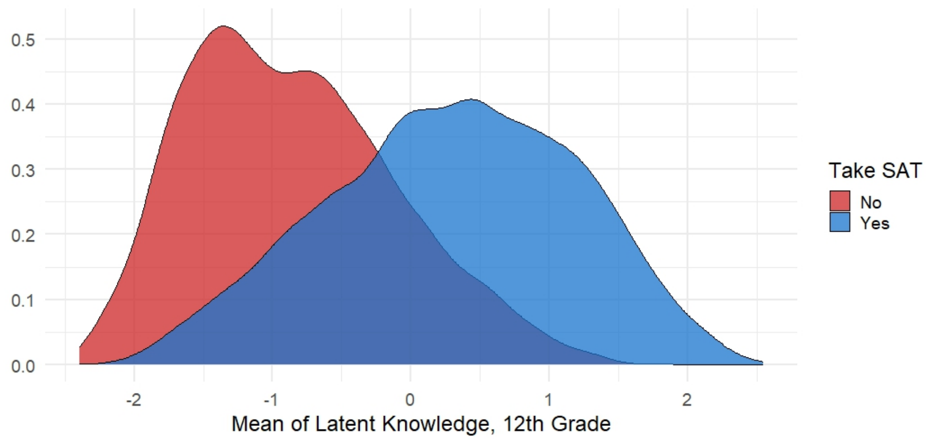
	URM	White & Asian
Knowledge(-1)	1.03	1.05
	(0.01)	(0.00)
Study, 10 hours/wk	0.09	0.08
	(0.01)	(0.01)
Private School	0.05	0.02
	(0.01)	(0.01)
Free Lunch	-0.11	-0.18
	(0.03)	(0.02)
Student Teacher Ratio	0.00	0.00
	(0.00)	(0.00)
Mother: High School	-0.01	0.00
	(0.02)	(0.01)
Mother: Some College	-0.02	-0.01
	(0.01)	(0.01)
Mother: Bachelors	-0.03	-0.02
	(0.02)	(0.02)
Mother: Postgraduate	-0.03	-0.02
	(0.02)	(0.02)
Constant	0.20	0.18
	(0.03)	(0.02)

The table displays estimates of parameters governing the technology of skill formation. Study refers to the effect of studying 10 hours per week on next year's skills. Free lunch is measured on a scale from 0 to 1. High school dropout is the omitted education category. More details on the technology are provided in section 3. SOURCE: (ELS 2002)

at poorer high schools, as indicated by the proportion of students qualifying for a free or reduced-price lunch, accumulate less knowledge. The effect of class size on skill development is insignificant. Attending a private school, either catholic or nondenominational, has a positive effect on knowledge accumulation for URMs. Mother's education has little effect on value added, in contrast to its effect on initial knowledge.

It is possible to use the dynamic factor model to analyze the distribution of knowledge at the end of high school. Figure 2 plots the densities of posterior knowledge means, $m_{i,12} = \mathbb{E}[K_{i,12} \mid \tilde{\Omega}_{i,12}]$, across students, grouped by whether the student took the SAT or not. The figure shows that SAT-takers typically have much greater knowledge than non-takers: The difference in means between the two groups is 1.15 sd. There are also very few students with knowledge one sd above the mean who fail to take the SAT. While these distributions are endogenous with respect to college admission policies, the large skill differences measured in the data will make it difficult for most non-takers to out-

Figure 2: Knowledge Distribution, 12th Grade



The figure shows the density of the mean of 12th grade knowledge, $m_{i,12} = \mathbb{E}[K_{i,12} \mid \tilde{\Omega}_{i,12}]$, by SAT-taking status. $m_{i,12}$ has been standardized to have zero mean and unit variance. SOURCE: (ELS 2002)

compete SAT-takers to gain admission to college under any policy that aims to admit knowledgeable students.

6.2 Estimated Preferences

Table 7 presents estimates of the preference parameters for students and colleges. The fixed effects for each college are ordered in a way that is consistent with selective universities being more highly valued. Distaste for distance is minor at the matriculation stage, although it will be significant at the application stage (Table 9). The coefficient on Net Tuition is positive, indicating that there may be some quality differences between less and more expensive colleges that cause matriculating students to prefer attending the latter.[32] Students exhibit a strong preference for college completion. Comparison with the fixed effects, $\overline{U}_1, \ldots, \overline{U}_6$, reveals that college is highly valued because of the degree that it confers rather than because of amenities unrelated to the degree.

Students dislike studying, but studying is less costly if they are female, attend private school, or have higher ninth grade skills. SAT-taking is also costly but greater income and logistical access make it less so. To put some of the numbers in Table 7 in perspective, the disutility of studying ten hours a week for a student with ninth grade knowledge one standard deviation above the mean is $(-0.14 + 0.07 \times 1) \times 10 = -0.70$. This is worth $0.70/3.50 = 0.20$ of a college degree. This student would therefore study ten additional hours a week if doing so increased her probability of attaining a college degree by 20% or more.

[32] Net tuition, like all monetary variables in the paper, is expressed in hundreds of dollars per week.

Table 7: Preference Parameters

	Value	Standard Error
Student Preferences for Universities		
College 1, $\overline{U}_1$	0.12	0.13
College 2, $\overline{U}_2$	-0.96	0.18
College 3, $\overline{U}_3$	-1.09	0.24
College 4, $\overline{U}_4$	0.22	0.08
College 5, $\overline{U}_5$	-0.71	0.07
College 6, $\overline{U}_6$	-0.86	0.19
Portfolio Shock, Scale	1.08	0.02
Distance, Public	-0.10	0.09
Distance, Private	-0.07	0.40
Net Tuition	0.40	0.17
Completion	3.50	0.30
Student Preferences for Outside Option		
Earnings Differential	0.14	0.17
Suburban	0.46	0.18
Rural	0.48	0.22
Student Preferences in High School		
Hours, γ_H	-0.14	0.04
Hours $\times$ Income, γ_H^{Inc}	0.00	0.00
Hours $\times$ Female, γ_H^{F}	0.05	0.03
Hours $\times$ Priv, γ_H^{Priv}	0.05	0.03
Hours $\times$ Knowledge, γ_H^{M9}	0.07	0.01
SAT, γ_S	-1.44	0.06
SAT $\times$ Access, γ_S^{Z}	0.08	0.07
SAT $\times$ Income, γ_S^{Inc}	0.20	0.03
College Preferences		
Preference for Knowledge, Tier 1	0.90	0.06
Preference for Knowledge, Tier 2	0.89	0.02
Preference for Knowledge, Tier 3	0.83	0.03
Preference for Knowledge, Tier 4	0.93	0.00
Preference for Knowledge, Tier 5	0.96	0.02
Preference for Knowledge, Tier 6	0.85	0.01

The table displays estimates of preference parameters of students and colleges in the model. Net Tuition and Income are measured in hundreds of dollars per week. Distance is measured in hundreds of miles. Earnings Differential is the gender-specific percent difference in average earnings between high school and college graduates in the student's county. More details are provided in section 3. SOURCE: (ELS 2002)

Table 8: Admissions Thresholds

	Raw		*Standardized*	
	White & Asian	URM	White & Asian	URM
Tier 1	2.50	1.81	1.07	0.54
Tier 2	1.43	0.86	0.24	-0.20
Tier 3	0.75	0.02	-0.29	-0.85
Tier 4	2.03	1.62	0.71	0.39
Tier 5	1.29	1.09	0.13	-0.03
Tier 6	0.47	-0.18	-0.51	-1.01

The table displays estimated admissions thresholds for each tier of college. Students who apply to a school with an application signal that exceeds their demographic-specific threshold are granted admission. Columns labeled Raw indicate the threshold in terms of $\log K_{i,12}$, while in the columns labeled Standardized, the thresholds have been normalized by subtracting the mean and dividing by the standard deviation of $\log K_{i,12}$. SOURCE: (ELS 2002)

Private colleges value knowledge less and diversity more that equally selective public colleges: Ten hundredths of top tier private universities' utility is derived from the diversity of its student body, while the corresponding fraction for top tier public universities (tier 4) is seven out of one hundred. State flagship universities (Tier 5) place the lowest weight on diversity. The weight placed on diversity causes a wedge to arise between the two admissions thresholds at each college. Table 8 shows that these thresholds are lower for URM than for white and Asian students within each college tier.

Table 9 shows estimates of the application cost parameters. Fixed costs – which vary by income, mother's education, distance, and a tier-specific constant – capture both monetary and nonmonetary deterrents to applications. A single application to a tier one school is over four times as costly as an application to a tier six school, all else equal. The signs on income show that higher income students prefer elite schools and state flagships relative to schools in tiers 3 and 6. At the same time, the evidence in Table 9 suggests that nonmonetary costs are more salient. Having a mother with a college degree dramatically reduces the fixed cost of applying to a tier one school. This effect, -1.07, is greater than the effect of moving from the poorest to the richest household in the sample $(-0.06 \times (9.62 - 0.24)) = -0.563$. Distance also influences where students apply, and it tends to be less salient at more selective schools. These parameter estimates are consistent with the existence of both information frictions and biased beliefs, which have been found to shape application choices (Hoxby and Avery 2012, Bleemer and Zafar 2018).

Table 9: Application Cost Parameters

	Fixed Cost				*Marginal Cost*	
	Constant	HH Income	Mother's Ed	Distance	Constant	HH Income
Tier 1	3.96	-0.06	-1.07	0.30	0.17	-0.06
	(0.16)	(0.01)	(0.10)	(0.09)	(0.39)	(0.08)
Tier 2	2.58	-0.06	-0.78	0.46	1.21	-0.10
	(0.05)	(0.02)	(0.10)	(0.25)	(0.05)	(0.03)
Tier 3	2.04	0.03	-0.25	1.35	1.61	-0.04
	(0.14)	(0.03)	(0.08)	(0.42)	(0.14)	(0.04)
Tier 4	2.64	-0.07	-0.79	0.58	1.18	-0.07
	(0.07)	(0.01)	(0.14)	(0.09)	(0.09)	(0.02)
Tier 5	1.72	-0.04	-0.41	1.11	2.41	-0.16
	(0.05)	(0.02)	(0.16)	(0.18)	(0.06)	(0.03)
Tier 6	0.84	0.04	-0.03	0.25	0.84	-0.03
	(0.06)	(0.03)	(0.25)	(0.07)	(0.04)	(0.03)

The table displays estimates of parameters governing the cost of applying to college. Income is measured in hundreds of dollars per week. Mother's Ed is an indicator for whether an individual's mother has a bachelor's degree. Distance is measured in hundreds of miles. Fixed and marginal costs are school-specific. More details on application costs are provided in section 3. SOURCE: (ELS 2002)

6.3 College Completion

Table 10 displays estimates of the college completion parameters. The estimates show that greater knowledge at the time of matriculation increases completion probabilities at all schools. Averaging across all individuals in the sample, the coefficients in Table 10 imply that increasing $\log K_{i,12}$ by one unit (the sd of $\log K_{i,12}$ is 1.28) would increase the probability of completion by anywhere between 6.7 pp at tier 4 schools and 12.7 pp at tier 6 schools. I do not find evidence of complementarities between student knowledge and college selectivity. Completion rates are actually higher for highly skilled students at *less* selective schools (tier 6), although these differences are not significant.[33]

Only tier 6 schools have a statistically significant constant, indicating that completion is lower at these schools after controlling for knowledge (Bound, Lovenheim, and Turner 2010). Although URMs are 11 pp less likely to complete college conditional on enrolling (Table 2), there is no significant difference in completion rates by demographic after controlling for knowledge and the college attended. I also find that having a college-educated

[33]The lack of complementarities between student quality and school selectivity in terms of degree attainment within eight years is consistent with Dillon and Smith (2020), who find evidence of such complementarities only in college completion within four years but not within longer time horizons.

Table 10: College Completion Model

	Estimate	Standard Error
Tier 1	-0.25	0.41
Tier 2	0.05	0.17
Tier 3	-0.15	0.10
Tier 4	0.11	0.23
Tier 5	-0.05	0.15
Tier 6	-0.35	0.07
$\log K_{i,12} \times$ Tier 1	0.28	0.13
$\log K_{i,12} \times$ Tier 2	0.25	0.07
$\log K_{i,12} \times$ Tier 3	0.26	0.06
$\log K_{i,12} \times$ Tier 4	0.18	0.09
$\log K_{i,12} \times$ Tier 5	0.25	0.07
$\log K_{i,12} \times$ Tier 6	0.34	0.04
URM	-0.03	0.05
HH Income	0.02	0.01
Mother has college degree	0.07	0.04

The table displays parameters of the college completion model, equation (20). The constant and slope with respect to knowledge, $\log K_{i,12}$, vary by college tier. Income is measured in hundreds of dollars per week. More details on college completion are provided in section 3. SOURCE: (ELS 2002)

mother has no effect on college completion and that family income is marginally significant. The results suggest that, if URMs attended the same colleges and had similar skills at the time of matriculation as white and Asian matriculants, the college completion gap would disappear.

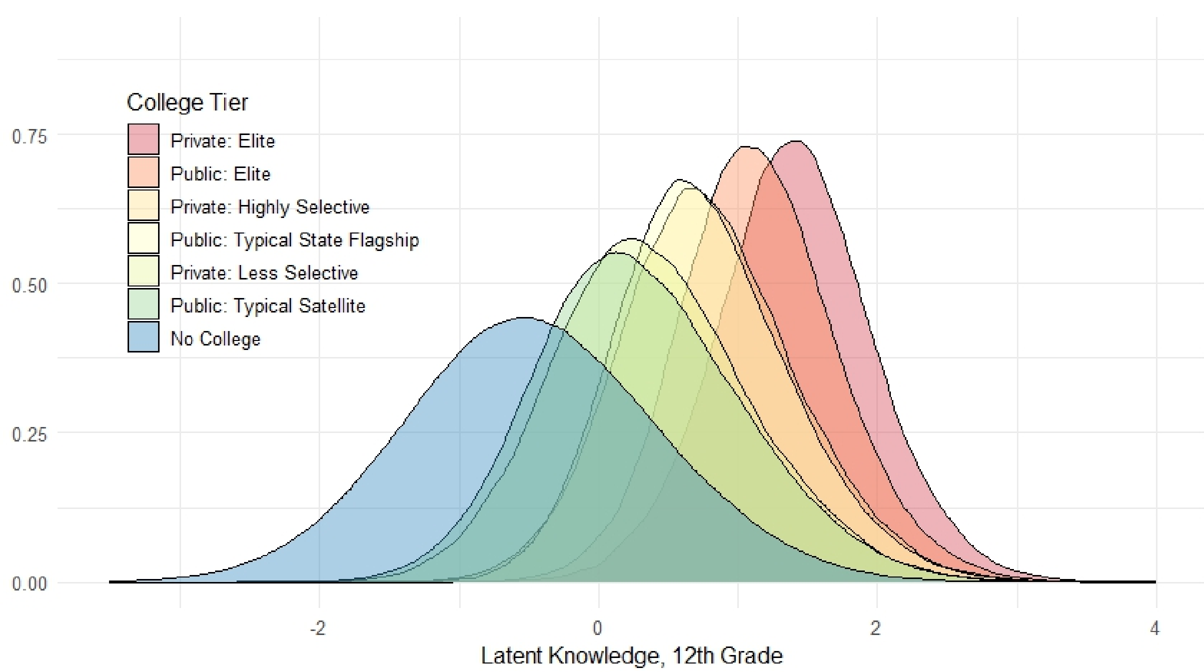
6.4 Model Fit

Table A-4 in Appendix A shows that the model successfully replicates many data moments. The model matches the proportions of students at each type of college, as well as the overall percentage of students, 44%, attending any four-year university. The model also closely replicates rates of college completion conditional on enrollment, both overall (68% in both the data and model simulations) and separately by demographic. The model also reproduces the clear pattern of sorting to universities by SAT scores and by household income that exists in the data.

The model matches hours spent studying quite well: 6.31 hours in the data versus 6.19 hours in model simulations. However, the model somewhat underpredicts study time by URMs (5.87 hours in the data versus 5.23 hours in the model) and overpredicts it among white and Asian students (6.49 hours in the data versus 6.58 hours in the model). The model matches SAT take-up overall, but slightly overpredicts take-up for URMs (by 5 pp). Put another way, by controlling for family income, exam access, and the chance of admission to college, the model can explain two thirds ($\frac{10}{15}$) of the 15 pp gap in SAT take-up by race.

Figure 3 depicts patterns of sorting by knowledge across college tiers, by plotting densities of $\log K_{i,12}$ for students attending each tier. The figure reinforces the patterns of sorting by SAT scores seen in Table A-4, as there is a definite ordering to the peaks of each density. The figure also reveals substantial overlap in the knowledge distribution at all colleges, even between students who attend no college and those who attend elite colleges. In the next section, I evaluate counterfactual policies that may send a new pool of students to college. The overlap of $\log K_{i,12}$ across colleges provides reassurance that the predictions of college completion in the next section have support in the data.

Figure 3: Latent Knowledge, by College Attended



The figure shows the simulated distribution of $\log K_{i,12}$ by college attended. Densities are computed using 200 simulated data sets. SOURCE: (ELS 2002)

7 Counterfactuals

This section simulates a ban on the SAT. Colleges rely on grades and the variables in $\Omega_{i,9}$ when determining whom to admit, while students respond by deciding where to apply and how much to study. The choice of whether to take the SAT is removed from students' choice set. Colleges optimally set admissions thresholds so that the expected demand for university enrollment equals capacity. A second policy mandates all students take the SAT. As before, students respond along the application and study margins, and colleges respond by adjusting thresholds.

7.1 Main Findings

Table 11 shows that banning the SAT causes a small increase in URM enrollment of half a percentage point, while mandating it causes a larger increase of 1.4 pp. These gains are driven by increased URM enrollment at less selective schools (tiers 3 and 6), with URM enrollment at selective and elite schools mostly declining under the ban. Both policies allow the predominantly poor and nonwhite students who do not take the SAT in the status quo to apply to college, but the increased application volume only translates into admissions at less selective schools for two reasons. First, the difference in skills between SAT-takers and non-takers is large (Figure 2), which means that non-takers can only outcompete existing matriculants for admission at schools with low admissions thresholds (Table 8). Second, URM applicants induced to apply by the SAT ban have lower knowledge than similarly-induced white and Asian applicants. After banning the SAT, applications from URM (white and Asian) students increase by 31% (18%). As more students apply, knowledge of the average applicant falls, with the mean of $\log K_{i,12}$ falling by 0.16 sd for URM applicants versus a 0.13 sd decline for white and Asian applicants.

Low-income enrollment fares better in the new admissions environments. The right panel of Table 11 shows that banning the SAT and SAT-for-All both generate similar increases in low-income enrollment of 2.8 pp, with nearly all of the gains again occurring at less selective schools. Low-income enrollment increases by more than URM enrollment, because URM non exam-takers are particularly disadvantaged: The mean of $\log K_{i,12}$ for low-income URM non-takers is 1.10 sd below the sample average, while it is only 0.82 sd below average for low-income white and Asian non-takers. The most knowledgeable non-takers (those with $m_{i,12}$ in the top tercile) are 73% white or Asian.

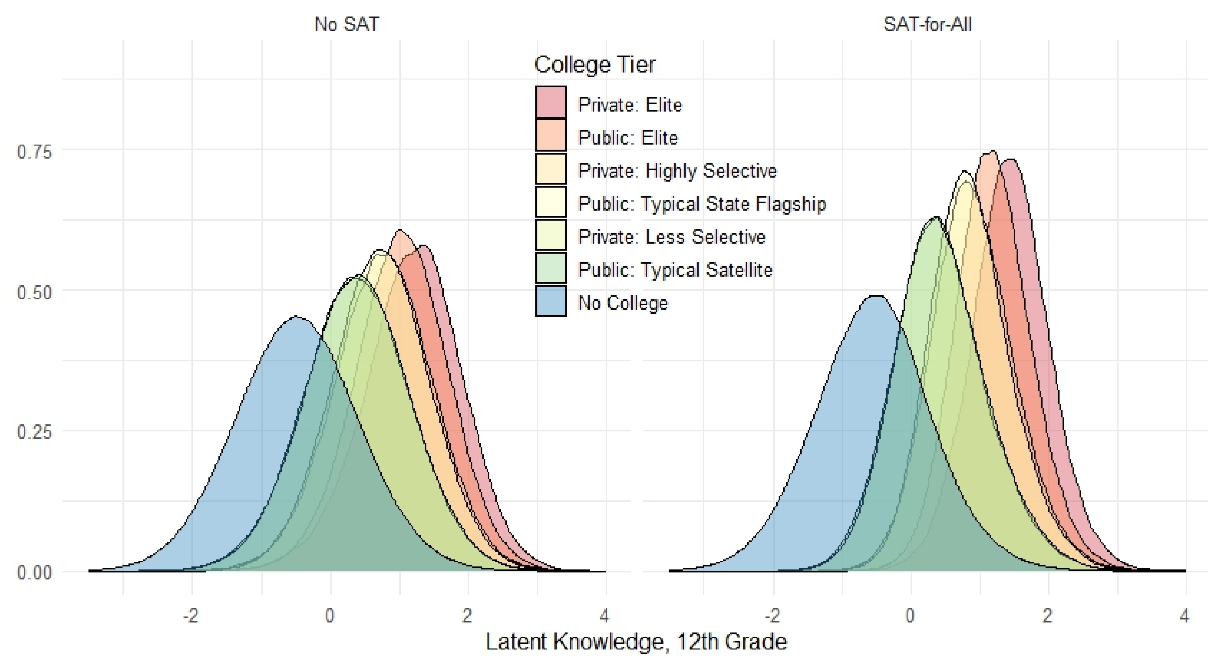
Figure 4 shows that the two policies generate markedly different effects on college sorting by knowledge. Eliminating the SAT renders the distributions of knowledge more homogeneous across colleges, while mandating the SAT has the opposite effect. By giving

Table 11: Access to College

	URM Attendance			Low-Income Attendance		
	Status Quo	No SAT	SAT-for-All	Status Quo	No SAT	SAT-for-All
Tier 1	0.014	0.016	0.014	0.008	0.008	0.008
Tier 2	0.048	0.044	0.044	0.034	0.035	0.033
Tier 3	0.051	0.061	0.062	0.046	0.057	0.058
Tier 4	0.025	0.020	0.024	0.018	0.018	0.020
Tier 5	0.038	0.031	0.033	0.038	0.039	0.037
Tier 6	0.153	0.161	0.166	0.151	0.166	0.169
Any College	0.329	0.334	0.343	0.295	0.323	0.324

The table displays the rates of attendance for URM and low-income students at each college tier under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit their scores with their applications. Low-income refers to students whose families earn less than the median ($52,500 per year). Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

Figure 4: Latent Knowledge Distribution, Counterfactuals



The figure shows the simulated distributions of $\log K_{i,12}$ by college attended under two counterfactual policies. The first eliminates the SAT in college admissions, while the second mandates that every high school student take the SAT and submit their scores when applying to college. Densities are computed using 200 simulated data sets. SOURCE: (ELS 2002)

each college access to two high-quality signals in the SAT math and verbal exams, colleges are better able to draw an inference on each applicant's latent knowledge. This allows

Table 12: Counterfactuals

	Status Quo			No SAT			SAT-for-All		
	All	URM	WA	All	URM	WA	All	URM	WA
Complete College									
Tier 1	0.763	0.705	0.776	0.736	0.650	0.759	0.767	0.700	0.782
Tier 2	0.763	0.701	0.779	0.757	0.694	0.773	0.774	0.721	0.788
Tier 3	0.640	0.565	0.666	0.640	0.568	0.666	0.649	0.582	0.673
Tier 4	0.766	0.730	0.773	0.759	0.714	0.766	0.770	0.733	0.777
Tier 5	0.728	0.683	0.737	0.725	0.668	0.734	0.739	0.702	0.745
Tier 6	0.599	0.510	0.628	0.612	0.521	0.642	0.621	0.540	0.649
All Schools	0.675	0.592	0.698	0.674	0.584	0.699	0.685	0.606	0.708
Yearly Household Income									
Tier 1	110K	99K	113K	109K	95K	112K	108K	95K	111K
Tier 2	96K	83K	99K	94K	83K	97K	94K	80K	97K
Tier 3	74K	61K	79K	71K	58K	76K	71K	58K	76K
Tier 4	103K	91K	105K	102K	92K	104K	100K	88K	102K
Tier 5	92K	80K	94K	90K	80K	92K	90K	78K	92K
Tier 6	72K	57K	77K	70K	56K	74K	70K	55K	74K
No College	52K	41K	58K	54K	42K	60K	54K	42K	61K
Attend Any College	0.437	0.329	0.481	0.449	0.334	0.496	0.444	0.343	0.485
Hours Study	6.19	5.23	6.58	6.30	5.32	6.69	6.30	5.31	6.70
$\mathbb{E}[\log K_{i,12}]$: Any 4-yr College	0.61	0.14	0.74	0.60	0.10	0.73	0.69	0.24	0.82
$\mathbb{E}[\log K_{i,12}]$: No College	-0.47	-0.83	-0.28	-0.49	-0.82	-0.31	-0.55	-0.91	-0.37

The table displays summary statistics under three separate policies: the status quo, a policy that bans the SAT, and a policy in which all students take the SAT and submit their scores with their applications. Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SOURCE: (ELS 2002)

more sought-after schools to better select highly-skilled applicants for admission, leading to greater assortative matching.

Changes in assortative matching cause changes in college completion. Table 12 shows that banning the SAT causes completion rates at tiers 1 and 4 to fall by as much as 2.7 pp. Tier 6 schools instead experience higher graduation rates, as they enroll stronger students who are turned away by more selective schools because of noisier application signals. By contrast, the SAT-for-All policy increases completion at lower-ranked schools without lowering it at elite schools. Rather than drawing students away from elite colleges, the SAT-for-All policy enables all schools, particularly those in tiers 2, 3, 5, and 6, to identify qualified students for admission among those who did not take the SAT and thus did not apply to college in the status quo. College completion rises by 1 pp overall.

Elite private and public colleges (tiers 1 and 4) have the most to lose from banning the SAT. Average knowledge and college completion decline at these schools without generating large increases in URM attendance. The increase in signal variance causes all colleges to inadvertently reject skilled candidates, but less selective colleges are able to enroll students rejected by higher tiers, while elite colleges have no higher tier to draw from. Hence, they experience the largest declines in the skill of their student body.

The mean of knowledge for students attending any four-year college barely budges when banning the SAT. This result may seem surprising: One might expect that the lower signal quality provided by grades would lead to the admission of weaker students. The next section explains that supply side responses by colleges in equilibrium forestall this outcome.[34]

7.2 Model Mechanisms

Eliminating the SAT causes four changes in the model: a shift in the set of measurements used to determine admission, endogenous applications, endogenous study decisions by high school students, and reoptimization by capacity-constrained colleges. To understand the quantitative importance of each of these elements, I simulate the model five times, starting from the status quo and adding one element at a time until I arrive at the full No-SAT counterfactual. Summary statistics for five major variables – URM attendance, knowledge, household income, college completion, and total attendance – under each simulation are presented in Table 13. The second column, labeled No Sat, holds applications and study effort fixed and removes the SAT from among the set of measurements used to determine admission. The third column allows for endogenous applications, while the fourth column adds endogenous study responses. The final column incorporates colleges' supply side responses in equilibrium.

A clear pattern emerges from the analysis. Knowledge and household income become more equalized across colleges in the No SAT and No SAT + Endogenous Applications simulations. The spread of $\log K_{i,12}$ between students at top tier colleges and those not attending college falls by 0.51 sd. The percentage of URMs attending college rises dramatically, from 32.9% to 42.3%. However, the introduction of endogenous effort raises sorting by knowledge and household income. Supply side responses reinforce this stratification and reduce URM enrollment, as marginal entrants are shut out of college by higher ad-

[34] Appendix I considers a ban on the SAT if the exam had been found to be biased against URM students and schools do not account for it. The Kalman Filter subtracts off any bias in the SAT when predicting students' latent knowledge, which has the effect of making bias irrelevant for admissions decisions. This would be problematic if Table 5 had found evidence of bias. The appendix instead assumes that universities do not correct for bias and asks how biased the SAT would need to be for eliminating it to increase URM enrollment by a specified amount.

Table 13: Model Mechanisms

	Status Quo	No Sat	+ Endogenous Apps	+ Endogenous Study	+ Supply Side
URM Attendance					
Tier 1	0.014	0.017	0.021	0.015	0.016
Tier 2	0.048	0.035	0.060	0.059	0.044
Tier 3	0.051	0.055	0.063	0.068	0.061
Tier 4	0.025	0.031	0.034	0.028	0.020
Tier 5	0.038	0.037	0.050	0.045	0.031
Tier 6	0.153	0.157	0.196	0.210	0.161
Any College	0.329	0.331	0.423	0.425	0.334
$\mathbb{E}[\log K_{i,12}]$					
Tier 1	1.41	1.03	0.92	1.20	1.20
Tier 2	0.79	0.70	0.58	0.67	0.77
Tier 3	0.38	0.28	0.27	0.29	0.38
Tier 4	1.15	0.83	0.78	0.99	1.03
Tier 5	0.76	0.61	0.55	0.64	0.73
Tier 6	0.31	0.27	0.20	0.22	0.39
No College	-0.47	-0.37	-0.45	-0.52	-0.49
Yearly HH Income					
Tier 1	110K	108K	101K	109K	109K
Tier 2	96K	92K	88K	91K	94K
Tier 3	74K	73K	69K	69K	71K
Tier 4	103K	93K	94K	100K	102K
Tier 5	92K	87K	85K	87K	90K
Tier 6	72K	70K	66K	66K	70K
No College	52K	55K	54K	53K	54K
Complete College					
Tier 1	0.764	0.718	0.698	0.736	0.735
Tier 2	0.768	0.753	0.738	0.751	0.766
Tier 3	0.640	0.624	0.622	0.627	0.647
Tier 4	0.770	0.736	0.737	0.757	0.763
Tier 5	0.728	0.706	0.700	0.715	0.732
Tier 6	0.598	0.591	0.575	0.581	0.609
Any College	0.674	0.66	0.647	0.653	0.672
Total Attendance					
Tier 1	0.023	0.027	0.029	0.023	0.023
Tier 2	0.068	0.063	0.078	0.078	0.068
Tier 3	0.057	0.052	0.065	0.071	0.057
Tier 4	0.045	0.048	0.054	0.048	0.045
Tier 5	0.067	0.066	0.078	0.078	0.067
Tier 6	0.178	0.172	0.210	0.227	0.178
Any College	0.437	0.428	0.515	0.525	0.437

Each column in the table presents moments from a different simulation. The leftmost column simulates the status quo policy. The second column removes the SAT from admissions but holds applications and study time fixed. The third column lets applications respond endogenously. The fourth column lets study time respond endogenously, and the final column imposes equilibrium in the college market. Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002) missions standards. The last line in the table shows how, in the absence of the capacity constraints imposed in equilibrium, college attendance would be 8.8 pp higher.

The reason why endogenous effort increases assortative matching in partial equilibrium is subtle. When the SAT is eliminated, average study time increases from 6.19 to 6.30 hours a week. This increase in average effort masks two countervailing changes: Students who did not attend college in the status quo increase their study hours from 5.28 to 5.62 hours per week. But, the reduction in signal quality arising from the SAT's elimination causes those formerly attending college in the status quo to reduce their hours worked, from 7.36 to 7.17 hours. Reduction in study hours pulls in the right tail of $\log K_{i,12}$, which explains why enrollment in Tier 1 colleges falls between columns 3 and 4 of Table 13. Lower-skilled and lower-income students originally attending elite public and private colleges narrowly lose out on admission because of the reduction in study effort, and students with higher skills and incomes remain.

When college sorting by knowledge changes, so do rates of college completion. Banning the SAT and allowing for endogenous applications reduce completion at all colleges by up to 6.2 pp. Adding endogenous effort and the supply side instead raises completion. In the full equilibrium without the SAT, elite private colleges have lower rates of completion, because noisier application signals cause them to lose out on some highly-skilled candidates, who then enroll in less selective universities. This, together with the admission of some strong students who do not take the SAT in the status quo, causes graduation rates at public satellite colleges (tier 6) to rise.

The trend moving from left to right in Table 13 of decreasing and then increasing stratification suggests that a lack of income at the application margin does not pose a prohibitive barrier to college access. The model allows the cost of application to vary with family income, but allowing for endogenous applications actually reduces stratification by income. Instead, unequal pre-college human capital investment generates a distribution of cognitive skills that results in students attending markedly different colleges based on their income.[35] Restrictive supply at four-year colleges exacerbates this trend.

7.3 Early Investments in Non SAT-Takers

Figure 2, which documents large skill differences between SAT-takers and non-takers, suggests that the effect of banning the SAT on URM enrollment depends on whether there is a sizable share of strong URM applicants who do not take the SAT but who would

[35] Cunha and Heckman (2007) emphasize how skill differences that open up early in life can be remediated by early and sustained investments. Lochner and Monge-Naranjo (2011) show how credit constraints prevent low-income households from investing in their children as much as they would like.

Table 14: No SAT Policy with Skill Investments in non SAT-Taking URMs

	Status Quo			*No SAT*		
	Attendance	$\mathbb{E}[\log K_{i,12}]$	Complete	Attendance	$\mathbb{E}[\log K_{i,12}]$	Complete
Tier 1	0.014	0.965	0.705	0.018	0.733	0.666
Tier 2	0.048	0.374	0.701	0.053	0.337	0.704
Tier 3	0.051	-0.077	0.565	0.072	-0.040	0.572
Tier 4	0.025	0.804	0.730	0.026	0.732	0.718
Tier 5	0.038	0.472	0.683	0.040	0.426	0.674
Tier 6	0.153	-0.136	0.510	0.190	0.026	0.545
Any College	0.329	0.138	0.592	0.399	0.174	0.601

The table displays attendance rates, average knowledge, and college completion rates for URM students at each college tier under the status quo and No-SAT policies in a hypothetical scenario in which $\log K_{i,9}$ for URM non SAT-takers is raised by 0.705 sd. Details regarding the scenario are provided in the text. Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

apply to college if an SAT score were no longer required. To investigate this, I simulate a policy that eliminates the SAT, but I raise $\log K_{i,9}$ for every URM student who does not take the exam by 0.705 sd so that their average knowledge in the ninth grade is level with typical white and Asian students who take the SAT.[36] Table 14 shows statistics on URM attendance, knowledge, and college completion when banning the SAT under this hypothetical scenario. The fraction of URMs attending any four-year college rises by nearly 7 pp, suggesting that if there were a large population of URM students with high skills who failed to take the SAT, then banning the exam could increase access to college for these students. However, the loss of information caused by banning the SAT in this setting still generates disparate effects across colleges, with elite schools enrolling less skilled students and less selective colleges enrolling stronger students.[37]

8 Conclusion

This paper has shown that when colleges stop using the SAT, they must focus more on other criteria, and this has immediate consequences for who attends college (holding applications fixed) and further affects who applies to college and how well-prepared

[36]While large, 0.705 sd is the average treatment effect on math test scores after one year's attendance in Boston charter schools estimated in Walters (2018).

[37]Appendix I simulates additional policies and economic environments including SAT-Optional admissions, a ban on affirmative action (AA), and an evaluation of an SAT ban if the exam had been found to be biased against URMs. Banning AA leads to a large decrease in URM enrollment of 6.4 pp, but neither an SAT ban nor SAT-for-All can counteract this drop. Both policies, however, can maintain low-income enrollment at pre-AA levels.

they are. Four mechanisms – the shift towards alternative admissions criteria, endogenous applications, endogenous human capital investment, and supply side responses by capacity-constrained colleges – are quantitatively important in shaping patterns of sorting to college. I find that endogenous human capital investment and supply side responses by colleges offset the initial benefit of letting greater numbers of students apply to college, leading to little overall change in URM enrollment after eliminating the SAT.

This paper has documented large skill differences between SAT-takers and non-takers and little evidence that the SAT is biased against URM students. This suggests that the number of non-takers who could out-compete existing SAT-takers for admission when SAT requirements are lifted is limited. Reducing admissions thresholds for disadvantaged applicants would result in greater diversity, although such a move could create real economic costs. Estimates from the model reveal that completion at all schools depends heavily on a student's knowledge at the time of matriculation. Accepting less skilled students may lower rates of completion.

There are several important extensions of this research. This paper has assumed that colleges have preferences over racial diversity and a single latent variable, which I have referred to as knowledge. These preferences can be extended to multiple skills and additional demographic groups. Preliminary evidence from Appendix K suggests that a two-skill model with noncognitive skills generates similar conclusions as the one-skill model regarding the bias and informativeness of the SAT and grades. Still, this does not rule out that other forms of skill that are not well-measured in the ELS 2002 may matter for college admissions and completion.

This paper has estimated universities' preferences for URM students, but it is straightforward to extend the model to include preferences for low-income students or interactions between race and income. This framework will likely prove useful as universities seek to find a way forward after the Supreme Court's decision in June 2023 to ban affirmative action on the basis of race (Students for Fair Admissions, Inc. v. President and Fellows of Harvard College, 600 U. S. 2023). University leaders and policymakers who wish to diversify access to college may wish to follow this paper's example and consider the properties of the measurements that will replace the SAT and how admissions criteria influence the potentially endogenous distribution of skills among likely applicants.

References

Agostinelli, Francesco, and Matthew Wiswall. 2020. "Estimating the Technology of Children's Skill Formation." Technical Report, Working Paper.

Akhtari, Mitra, Natalie Bau, and Jean-William P Laliberté. 2020. "Affirmative Action and Pre-College Human Capital." Technical Report, National Bureau of Economic Research.

Anderson, TW, and Herman Rubin. 1956. "Statistical Inference in Factor Analysis." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability: Held at the Statistical Laboratory, University of California, December, 1954, July and August, 1955*, Volume 1. Univ of California Press, 111.

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5): 1477–1524.

Arcidiacono, Peter, Esteban Aucejo, Patrick Coate, and V Joseph Hotz. 2014. "Affirmative Action and University Fit: Evidence from Proposition 209." *IZA Journal of Labor Economics* 3:1–29.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom. 2016. "College Attrition and the Dynamics of Information Revelation." Technical Report, National Bureau of Economic Research.

Arcidiacono, Peter, Esteban M Aucejo, Hanming Fang, and Kenneth I Spenner. 2011. "Does Affirmative Action Lead to Mismatch? A New Test and Evidence." *Quantitative Economics* 2 (3): 303–333.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2020. "Ex Ante Returns and Occupational Choice." *Journal of Political Economy* 128 (12): 4475–4522.

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2022a. "Asian American Discrimination in Harvard Admissions." *European Economic Review*, vol. 144.

———. 2022b. "Recruit to Reject? Harvard and African American Applicants." *Economics of Education Review*, vol. 88.

———. 2023. "What the Students for Fair Admissions Cases Reveal About Racial Preferences." *Journal of Political Economy Microeconomics* 1, no. 4.

Belasco, Andrew S, Kelly O Rosinger, and James C Hearn. 2015. "The Test-Optional Movement at America's Selective Liberal Arts Colleges: A Boon for Equity or Something Else?" *Educational Evaluation and Policy Analysis* 37 (2): 206–223.

Bennett, Christopher T. 2022. "Untested Admissions: Examining Changes in Application Behaviors and Student Demographics under Test-Optional Policies." *American Educational Research Journal* 59 (1): 180–216.

Bleemer, Zachary. 2022. "Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209." *The Quarterly Journal of Economics* 137 (1): 115–160.

Bleemer, Zachary, and Basit Zafar. 2018. "Intended College Attendance: Evidence from an Experiment on College Returns and Costs." *Journal of Public Economics* 157:184–211.

Bodoh-Creed, Aaron, and Brent Richard Hickman. 2017. "Pre-college Human Capital Investments and Affirmative Action: A Structural Policy Analysis of US College Admissions." *Becker Friedman Institute for Research in Economics Working Paper*, no. 2017-10.

Bond, Timothy N, George Bulman, Xiaoxiao Li, and Jonathan Smith. 2018. "Updating Human Capital Decisions: Evidence from SAT Score Shocks and College Applications." *Journal of Labor Economics* 36 (3): 807–839.

Botelho, Fernando, Ricardo A Madeira, and Marcos A Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7 (4): 37–52.

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): 129–57.

Bowen, William G, and Derek Bok. 2016. *The Shape of the River*. Princeton University Press.

Bulman, George. 2015. "The Effect of Access to College Assessments on Enrollment and Attainment." *American Economic Journal: Applied Economics* 7 (4): 1–36.

Camara, Wayne J, and Amy Elizabeth Schmidt. 1999. "Group Differences in Standardized Testing and Social Stratification. Report No. 99-5." *College Entrance Examination Board*.

Cameron, Stephen V, and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints using Returns to Schooling." *Journal of Political Economy* 112 (1): 132–182.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *The Quarterly Journal of Economics* 134 (3): 1163–1224.

Carneiro, Pedro, Karsten T Hansen, and James J Heckman. 2003. "2001 Lawrence R. Klein lecture: Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review* 44 (2): 361–422.

Carneiro, Pedro, and James J Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *The Economic Journal* 112 (482): 705–734.

Caucutt, Elizabeth M, and Lance Lochner. 2020. "Early and Late Human Capital Investments, Borrowing Constraints, and the Family." *Journal of Political Economy* 128 (3): 1065–1147.

Chan, Jimmy, and Erik Eyster. 2003. "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93 (3): 858–872.

Cotton, Christopher S, Brent R Hickman, and Joseph P Price. 2022. "Affirmative Action and Human Capital Investment: Evidence from a Randomized Field Experiment." *Journal of Labor Economics* 40 (1): 157–185.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

Delavande, Adeline, and Basit Zafar. 2019. "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints." *Journal of Political Economy* 127 (5): 2343–2393.

del Rio, Giulia McDonnell Nieto. 2021. "University of California Will No Longer Consider SAT and ACT Scores." *The New York Times*.

Dillon, Eleanor Wiske, and Jeffrey Andrew Smith. 2017. "Determinants of the Match between Student Ability and College Quality." *Journal of Labor Economics* 35 (1): 45–66.

———. 2020. "The Consequences of Academic Match between Students and Colleges." *Journal of Human Resources* 55 (3): 767–808.

Donni, Olivier. 2015. "Measuring the Cost of Children." *IZA World of Labor*.

Dynarski, Susan M. 2018. "ACT/SAT for All: A Cheap, Effective Way to Narrow Income Gaps in College." *Brookings Institution*.

Epple, Dennis, Richard Romano, and Holger Sieg. 2006. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica* 74 (4): 885–928.

———. 2008. "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* 10 (4): 475–501.

Epstein, Jonathan P. 2009. "Behind the SAT-Optional Movement: Context and Controversy." *Journal of College Admission* 204:8–19.

Freedle, Roy. 2003. "Correcting the SAT's Ethnic and Social-Class Bias: A Method for Reestimating SAT Scores." *Harvard Educational Review* 73 (1): 1–43.

Freeman, Mark, Brian Heseung Kim, Preston Magouirk, and Trent Kajikawa. 2022. "Deadline Update: First-year Application Trends Through March 15." Technical Report, Common App.

Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122 (2): 225–281.

Golightly, Eleanor. 2019. "Does College Access Increase High School Effort? Evaluating the Impact of the Texas Top 10% Rule." Technical Report.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. 2020. "Take Two! SAT Retaking and College Enrollment Gaps." *American Economic Journal: Economic Policy* 12 (2): 115–58.

Goodman, Sarena. 2016. "Learning from the Test: Raising Selective College Enrollment by Providing Information." *Review of Economics and Statistics* 98 (4): 671–684.

Grau, Nicolás. 2018. "The Impact of College Admissions Policies on the Academic Effort of High School Students." *Economics of Education Review* 65:58–92.

Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman. 2015. "The Effects of Earnings Disclosure on College Enrollment Decisions." Technical Report, National Bureau of Economic Research.

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.

Hoxby, Caroline M, and Christopher Avery. 2012. "The Missing "One-offs": The Hidden Supply of High-Achieving, Low Income Students." Technical Report, National Bureau of Economic Research.

Hurwitz, Michael, Jonathan Smith, Sunny Niu, and Jessica Howell. 2015. "The Maine Question: How is 4-year College Enrollment Affected by Mandatory College Entrance Exams?" *Educational Evaluation and Policy Analysis* 37 (1): 138–159.

Hyman, Joshua. 2017. "ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice." *Education Finance and Policy* 12 (3): 281–311.

Kapor, Adam. 2020. "Distributional Effects of Race-Blind Affirmative Action." Technical Report, Working Paper.

Klasik, Daniel. 2013. "The ACT of Enrollment: The College Enrollment Effects of State-Required College Entrance Exam Testing." *Educational Researcher* 42 (3): 151–160.

Leeds, Daniel M, Isaac McFarlin, and Lindsay Daugherty. 2017. "Does Student Effort Respond to Incentives? Evidence from a Guaranteed College Admissions Program." *Research in Higher Education* 58 (3): 231–243.

Lochner, Lance J, and Alexander Monge-Naranjo. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Review* 101 (6): 2487–2529.

Otero, Sebastián, Nano Barahona, and Cauê Dobbin. 2021. "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil." Technical Report, Working paper.

Pallais, Amanda. 2015. "Small Differences that Matter: Mistakes in Applying to College." *Journal of Labor Economics* 33 (2): 493–520.

Putnick, Diane L, and Marc H Bornstein. 2016. "Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research." *Developmental Review* 41:71–90.

Rauschenberg, Samuel. 2014. "How Consistent Are Course Grades? An Examination of Differential Grading." *Education Policy Analysis Archives* 22, no. 92.

Rosinger, Kelly Ochs, and Karly S Ford. 2019. "Pell Grant Versus Income Data in Postsecondary Research." *Educational Researcher* 48 (5): 309–315.

Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135–146.

Saboe, Matt, and Sabrina Terrizzi. 2019. "SAT Optional Policies: Do They Influence Graduate Quality, Selectivity or Diversity?" *Economics Letters* 174:13–17.

Saltiel, Fernando. 2023. "Multi-Dimensional Skills and Gender Differences in Stem Majors." *The Economic Journal* 133 (651): 1217–1247.

Schmill, Stu. 2022. "We Are Reinstating Our SAT/ACT Requirement for Future Admissions Cycles." https://mitadmissions.org/blogs/entry/we-are-reinstating-our-sat-act-requirement-for-future-admissions-cycles/. Accessed March 28, 2022.

Soares, Joseph A. 2020. "The Scandal of Standardized Tests: Why We Need to Drop the SAT and ACT." *Teachers College Press*.

Stinebrickner, Ralph, and Todd R Stinebrickner. 2004. "Time-Use and College Outcomes." *Journal of Econometrics* 121 (1-2): 243–269.

———. 2008. "The Causal Effect of Studying on Academic Performance." *The BE Journal of Economic Analysis & Policy* 8, no. 1.

Stinebrickner, Todd, and Ralph Stinebrickner. 2012. "Learning about Academic Ability and the College Dropout Decision." *Journal of Labor Economics* 30 (4): 707–748.

Sweitzer, Kyle, A Emiko Blalock, and Dhruv B Sharma. 2018. "The Effect of Going Test-Optional on Diversity and Admissions." *Measuring Success: Testing, Grades, and the Future of College Admissions*, pp. 288–307.

Tincani, Michela M, Fabian Kosse, and Enrico Miglino. 2021. "Subjective Beliefs and Inclusion Policies: Evidence from College Admissions." Technical Report, Working Paper.

Van den Bergh, Linda, Eddie Denessen, Lisette Hornstra, Marinus Voeten, and Rob W Holland. 2010. "The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap." *American Educational Research Journal* 47 (2): 497–527.

Walters, Christopher R. 2018. "The Demand for Effective Charter Schools." *Journal of Political Economy* 126 (6): 2179–2223.

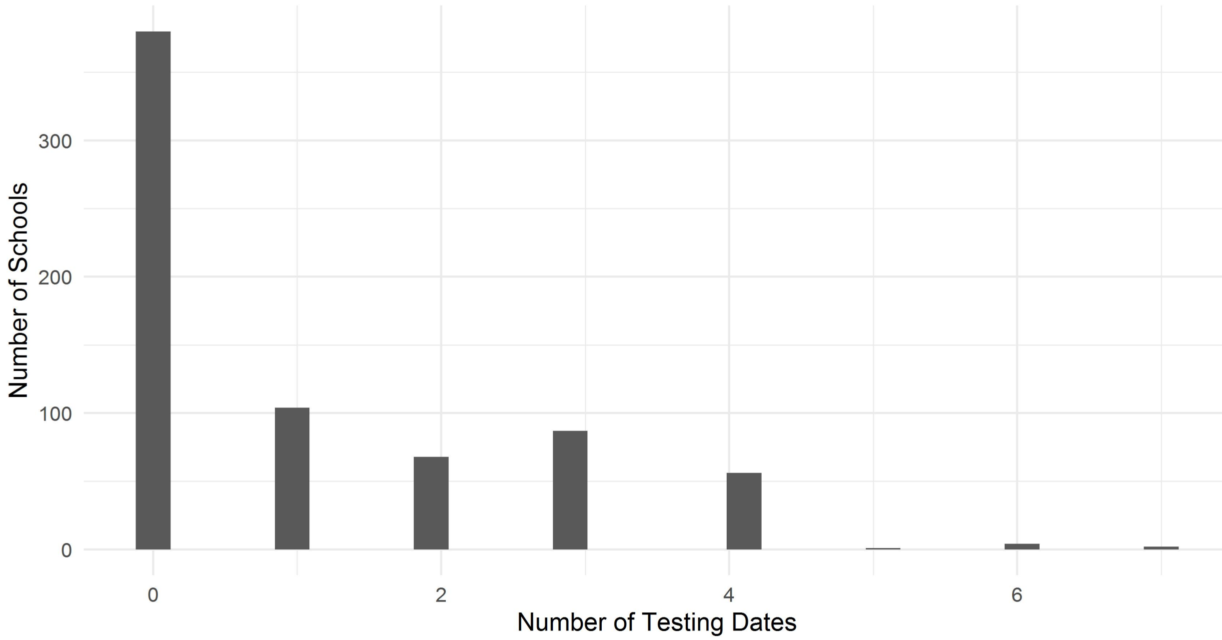
Williams, Benjamin. 2020. "Identification of the Linear Factor Model." *Econometric Reviews* 39 (1): 92–109.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2): 791–824.

Appendices

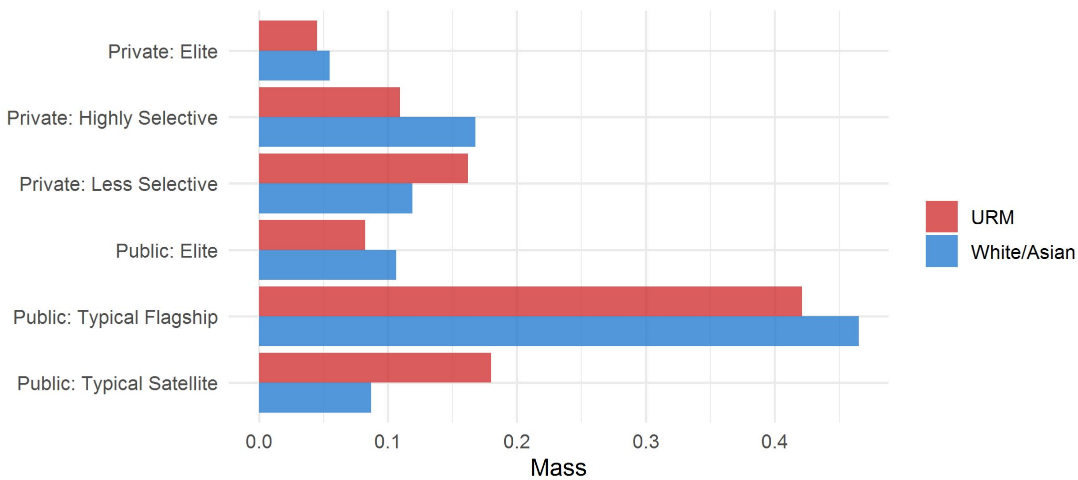
A Additional Tables and Figures

Figure A-1: SAT and ACT Testing Dates per School



The figure shows the number of dates reserved for SAT or ACT examinations during the spring of 2003 at schools in the ELS 2002. SOURCES: College Board, ACT, Inc., and U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

Figure A-2: Type of College Attended among College-Goers



The figure shows the fraction of all college students of each demographic group that are enrolled at each college tier. The grouping of colleges into tiers is described in section 3.7. SOURCE: (ELS 2002)

Table A-1: Effect of Testing Availability on Exam-Taking and Applications

	Take SAT			*Apply to College*		
Testing Dates	0.055 (0.016)	0.056 (0.016)	0.042 (0.017)	0.057 (0.015)	0.059 (0.015)	0.047 (0.015)
School Type	Yes	Yes	Yes	Yes	Yes	Yes
School Geography	Yes	Yes	Yes	Yes	Yes	Yes
School Size	Yes	Yes	Yes	Yes	Yes	Yes
Local Poverty Rate	No	Yes	Yes	No	Yes	Yes
Mother's Education	No	No	Yes	No	No	Yes
N	9910	9910	9910	9910	9910	9910

The table displays the results of logit regressions of binary indicators for taking the SAT and applying to college on the number of SAT testing dates in one's own school during spring of the junior year of high school. Controls for the type of school, school geography (urban, rural, or suburban), enrollment, district level poverty rates, and mother's education (5 categories) are also included. SOURCE: (ELS 2002)

Table A-2: Effect of Distance on Applications

	Apply to College					
		Private			*Public*	
	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5	Tier 6
Distance : 50-150 miles	-0.0121***	-0.0156***	-0.0122***	-0.0479***	-0.0686***	-0.0645***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039	(0.0020)
Distance : 150 - 250 miles	-0.0155***	-0.0203***	-0.0145***	-0.0735***	-0.0942***	-0.0773***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039)	(0.0020)
Distance : 250 - 500 miles	-0.0179***	-0.0221***	-0.0152***	-0.0800***	-0.1034***	-0.0807***
	(0.0016)	(0.0009)	(0.0006)	(0.0042)	(0.0039	(0.0020)
Distance : over 500 miles	-0.0199***	-0.0230***	-0.0153***	-0.0893***	-0.1056***	-0.0817***
	(0.0016)	(0.0009)	(0.0006)	(0.0043)	(0.0039)	(0.0020)
N	9910	9910	9910	9910	9910	9910
Num Schools	60	270	480	40	80	380
Individual FEs	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SEs	Individual	Individual	Individual	Individual	Individual	Individual

The table shows the coefficients from linear probability models of application decisions on distance to college. The unit of observation is an individual-college. All four-year public and private non-profit colleges are included. Distance is calculated from the centroid of the individual's home census block while in high school to the latitude and longitude of each college in IPEDs. Each regression includes individual fixed effects and controls for tuition, acceptance rates, and interactions between acceptance rates and an individual's performance on the 12th grade NCES math exam. SOURCE: (ELS 2002)

Table A-3: Proportion of Variance Due to Signal

	URM	White & Asian	Difference
GPA 9	0.285	0.417	-0.132
GPA 10	0.267	0.397	-0.130
GPA 11	0.211	0.319	-0.108
GPA 12	0.134	0.202	-0.068
SAT Math	0.687	0.735	-0.048
SAT Verbal	0.546	0.538	0.008
NCES10 Read	0.495	0.464	0.031
NCES10 Math	0.818	0.739	0.079
NCES12 Math	0.803	0.758	0.045

The table displays estimates of the proportion of variance of each measurement due to the signal for URM and white and Asian students. For each individual, the proportion of variance coming from signal j is given by $\frac{{\alpha_{t,j}^{R}}^2 var(\log K_{i,t})}{{\alpha_{t,j}^{R}}^2 var(\log K_{i,t}) + {\sigma_{t,j}^{R}}^2}$ for $R = URM, WA$. The individual-specific proportions are then averaged for the URM and WA populations to create the numbers in the table. SOURCE: (ELS 2002)

Table A-4: Model Fit

	Data			Model		
	All	URM	WA	All	URM	WA
Fraction Attending						
Tier 1 School	0.023	0.015	0.027	0.023	0.014	0.026
Tier 2 School	0.069	0.036	0.082	0.068	0.048	0.075
Tier 3 School	0.057	0.053	0.058	0.057	0.051	0.059
Tier 4 School	0.045	0.027	0.052	0.045	0.025	0.053
Tier 5 School	0.069	0.035	0.083	0.067	0.038	0.079
Tier 6 School	0.178	0.161	0.185	0.178	0.153	0.188
Any College	0.441	0.327	0.487	0.437	0.329	0.481
SAT Math Score						
Tier 1	1.30	0.83	1.41	1.11	0.57	1.24
Tier 2	0.51	0.21	0.56	0.44	-0.01	0.56
Tier 3	-0.12	-0.60	0.05	0.03	-0.43	0.19
Tier 4	0.81	0.40	0.90	0.84	0.41	0.92
Tier 5	0.42	0.09	0.48	0.42	0.09	0.48
Tier 6	-0.10	-0.51	0.05	-0.04	-0.47	0.10
No College	-0.54	-0.92	-0.38	-0.87	-1.19	-0.72
SAT Verbal Score						
Tier 1	1.34	0.96	1.42	1.01	0.54	1.12
Tier 2	0.62	0.41	0.66	0.42	0.01	0.52
Tier 3	-0.05	-0.42	0.09	0.04	-0.38	0.19
Tier 4	0.74	0.35	0.83	0.77	0.41	0.84
Tier 5	0.38	0.04	0.43	0.39	0.11	0.44
Tier 6	-0.09	-0.48	0.05	-0.02	-0.42	0.11
No College	-0.51	-0.86	-0.37	-0.77	-1.09	-0.62
Yearly Household Income						
Tier 1	107	78	114	110	99	113
Tier 2	95	82	97	96	83	99
Tier 3	73	55	79	74	61	79
Tier 4	98	87	100	103	91	105
Tier 5	88	67	92	92	80	94
Tier 6	70	57	75	72	57	77
No College	54	43	60	52	41	58
Complete College	0.676	0.587	0.699	0.675	0.592	0.698
Hours Study	6.31	5.87	6.49	6.19	5.23	6.58
Take SAT	0.75	0.64	0.79	0.76	0.69	0.79

The table compares moments in the data with their model counterparts by simulating the model according to the estimated parameters. Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SAT scores are normalized by the sample mean and standard deviation in the data. Households income is measured in thousands of dollars. SOURCE: (ELS 2002)

Table A-5: URM Access to College by Income

	Status Quo		No SAT		SAT-for-All	
	Low Inc	High Inc	Low Inc	High Inc	Low Inc	High Inc
Tier 1	0.003	0.011	0.004	0.012	0.004	0.011
Tier 2	0.016	0.032	0.015	0.029	0.016	0.028
Tier 3	0.026	0.025	0.033	0.028	0.034	0.028
Tier 4	0.007	0.018	0.006	0.015	0.007	0.017
Tier 5	0.014	0.024	0.011	0.019	0.013	0.020
Tier 6	0.084	0.069	0.091	0.070	0.095	0.071
Any College	0.150	0.179	0.160	0.174	0.168	0.175

The table displays the rates of attendance for low-income and high-income URM students at each college tier under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit their scores with their applications. Low-income refers to students whose families earn less than the median ($52,500 per year). High-income families earn more than the median. Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

Table A-6: White and Asian Access to College by Income

	Status Quo		No SAT		SAT-for-All	
	Low Inc	High Inc	Low Inc	High Inc	Low Inc	High Inc
Tier 1	0.004	0.023	0.004	0.021	0.004	0.022
Tier 2	0.016	0.059	0.017	0.058	0.016	0.053
Tier 3	0.020	0.039	0.025	0.044	0.024	0.044
Tier 4	0.009	0.044	0.010	0.043	0.010	0.042
Tier 5	0.019	0.059	0.021	0.059	0.019	0.055
Tier 6	0.067	0.122	0.073	0.121	0.074	0.122
Any College	0.135	0.346	0.150	0.346	0.147	0.338

The table displays the rates of attendance for low-income and high-income white and Asian students at each college tier under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit their scores with their applications. Low-income refers to students whose families earn less than the median ($52,500 per year). High-income families earn more than the median. Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

B Sample Selection Criteria

Of the 16,200 students who were initially sampled in the ELS 2002, 12,880 responded in both the base year and the first follow-up survey in 2004. The most common reasons for exclusion are nonresponse (12.05% of the sample), dropping out of high school between 2002 and 2004 (3.76% of the sample), and graduating early (2.43% of the sample). Other, infrequent, reasons include being out of the country, language difficulties making the survey impossible, and death. 370 students lack information on the amount of time they spend studying. A further 2,370 students either lack information on GPA, school characteristics, or geocode data. 610 individuals lack an SAT score despite applying to colleges that required the exam. Finally, I exclude a small number of individuals (< 10) with extremely low grades and SAT scores who are admitted to elite colleges, possibly because of athletics. These students cause the model likelihood function to return infinite values for large regions of the parameter space. These exclusions result in a sample of 9,910 observations.

During the second follow-up survey in 2006, students report the full list of colleges they applied to, where they were admitted, and where they first matriculated. To address concerns regarding whether use of this self-reported measure may result in biased admissions probabilities, I compare admission probabilities derived from the survey responses in the ELS 2002 to acceptance rates in the Integrated Postsecondary Education Data System (IPEDS) for the same year (2004/05 cohort) in Table B-1. The table disaggregates admission probabilities by type (private non-profit vs public) and Barron's Selectivity Ranking. The IPEDs statistics are weighted by the enrollment of the institution. The table shows that admissions rates are somewhat lower in IPEDS than in the ELS 2002. The differences, of eleven to thirteen percentage points (pp) for private colleges, and between six and eight pp for public colleges, indicate that students in the ELS 2002 selectively under-report applications to colleges at which they are rejected. The implication for the empirical results is that the cost of applying to college, which is identified by the number of applications to college, could be over-estimated in the model. This suggests that model forecasts of application growth resulting from policy changes may be a lower bound for the true application growth.

C Initial Conditions and Kalman Filter

The first observed measure of knowledge for students in the ELS 2002 is in the ninth grade, and it is unlikely that all students begin high school on a level playing field. Pro-

Table B-1: Admission Rates, IPEDS and ELS 2002

			Rate of Admission	
Type	Tier	Barron's Rank	IPEDS	ELS 2002
Private	1	1	0.29	0.40
	2	2 and 3	0.66	0.79
	3	4, 5 and 6	0.72	0.84
Public	4	1 and 2	0.53	0.61
	5	3	0.69	0.75
	6	4, 5, and 6	0.72	0.80

The table compares rates of admission at six college tiers based on self-reported application and admission information in the ELS 2002 together with data from IPEDs for the same set of schools. IPEDS admission rates for each tier are weighted by the enrollment at each school within the tier. SOURCES: IPEDS and National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002), "Base Year through Second Follow-up, 2002-2006."

vided that the skill accumulation equation in (1) holds in middle and elementary school, one can use backwards substitution to write knowledge in the ninth grade as a function of an entire history of study decisions and educational inputs as follows:

$$\log K_{i,9} = f(h_{i,1}, \ldots, h_{i,9}, \mathbf{I}_{i,1}, \ldots, \mathbf{I}_{i,9}, K_{i,0}) \ .$$

This equation means that even if children were born with equal endowments, $K_{i,0}$, a history of unequal investments would generate differences in the distribution of $K_{i,9}$. While the ELS 2002 does not record the entire history of inputs and study decisions prior to high school, I *do* allow the distribution of $K_{i,9}$ to vary according to a set of predetermined covariates, $\mathbf{W}_i$, that are likely to be correlated with prior investments. Hence, rather than imposing the normalization that $\log K_{i,9} \sim N(0, \sigma_k^2)$, as is common in the literature estimating dynamic factor models, I instead allow the distribution of initial conditions to vary by $\mathbf{W}_i$ as follows:

$$\begin{pmatrix} \log K_{i,9} \\ \mathbf{y}_{i,9}^R \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{W}_i'\mathbf{a} \\ \boldsymbol{\mu}_9^R + \boldsymbol{\alpha}_9^R \mathbf{W}_i'\mathbf{a} \end{pmatrix}, \begin{pmatrix} \sigma_k^2(\mathbf{W}_i) & \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'} \\ \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R & \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R\boldsymbol{\alpha}_9^{R'} + \boldsymbol{\Sigma}_9^R \end{pmatrix}\right) \ ,$$

where the variance of ninth grade knowledge is given by $\sigma_k^2(\mathbf{W}_i) = \exp(\mathbf{W}_i'\mathbf{b})$ and $\boldsymbol{\Sigma}_9^R = \mathbb{E}(\boldsymbol{\varepsilon}_{i,9}^R \boldsymbol{\varepsilon}_{i,9}^{R'})$.

As in the main text, let $URM_i \in \{0, 1\}$ indicate whether a student belongs to an underrepresented minority, and define the initial information set by

$$\Omega_{i,9} := \{URM_i, \mathbf{W}_i, \mathbf{y}_{i,9}, \{\mathbf{I}_{i,k}\}_{k=10}^{12}\} \ ,$$

and subsequent updates as $\Omega_{i,t} := \{\Omega_{i,t-1}, \mathbf{y}_{i,t}, h_{i,t}\}$. The Kalman Filter yields the following update for knowledge after observing the initial conditions and ninth grade GPA:

$$\log K_{i,9} \mid \Omega_{i,9} \sim N(m_{i,9}, P_{i,9}) \ ,$$

where

$$\begin{aligned}
m_{i,9} &:= \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'}\mathbf{F}_{i,9}^{-1}\left(\mathbf{y}_{i,9} - (\boldsymbol{\mu}_9^R + \boldsymbol{\alpha}_9^R\mathbf{W}_i'\mathbf{a})\right) \ , \\
P_{i,9} &:= \sigma_k^2(\mathbf{W}_i) - \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^{R'}\mathbf{F}_{i,9}^{-1}\boldsymbol{\alpha}_9^R\sigma_k^2(\mathbf{W}_i) \ , \\
F_{i,9} &:= \sigma_k^2(\mathbf{W}_i)\boldsymbol{\alpha}_9^R\boldsymbol{\alpha}_9^{R'} + \boldsymbol{\Sigma}_9^R \ .
\end{aligned}$$

For an individual with information set $\Omega_{i,t-1}$ who chooses to study $h_{i,t}$ hours, the prediction for period t knowledge is

$$\begin{aligned}
\log K_{i,t} \mid \Omega_{i,t-1}, h_{i,t} &\sim N(m_{i,t|t-1}, P_{i,t|t-1}) \ , \\
m_{i,t|t-1} &:= \gamma^{K,R}m_{i,t-1} + \beta^{H,R}h_{i,t} + \mathbf{I}_{i,t}'\boldsymbol{\beta}^{I,R} \ , \\
P_{i,t|t-1} &:= \gamma^{K,R^2}P_{i,t-1} \ ,
\end{aligned}$$

and the subsequent update for period t knowledge is

$$\begin{aligned}
\log K_{i,t} \mid \Omega_{i,t} &\sim N(m_{i,t}, P_{i,t}) \ , \\
m_{i,t} &:= P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'}\mathbf{F}_{i,t}^{-1}\left(\mathbf{y}_{i,t} - (\boldsymbol{\mu}_t^R + \boldsymbol{\alpha}_t^R m_{i,t|t-1})\right) \ , \\
P_{i,t} &:= P_{i,t|t-1} - P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'}\mathbf{F}_{i,t}^{-1}\boldsymbol{\alpha}_t^R P_{i,t|t-1} \ , \\
\mathbf{F}_{i,t} &:= \boldsymbol{\alpha}_t^R P_{i,t|t-1}\boldsymbol{\alpha}_t^{R'} + \boldsymbol{\Sigma}_t^R \ .
\end{aligned}$$

D Financial Aid

The model allows financial aid to shape matriculation through its contribution to Net Tuition in equation (8). Any study of college attendance must deal with the fact that aid is unobserved at schools to which the student does not apply. Another challenge is that the ELS 2002 contains data on federal financial aid, but not state or institutional aid.

I address this problem of partially observed data by training a random forest on data from a more recent NCES educational survey, the High School Longitudinal Study of 2009 (HSLS 2009). The HSLS 2009 is a longitudinal survey of the transition from high school to college with many of the same measurements as the ELS 2002 (GPA, SAT scores, college attended), but, unlike the ELS 2002, the HSLS 2009 records federal, state, and institutional financial aid awarded to each student.[38] I train a random forest on the HSLS 2009 using five-fold cross validation to predict the proportion of tuition covered by financial aid at each college tier. I then develop a crosswalk (available upon request) between the two data sets and use the random forest to predict the proportion of tuition that would be covered by financial aid for each individual in the ELS 2002 at each college tier. I work with proportions rather than aid dollars to control for the growth in college tuition between the 2002 and 2009 cohorts. $Aid_{i,c}$ is therefore the predicted proportion of tuition at school c that would be covered by financial aid awarded to individual i. I then compute predicted aid *dollars* for students in the ELS 2002 by multiplying the predicted aid proportion by the dollar amount of tuition for their reference college in 2004, when these students are matriculating to college. Net tuition is just the difference between posted tuition and predicted aid dollars: $NetTuition_{i,c} = Tuition_{i,c}^{2004}(1 - Aid_{i,c})$. Net tuition is measured in hundreds of dollars per week (divide by 52×100).[39]

The random forest uses 1000 trees, samples six variables at each node, and sets the minimum node size to six. These parameters were selected to optimize the out-of-sample predictions across the six prediction models (one for each tier of colleges). The out-of-bag R^2 for these prediction models are 0.306, 0.151, 0.095, 0.195, 0.10, and 0.071.

E Knowledge Predicts Admission

This paper has modeled each student's probability of admission to college as a function of her demographic and her knowledge at the time of application, $\log K_{i,12}$, as filtered through the observable measurements in $\Omega_{i,12}$. This section shows that knowledge filtered in this way is highly predictive of admission.

Figures E-1 through E-6 provide estimates of nonparametric local linear regressions of admission to college as a function of the mean of each student's twelfth grade knowledge as derived from the model, $\mathbb{E}[\log K_{i,12} \mid \Omega_{i,12}]$. The regressions are all estimated using an Epanechnikov kernel and a bandwidth of one. The nonparametric functions are plotted only over the support of $\mathbb{E}[\log K_{i,12} \mid \Omega_{i,12}]$ among applicants to each school in the ELS 2002. The figures show that admission probabilities are increasing in twelfth grade knowledge at all schools. While many colleges observe characteristics that are not in the ELS 2002, such as writing samples and teacher recommendations, the nonparametric re-

[38]The main disadvantage of the HSLS 2009 relative to the ELS 2002 and the reason it was not used for this study is that it does not record each student's entire admissions and acceptance portfolios.

[39]The model gives students perfect foresight over financial aid. If students instead had to form expectations over aid, their applications would likely appear more random and less targeted to their preferred school.

gressions provide reassurance that the measurements observed in the data are still highly predictive of admission.

Figure E-1: Probability of Admission, Elite Private Colleges

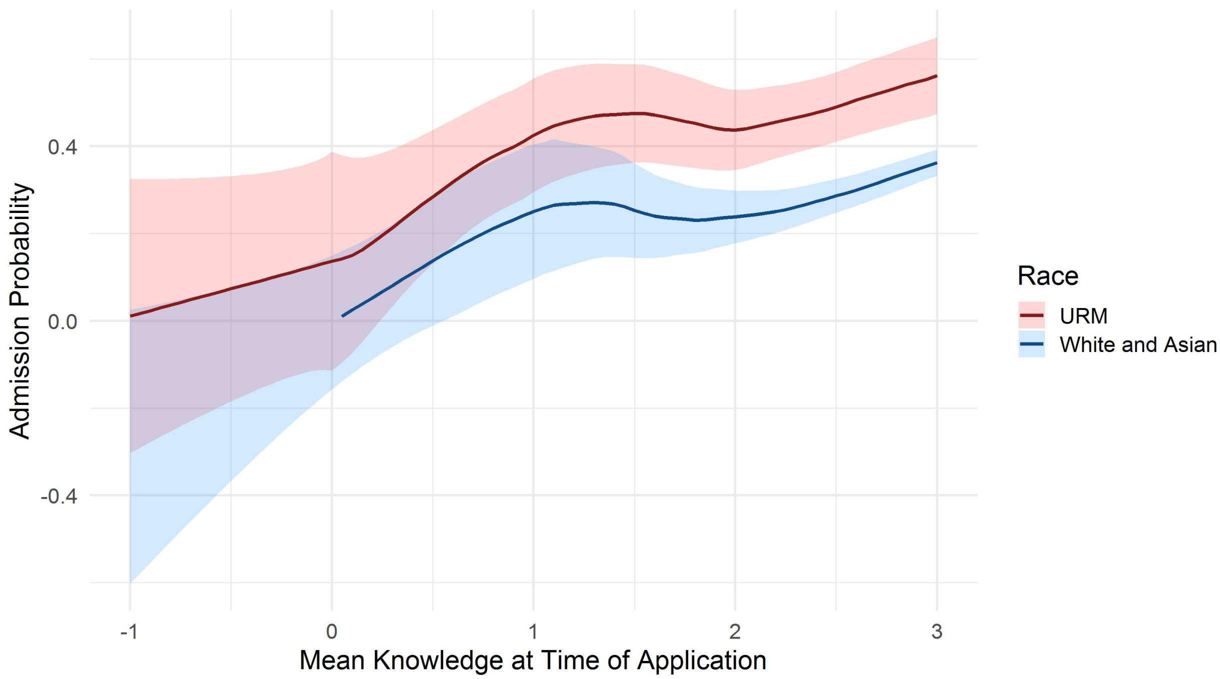


Figure E-2: Probability of Admission, Highly Selective Private Colleges

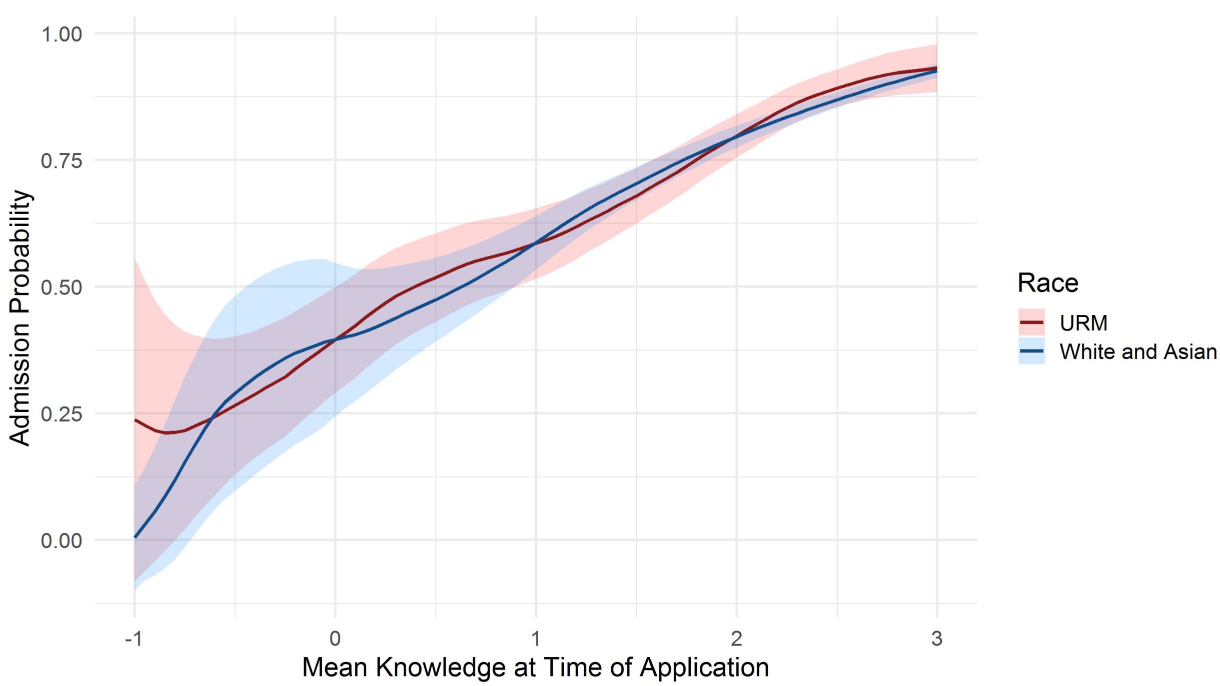


Figure E-3: Probability of Admission, Less Selective Private Colleges

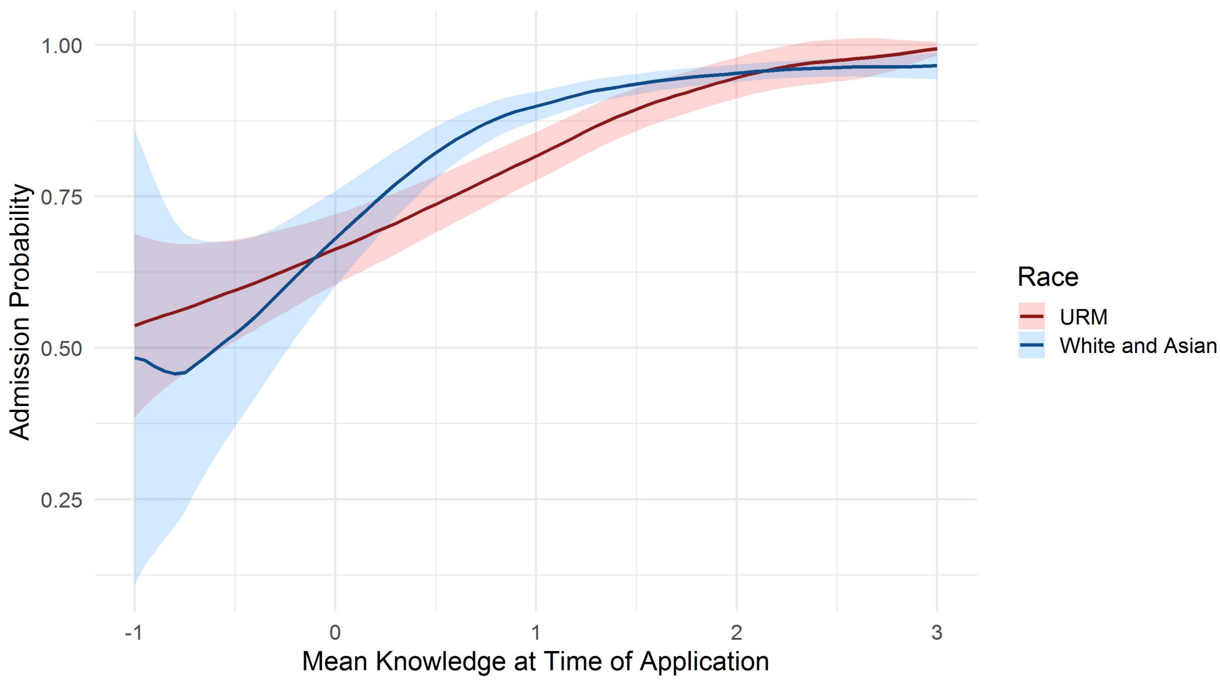


Figure E-4: Probability of Admission, Elite Public Colleges

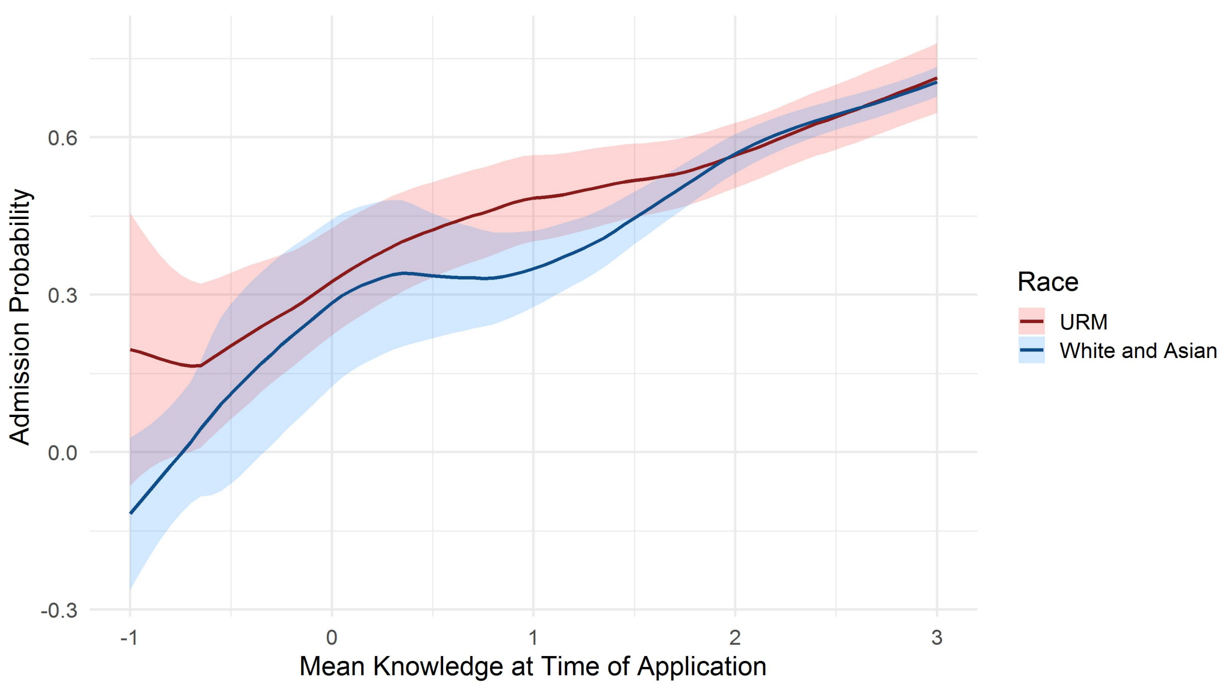


Figure E-5: Probability of Admission, Typical State Flagships

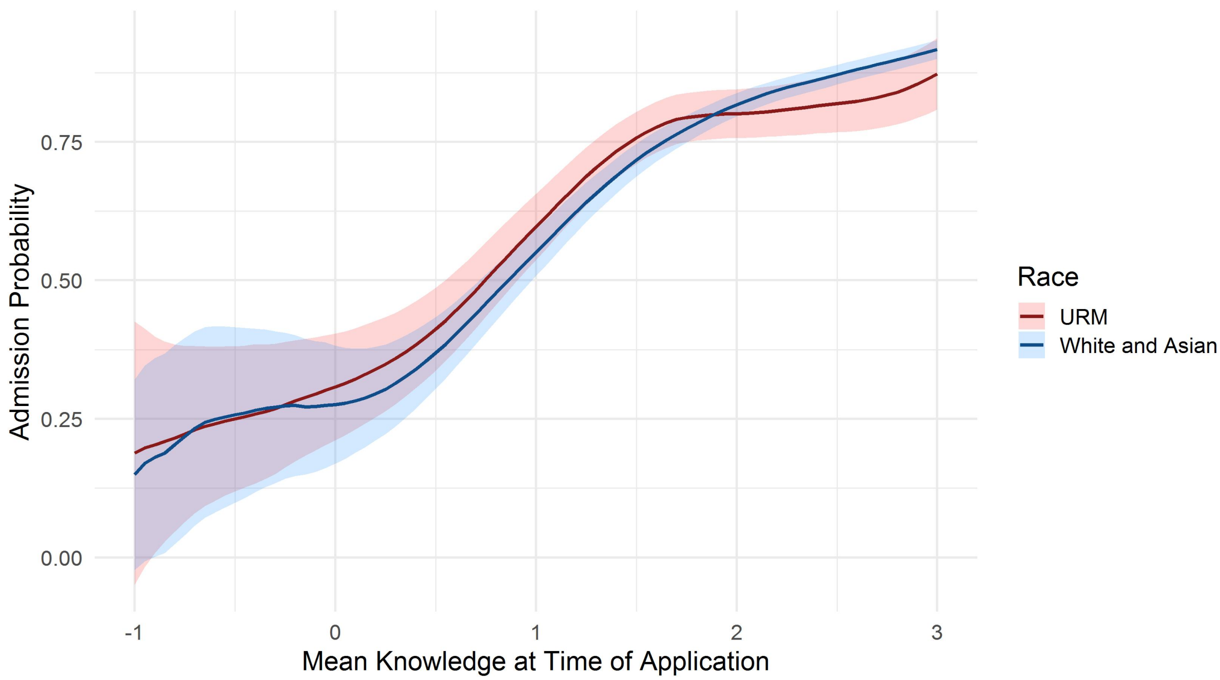
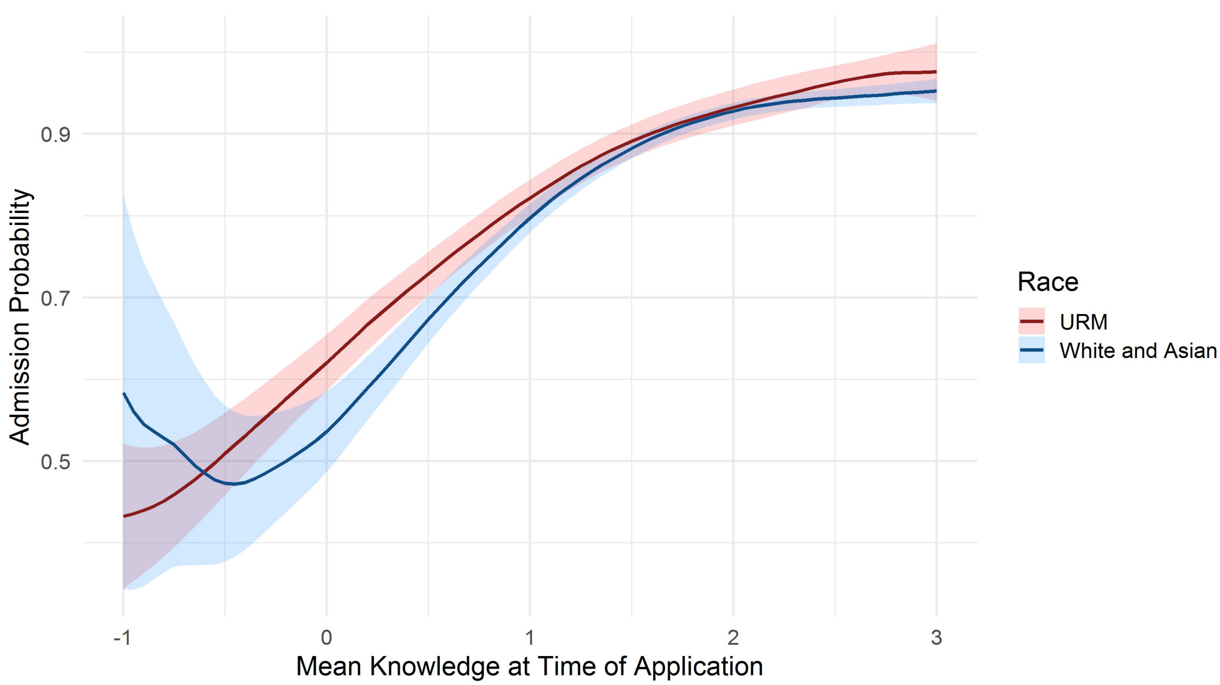


Figure E-6: Probability of Admission, Typical State Satellites



F Alternative Normalizations and Specifications for the Dynamic Factor Model

The identification of dynamic factor models requires normalizations. In this paper, I exploit the fact that the tenth and twelfth grade NCES math exams are scored on the same vertical scale to reduce the number of necessary normalizations. However, as I explain in section 4.1, since both the parameters of the measurement system and the mean of the initial conditions vary by whether a student belongs to an under-represented minority, an additional normalization is required. The approach I adopt in this paper is to impose that the NCES math exams have the same constants in equation (3), namely that $\mu_{10,j}^{URM} = \mu_{10,j}^{WA}$ and $\mu_{12,j}^{URM} = \mu_{12,j}^{WA}$ for j equal to the NCES math exam. In this section, I explore whether the inferences drawn from the dynamic factor model are robust to alternative normalizations and alternative specifications for the technology of skill formation.

Table F-1 presents estimates from a dynamic factor model that instead imposes the normalization that GPA in the ninth grade has the same constant for URM students as it does for white and Asian students, $\mu_{9,j}^{URM} = \mu_{9,j}^{WA}$ for j equal to the ninth grade GPA. The NCES math exams are now permitted to have different constants by URM status. The estimates of bias in the left panel are qualitatively and quantitavely very similar to the main specification. Despite using a different normalization, I estimate that neither of the SAT exams, nor any of the NCES exams are biased. Similar to the main specification, the only evidence of bias is in GPA in the twelfth grade, which appears to be biased against URMs. The estimates of signal-to-noise ratios in the right panel of Table F-1 are qualitatively similar to those from the main specification. I estimate that GPAs are less informative for URMs during each year of high school, and I find that the standardized exams are typically more informative for URMs than for white and Asian students.

The model in the main part of the paper specificies a deterministic skill technology (equation 1). I now explore whether the results are robust to the inclusion of a shock in this equation. Because of the inclusion of the shock, the identification argument in section 4.1 breaks down, and additional normalizations are needed. Since there is only one measurement in grade nine, it is not possible to separately identify the variance of the technology shock, the variance of the measurement shock, and the factor loading, so I normalize the factor loading on GPA in the ninth grade to equal the factor loading on GPA in the tenth grade, $\alpha_{9,j}^{R} = \alpha_{10,j}^{R}$ for $j = GPA$ and $R = URM, WA$. For the same reason, I normalize the factor loadings on eleventh and twelfth grade GPAs to be the same, $\alpha_{11,j}^{R} = \alpha_{12,j}^{R}$, again for $j = GPA$ and $R = URM, WA$.

Table F-2 presents estimates from this dynamic factor model. The estimated variance

Table F-1: Bias and Signal-to-Noise Ratios

	$\boldsymbol{\mu}_t^R$			$\frac{\alpha_{t,j}}{\sigma_{t,j}}$		
	URM	WA	Difference	URM	WA	Difference
GPA, 9th grade	-0.14	-0.14	0	0.75	0.90	-0.15
	(0.02)	(0.02)	(-)	(0.03)	(0.03)	(0.03)
GPA, 10th grade	-0.26	-0.26	0.00	0.70	0.86	-0.16
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
GPA, 11th grade 11	-0.38	-0.36	-0.02	0.60	0.73	-0.13
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
GPA, 12th grade 12	-0.49	-0.37	-0.12	0.46	0.53	-0.08
	(0.03)	(0.02)	(0.05)	(0.01)	(0.02)	(0.02)
SAT Math	-0.91	-1.02	0.11	1.72	1.78	-0.06
	(0.04)	(0.03)	(0.03)	(0.08)	(0.06)	(0.07)
SAT Verbal	-0.83	-0.89	0.06	1.28	1.15	0.13
	(0.04)	(0.03)	(0.03)	(0.08)	(0.04)	(0.07)
NCES Reading, 10th grade	-0.28	-0.31	0.03	1.15	0.99	0.16
	(0.04)	(0.03)	(0.03)	(0.05)	(0.03)	(0.04)
NCES Math, 10th grade	-0.33	-0.39	0.06	2.46	1.80	0.66
	(0.04)	(0.03)	(0.03)	(0.13)	(0.06)	(0.01)
NCES Math, 12th grade	-0.33	-0.39	0.06	2.34	1.89	0.45
	(0.04)	(0.03)	(0.03)	(0.09)	(0.06)	(0.06)

The table displays estimates of bias (in the left panel) and signal-to-noise ratios (in the right panel) when ninth grade GPA is normalized to have no bias. These correspond to $\boldsymbol{\mu}_t^R$ and $\frac{\alpha_{t,j}}{\sigma_{t,j}}$ in equation (3). SOURCE: (ELS 2002)

of the shock is not statistically different from zero, and the rest of the parameters are nearly identical to the main specification in the paper (Table 6). The results presented here should alleviate concern that the use of a deterministic skill formation equation, which aids in identification, significantly biases the findings in the main text.

G Solving the Model

I use Bayes' rule to compute diversity at each school:

$$\begin{aligned} \lambda^c_{URM} :=& \mathbb{P}(URM_i = 1 \mid Attend_{i,c} = 1) , \\ =& \frac{\mathbb{P}(Attend_{i,c} = 1 \mid URM_i = 1)\lambda_{URM}}{\mathbb{P}(Attend_{i,c} = 1)} , \end{aligned}$$

where λ_{URM} is the population fraction of under-represented minorities (which I estimate by the fraction of students in the ELS 2002 that are URMs).

Table F-2: Technology of Skill Formation

	URM	White & Asian
Knowledge(-1)	1.03	1.04
	(0.00)	(0.02)
Study, 10 hours/wk	0.09	0.08
	(0.01)	(0.01)
Private School	0.05	0.02
	(0.01)	(0.01)
Free Lunch	-0.11	-0.17
	(0.02)	(0.03)
Student Teacher Ratio	0.00	0.00
	(0.00)	(0.00)
Mother: High School	-0.01	0.00
	(0.01)	(0.02)
Mother: Some College	-0.02	-0.01
	(0.01)	(0.01)
Mother: Bachelors	-0.03	-0.02
	(0.02)	(0.02)
Mother: Postgraduate	-0.03	-0.02
	(0.02)	(0.02)
Constant	0.20	0.18
	(0.02)	(0.03)
σ_k^2	0.00	0.00
	(0.38)	(0.38)

The table displays estimates of parameters of the technology of skill formation with a stochastic shock. Study refers to the effect of studying 10 hours per week on next year's skills. Free lunch is measured on a scale from 0 to 1. High school dropout is the omitted education category. SOURCE: (ELS 2002)

G.1 Solving the Model in Counterfactual Simulations

In counterfactual simulations, the distribution of skills and the pattern of college applications may differ from the data, so they cannot be conditioned on, but rather must be integrated over when solving for the equilibrium thresholds. In these simulations, the probability of attending school c is

$$\begin{aligned}
\mathbb{P}(Attend_{i,c}=1) &= \int \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,9}) dF(\Omega_{i,9}) \;, \\
&= \frac{1}{N}\sum_i \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,9}) \;, \\
&= \frac{1}{N}\sum_i \sum_a \int \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12}) dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) \;. \qquad \text{(G-1)}
\end{aligned}$$

The innermost integral requires integrating $\mathbb{P}(Attend_{i,c} = 1 \mid \Omega_{i,12})$ only over the distribution of $(m_{i,12}, P_{i,12})$, which govern admissions chances for individual i, so $f(\Omega_{i,12} \mid a, \Omega_{i,9})$ simplifies to $f(m_{i,12}, P_{i,12} \mid a, \Omega_{i,9})$. Note that $P_{i,12}$, the conditional variance of

$\log K_{i,12}$, is deterministic conditional on the choice, a, so $f(m_{i,12}, P_{i,12} \mid a, \Omega_{i,9}) = f(m_{i,12} \mid a, \Omega_{i,9})\delta_{P_{i,12}(a)}$, where $\delta_{P_{i,12}(a)}$ is the Dirac delta function that places infinite mass on the point $P_{i,12}(a)$.

The expression in (G-1) shows that expected attendance at college c can be calculated as the average across all students of the integral of the probability of attending college conditional on the state variables in the 12th grade, where the integral is over (i) the single continuous state variable $m_{i,12}$ conditional on choices and initial conditions $(a, \Omega_{i,9})$ and (ii) the choices (a) the student made in high school whose Logit probabilities are given by the solution to (19).[40]

The probability of attending college c conditional on the state variables in the 12th grade can be expressed as follows:

$$\begin{aligned}\mathbb{P}(Attend_{i,c} = 1 \mid \Omega_{i,12}) = \mathbb{P}(&Matriculate_{i,c} = 1 \mid Admit_{i,c} = 1, Apply_{i,c} = 1, \Omega_{i,12}) \times \\ &\mathbb{P}(Admit_{i,c} = 1 \mid Apply_{i,c} = 1, \Omega_{i,12}) \times \mathbb{P}(Apply_{i,c} = 1 \mid \Omega_{i,12}) \\ &= \sum_{B \subseteq A} \mathbb{P}(C_i = C \mid B, \Omega_{i,12}) \times \sum_{A \in \mathcal{A}(SAT_i(a))} \mathbb{P}(B \mid A, \Omega_{i,12}) \mathbb{P}(A \mid \Omega_{i,12}) \ ,\end{aligned}$$

where $\mathbb{P}(C_i = c \mid B, \Omega_{i,12})$ is given by equation (11), $\mathbb{P}(B \mid A, \Omega_{i,12})$ is given by equation (13), and $\mathbb{P}(A \mid \Omega_{i,12})$ is given by equation (17). Admissions thresholds influence both $\mathbb{P}(B \mid A, \Omega_{i,12})$ and $\mathbb{P}(A \mid \Omega_{i,12})$.

When there are multiple colleges per tier, as in estimation, I modify $\mathbb{P}(B \mid A, \Omega_{i,12})$ as follows:

$$P(B \mid A, \Omega_{i,12}) = \prod_{c=1}^{C} \binom{A(c)}{B(c)} \mathbb{P}(S_i^c > S_{URM_i}^{c*} \mid \Omega_{i,12})^{B(c)} \mathbb{P}(S_i^c < S_{URM_i}^{c*} \mid \Omega_{i,12})^{A(c)-B(c)} \ ,$$

where $A(c)$ denotes the number of applications to schools in tier c in portfolio A, and $B(c)$ is defined analogously for acceptances.

The first part of colleges' objective function, $\mathbb{E}[\log K_{i,12} \mid Attend_{i,c} = 1]$ can be expressed as follows:

$$\mathbb{E}[\log K_{i,12} \mid Attend_{i,c} = 1] = \frac{\mathbb{E}[\log K_{i,12} \mathbb{1}(Attend_{i,c} = 1)]}{\mathbb{P}(Attend_{i,c} = 1)} \ ,$$

[40] The formula for $\mathbb{P}(Attend_{i,c} = 1 \mid URM_i = 1)$ is analogous to equation (G-1), except that it averages only over URM students.

where the numerator equals:

$$
\begin{aligned}
\mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)] &= \frac{1}{N}\sum_i \mathbb{E}[\log K_{i,12}\mathbb{1}_{Attend_{i,c}=1} \mid \Omega_{i,9}] \ , \\
&= \frac{1}{N}\sum_i\sum_a \int \mathbb{E}[\log K_{i,12}\mathbb{1}_{Attend_{i,c}=1} \mid \Omega_{i,12}) dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) \ , \\
&= \frac{1}{N}\sum_i\sum_a \int m_{i,12}\mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12}) dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) \ ,
\end{aligned}
$$

where the third equality follows from independence between $\log K_{i,12}$ and the shocks governing matriculation, admission, and application conditional on $\Omega_{i,12}$.

The partial derivatives of diversity with respect to the admissions thresholds satisfy:

$$
\frac{\frac{\partial \mathbb{P}(URM_i=1|Attend_{i,c}=1)}{\partial S_j^{c*}}}{\mathbb{P}(URM_i=1 \mid Attend_{i,c}=1)} = \frac{\frac{\partial \mathbb{P}(Attend_{i,c}=1|URM_i=1)}{\partial S_j^{c*}}}{\mathbb{P}(Attend_{i,c}=1 \mid URM_i=1)} - \frac{\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial S_j^{c*}}}{\mathbb{P}(Attend_{i,c}=1)} \ ,
$$

for $j = URM, WA$. Note that if $j = WA$, then $\frac{\partial \mathbb{P}(Attend_{i,c}=1|URM_i=1)}{\partial S_j^{c*}} = 0$.

The remaining equations needed to compute the first-order conditions in section 3.4 and solve for the equilibrium are copied below:

$$
\begin{aligned}
\frac{\partial \mathbb{E}[\log K_{i,12} \mid Attend_{i,c}=1]}{\partial S_j^{c*}} &= \frac{\frac{\partial \mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)]}{\partial S_j^{c*}}\mathbb{P}(Attend_{i,c}=1) - \mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{S_j^{c*}}}{\mathbb{P}(Attend_{i,c}=1)^2} \ , \\
\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial S_j^{c*}} &= \frac{1}{N}\sum_i\sum_a \int \frac{\partial \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12})}{\partial S_j^{c*}} dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) \ , \\
\frac{\partial \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12})}{\partial S_j^{c*}} &= \sum_{B \subseteq A} \mathbb{P}(C_i = C | B, \Omega_{i,12}) \times \\
&\quad \sum_{A \in \mathcal{A}(SAT_i(a))} \frac{\partial \mathbb{P}(B|A,\Omega_{i,12})}{\partial S_j^{c*}}\mathbb{P}(A \mid \Omega_{i,12}) + \mathbb{P}(B|A,\Omega_{i,12})\frac{\partial \mathbb{P}(A \mid \Omega_{i,12})}{\partial S_j^{c*}} \ , \\
\frac{\partial \mathbb{P}(B|A,\Omega_{i,12})}{\partial S_j^{c*}} &= \begin{cases} \mathbb{P}(B \mid A, \Omega_{i,12})\phi(S_j^{c*}) \left[-\frac{B(c)}{\mathbb{P}(S_i^c > S_j^{c*}|\Omega_{i,12})} + \frac{A(c)-B(c)}{\mathbb{P}(S_i^c < S_j^{c*}|\Omega_{i,12})} \right] & \text{for } URM_i = j \\ 0 \quad \text{for } URM_i \neq j \end{cases} \\
\frac{\partial \mathbb{P}(A \mid \Omega_{i,12})}{\partial S_j^{c*}} &= \frac{\mathbb{P}(A \mid \Omega_{i,12})}{\lambda_A} \Bigg(\sum_{B \subseteq A} \frac{\partial \mathbb{P}(B|A,\Omega_{i,12})}{\partial S_j^{c*}} U_{i,B} - \\
&\quad \sum_{A' \in \mathcal{A}(SAT_i(a))} \mathbb{P}(A' \mid \Omega_{i,12}) \sum_{B' \subseteq A'} \frac{\partial \mathbb{P}(B'|A',\Omega_{i,12})}{\partial S_j^{c*}} U_{i,B'} \Bigg) \ ,
\end{aligned}
$$

where $U_{i,B}$ comes from equation (12).

$$\begin{aligned}
\frac{\partial \mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)]}{\partial S_j^{c*}} &= \frac{1}{N}\sum_i \frac{\partial \mathbb{E}[\log K_{i,12}\mathbb{1}_{Attend_{i,c}=1} \mid \Omega_{i,9}]}{\partial S_j^{c*}} , \\
&= \frac{1}{N}\sum_i \sum_a \int \frac{\partial \mathbb{E}[\log K_{i,12}\mathbb{1}_{Attend_{i,c}=1} \mid \Omega_{i,12})}{\partial S_j^{c*}} dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) , \\
&= \frac{1}{N}\sum_i \sum_a \int m_{i,12} \frac{\partial \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12})}{\partial S_j^{c*}} dF(\Omega_{i,12} \mid a, \Omega_{i,9}) p(a \mid \Omega_{i,9}) ,
\end{aligned}$$

Since each tier contains of a continuum of colleges, schools do not take into account how their choice of admissions thresholds influences students' choices while in high school. This explains why none of the first-order conditions from the college problem contain partial derivatives of $p(a \mid \Omega_{i,12})$ with respect to the admissions thresholds.

G.2 Solving the Model in Estimation

In estimation, colleges take as given the pattern of applications $\{A_i\}_{i=1}^N$ and distribution of state variables $\{\Omega_{i,12}\}_{i=1}^N$ in the data. The only uncertainty when setting thresholds is over the probability of attendance conditional on admission. Several of the expressions in the previous section are therefore greatly simplified. The expressions that simplify are as follows:

$$\begin{aligned}
\mathbb{P}(Attend_{i,c}=1) &= \frac{1}{N}\sum_i \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12}) , \\
\mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)] &= \frac{1}{N}\sum_i m_{i,12}\mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12}) , \\
\frac{\partial \mathbb{P}(Attend_{i,c}=1)}{\partial S_j^{c*}} &= \frac{1}{N}\sum_i \frac{\partial \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12})}{\partial S_j^{c*}} , \\
\frac{\partial \mathbb{E}[\log K_{i,12}\mathbb{1}(Attend_{i,c}=1)]}{\partial S_j^{c*}} &= \frac{1}{N}\sum_i m_{i,12}\frac{\partial \mathbb{P}(Attend_{i,c}=1 \mid \Omega_{i,12})}{\partial S_j^{c*}} .
\end{aligned}$$

G.3 Computation

Since colleges take $\{\Omega_{i,12}, A_i\}_{i=1}^N$ as given, solving for the equilibrium thresholds in the NFXP algorithm is not computationally costly: I need only to calculate the probabilities of admission and matriculation given the observed application and admission portfolios. However, evaluating $P(a_i \mid \Omega_{i,12}, \theta)$ *is* very costly, because for every student and every action she may take, I must integrate over the distribution of $(m_{i,12}, P_{i,12})$ given $(\Omega_{i,9}, a)$, and compute the inclusive value, $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i(a))$, which depends on all 584 possible application portfolios. To speed up computation, I precompute $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i)$ on a grid for each student and use interpolation to simulate the values between grid points. $\overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i)$ varies smoothly with $m_{i,12}$, and Monte Carlo simulations (available upon request) reveal that interpolation introduces negligible error.

Solving for the admissions thresholds in counterfactual policy simulations is also computationally intensive, because the applications in these counterfactual scenarios do not yet exist and cannot be conditioned on, as in estimation. Instead, these counterfactual application portfolios must be integrated over. This requires evaluating, for each student, for each action they take in high school, the probability of attending each school, which involves integrating over the realization of the continuous variable $m_{i,12}$, the probability of applying to each of the 584 application portfolios, and the probability of matriculating to each school conditional on each possible admissions portfolio. A single function evaluation can take over 24 hours, and solving for the equilibrium would be computationally infeasible on all but the most advanced machines.

Careful inspection of the model reveals that, for each set of admissions thresholds, $\{\{S_{URM}^{c*}\}_{URM=0}^{1}\}_{c=1}^{C}$, the $2 \times C$ equations in section 3.4 that characterize the equilibrium depend only on the following $2 \times C + 1$ quantities

$$\begin{aligned} \mathbb{P}(Attend_{i,c} = 1) \text{ for } c = 1, \ldots, C \ , \\ \frac{\partial \mathbb{P}(Attend_{i,c} = 1)}{\partial S_{URM}^{c*}} \text{ for } c = 1, \ldots, C \ , \\ \overline{V}_i^{Coll}(m_{i,12}, P_{i,12}, SAT_i) \ , \end{aligned} \tag{G-2}$$

for every individual $i = 1, \ldots, N$. The values in (G-2) are time-consuming to compute, but they are themselves deterministic functions of the following variables:

$$\Upsilon_i = \Bigg(\left\{ \mathbb{P}(S_i^c > S_{URM_i}^c{}^* \mid m_{i,12}, P_{i,12}), \frac{\partial \mathbb{P}(S_i^c > S_{URM_i}^c{}^* \mid m_{i,12}, P_{i,12})}{\partial S_{URM_i}^c{}^*}, U_{i,c}, Dist_{i,c} \right\}_{c=1}^{C}, \\ U_{i,0}, Inc_i, MomCollege_i \Bigg) \ .$$

I use lasso regression to estimate the equations in (G-2) as linear functions of a second-order polynomial of the variables in Υ_i. Then, when optimizing the equilibrium mapping to solve for the admissions thresholds, S_j^{c*} for $j = 0, 1$, I recompute the values of Υ_i that change when the thresholds are updated, and use the regression coefficients obtained by lasso to quickly project the values of (G-2) before evaluating the equilibrium mapping.

The R^2 from the $2 \times C + 1$ lasso regressions used to predict the values in (G-2) average

0.97. This method shortens the computation time for evaluating the equilibrium from over 24 hours on a standard PC to 12 seconds while introducing negligible error.

H Standard Errors

Section 5 describes how the model is estimated in two steps. Standard errors for θ_1, which include the parameters governing the dynamic factor model and college completion model are obtained from the Hessian of the partial likelihood function in (23). I then use the delta method to obtain standard errors for, θ_2, the preference parameters estimated by optimizing equation (24).

The delta method states that if we have a set of parameters, θ_1, distributed normally with variance V as follows

$$\theta_1 \sim N(0, V) \ ,$$

then some (possibly) vector-valued function of θ_1, $h(\cdot)$, has the following distribution

$$h(\theta_1) \sim N\left(0, \frac{\partial h}{\partial \theta_1'} V \frac{\partial h}{\partial \theta_1}\right) \ . \tag{H-1}$$

To use the delta method, I write the solution to the second step of the optimization as a function of θ_1 so that $\theta_2^* = h(\theta_1)$, where

$$h(\theta_1) = \arg\max_{\theta_2} l(\theta_1, \theta_2) \ .$$

The Jacobian of h with respect to θ_1 is then

$$\frac{\partial h}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} \arg\max_{\theta 2} l(\theta_1, \theta_2) \ .$$

I use a Taylor expansion around the solution to the second step of optimization to obtain an expression for this Jacobian.

At the optimum of the second stage of estimation, we know that

$$\frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) = 0 \ .$$

A first-order Taylor expansion of this equation around $(\theta_1, \theta_2) = (a, b)$ yields:

$$\frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) = \frac{\partial}{\partial \theta_2} l(a, b) + \frac{\partial^2}{\partial \theta_2 \partial \theta_1} l(a, b)(\theta_1 - a) + \frac{\partial^2}{\partial \theta_2^2} l(a, b)(\theta_2 - b) + R_1(\theta_1, \theta_2) \ .$$

Through rearranging and ignoring higher-order terms, we get that

$$\begin{aligned} \theta_2 - b &= \left(\frac{\partial^2}{\partial \theta_2^2} l(a, b) \right)^{-1} \left(\frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) - \frac{\partial}{\partial \theta_2} l(a, b) - \frac{\partial^2}{\partial \theta_2 \partial \theta_1} l(a, b)(\theta_1 - a) \right) \ , \\ &\approx \left(\frac{\partial^2}{\partial \theta_2^2} l(a, b) \right)^{-1} \left(- \frac{\partial^2}{\partial \theta_2 \partial \theta_1} l(a, b)(\theta_1 - a) \right) \ , \end{aligned}$$

where the second equality follows because the first-order conditions of the likelihood are near zero in the vicinity of the solution. Rewriting the above equation in terms of differentials yields

$$dh = \left(\frac{\partial^2}{\partial \theta_2^2} l(a, b) \right)^{-1} \left(- \frac{\partial^2}{\partial \theta_2 \partial \theta_1} l(a, b) d\theta_1 \right) \ ,$$

and so

$$\frac{\partial h}{\partial \theta_1} = \left(\frac{\partial^2}{\partial \theta_2^2} l(a, b) \right)^{-1} \left(- \frac{\partial^2}{\partial \theta_2 \partial \theta_1} l(a, b) \right) \tag{H-2}$$

The first part of equation (H-2) is just the Hessian from the second-step of estimation. I estimate the second part numerically by perturbing the joint likelihood. With an estimate of $\frac{\partial h}{\partial \theta_1}$, I compute the standard errors for θ_2 according to (H-1). The standard errors in Tables 7 and 9 are computed according to this method. The standard errors for θ_1, in the remaining tables, are obtained immediately from the Hessian from the first stage of optimization.

I Additional Policy Counterfactuals

I.1 SAT-Optional Policy at Elite Private Colleges Only

The policies examined in the main text of the paper analyze educational outcomes when all colleges pursue the same admissions policy. This section instead evaluates what happens when a single college deviates and pursues a different policy than the rest. I consider a policy in which elite private colleges permit applications from non SAT-takers, while other colleges continue to require the SAT. Students who do not take the exam can either apply to elite private colleges or to no college at all. This policy therefore aims to mimic the behavior of small numbers of liberal arts colleges that started going SAT-Optional in 2005, while the rest of the college market continued to require the exam (Epstein 2009). For computational reasons, I abstract from the strategic decision of whether an SAT-taker should send an SAT score and instead assume that all students who take the SAT send their scores.[41]

Table I-1 shows the pattern of college sorting by knowledge under the SAT-Optional policy. The first two sets of columns show sorting under the status quo and under the No-SAT policy analyzed in the main text for comparison. Relative to the status quo, there is little change in college sorting by knowledge when elite private colleges go SAT-Optional. Average knowledge at tier one schools actually increases. This result contrasts with the marked decline in average knowledge under the No-SAT policy, and it arises because elite colleges continue to receive applications from students who applied in the status quo with an SAT score, for whom there is no loss of information, plus additional applications from students who did not take the SAT. Admissions thresholds at tier one schools increase by over one-tenth of a standard deviation (Table I-2), underscoring how policies that lead to more applications can increase selectivity. The effect of raising thresholds offsets the cost of not having the SAT for some of the applicants and leaves elite private colleges with similarly skilled students.[42]

Table I-3 displays estimates of racial diversity, λ^c_{URM}, and sorting to college by household income under the SAT-Optional policy. Relative to the status quo, racial diversity falls slightly at Tier 1 colleges under the SAT-Optional policy, from 18.2% to 16.5% of students. SAT-Optional admissions does, however, help elite colleges to enroll more low-income students and the average household income of students attending these colleges falls, from \$110,000 to \$105,000. Even though diversity falls, the increase in the average knowledge of matriculating students is large enough to raise the objective function of Tier 1 colleges ($\kappa_1 \cdot 1.45 + (1-\kappa_1) \cdot \log(0.165) > \kappa_1 \cdot 1.41 + (1-\kappa_1) \cdot \log(0.182)$). The fact that a single college tier can improve its objective function by going SAT-Optional suggests that requiring the SAT may not have been a Nash equilibrium and provides one explanation for the unraveling of SAT requirements since 2005.

[41] Allowing students to decide whether to send their scores to elite private colleges increases the size of the application portfolio from 3^6 to 2×3^6, because there are now two separate applications to elite private colleges, those that contain SAT scores and those that omit them.

[42] Appendix J shows that these findings are consistent with papers in the education literature that evaluate SAT-Optional policies at liberal arts colleges using difference-in-differences.

Table I-1: SAT-Optional at Elite Private Colleges

	Status Quo			No SAT			SAT-Optional		
$\mathbb{E}[\log K_{i,12}]$	All	URM	WA	All	URM	WA	All	URM	WA
Tier 1	1.41	0.96	1.51	1.20	0.66	1.34	1.45	1.00	1.54
Tier 2	0.79	0.37	0.90	0.77	0.33	0.87	0.76	0.34	0.88
Tier 3	0.38	-0.08	0.54	0.38	-0.07	0.54	0.39	-0.08	0.55
Tier 4	1.15	0.80	1.22	1.03	0.60	1.10	1.14	0.79	1.21
Tier 5	0.76	0.47	0.82	0.73	0.39	0.79	0.73	0.45	0.79
Tier 6	0.31	-0.14	0.45	0.39	-0.06	0.54	0.30	-0.15	0.45
Any College	0.61	0.14	0.74	0.6	0.10	0.73	0.60	0.13	0.73
No College	-0.47	-0.83	-0.28	-0.49	-0.82	-0.31	-0.47	-0.83	-0.28

The table presents simulated estimates of mean knowledge by college tier under the status quo, No SAT, and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite private colleges (Tier 1). Simulated moments are computed using 200 simulated data sets. WA refers to the population of white and Asian students. SOURCE: (ELS 2002)

Table I-2: Admissions Thresholds, SAT-Optional

	Status Quo		SAT-Optional	
	White & Asian	URM	White & Asian	URM
Tier 1	1.07	0.54	1.18	0.73
Tier 2	0.25	-0.20	0.19	-0.26
Tier 3	-0.29	-0.85	-0.31	-0.87
Tier 4	0.71	0.39	0.70	0.38
Tier 5	0.13	-0.02	0.07	-0.08
Tier 6	-0.51	-1.01	-0.54	-1.04

The table presents admissions thresholds at each college tier under the status quo and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite private colleges (Tier 1). Thresholds have been standardized by the mean and sd of $\log K_{i,12}$. SOURCE: (ELS 2002)

I.2 How Biased would the SAT Need to Be to Justify Abandoning It?

Table 5 in the main text found no evidence that the SAT math and verbal examinations were biased against URM students, but assessing this required assuming scalar invariance of the NCES math exams, because it is not possible to distinguish a difference in latent skills by demographic from a level shift in *all* the intercepts mapping those latent skills to the observed measurements. This means that if all measurements were biased against URM students, the tests in Table 5 would be unable to detect any bias. Even if the exams were biased against URMs, this would not affect their chance of admission *in the model*, because colleges in the model use the Kalman Filter, which corrects for any biases when using the measurements to predict each students' knowledge. This section therefore explores a setting in which the SAT is biased against URMs and colleges do not

Table I-3: Diversity and Household Income by College

	Diversity (λ^c_{URM})			Household Income		
	Status Quo	No SAT	SAT-Optional	Status Quo	NoSAT	SAT-Optional
Tier 1	0.182	0.205	0.165	110K	108K	105K
Tier 2	0.207	0.194	0.209	96K	94K	95K
Tier 3	0.260	0.266	0.256	74K	71K	74K
Tier 4	0.160	0.135	0.163	103K	100K	103K
Tier 5	0.162	0.134	0.164	92K	90K	92K
Tier 6	0.247	0.252	0.247	72K	70K	72K

The table presents simulated estimates of the fraction of URMs attending each tier of college and mean household income under the status quo, No SAT, and SAT-Optional policies. The SAT-Optional Policy allows students who have not taken the SAT to apply to elite private colleges (Tier 1). Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

correct for it. The exercise asks, How biased would the SAT need to be for banning it to raise URM enrollment?.

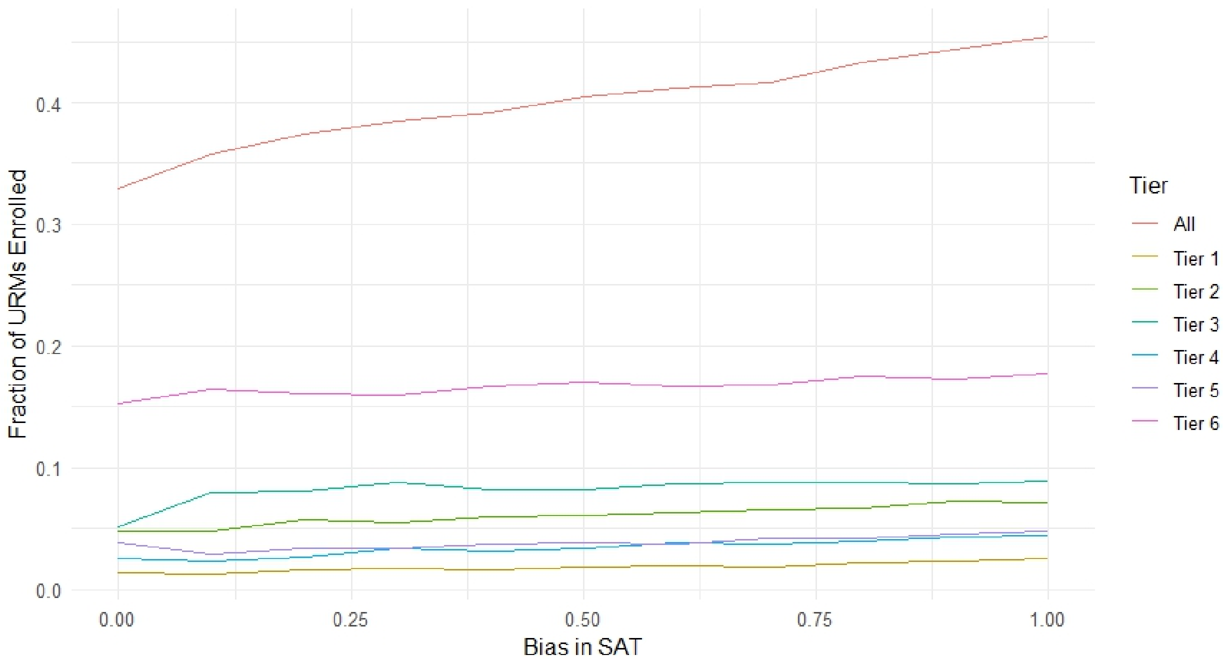
To evaluate this scenario, I solve for each university's preferences for knowledge and diversity, given by κ_c, under assumptions that both the SAT math and verbal exams are biased by x standard deviations, where I vary x between 0 and 1. When these exams are biased against URMs and universities fail to take this into account, the admissions decisions observed in the data can only be rationalized by a higher preference for diversity (lower κ_c). With this new value of κ_c, I use the model to simulate banning the SAT and plot university enrollment in Figure I-1. The figure shows that, as the bias against URMs increases, banning it causes URM access to increase from 32.9% of the sample enrolled in four-year colleges, when there is no bias in the exam, to 45.4%, when the SAT math and verbal exams are both biased by one standard deviation. URM students score 0.6 sd lower than white and Asian students on the SAT math and verbal exams in the ELS 2002. The figure suggests that, if there were no population differences in skills and the entire difference in SAT scores were due to bias that universities fail to account for, banning the SAT would raise URM enrollment by 8.2 pp.

I.3 Beyond Affirmative Action

The model in section 3 allows colleges to practice affirmative action (AA) in two related ways: through a preference for diversity, κ_c in equation (7), and through having separate demographic-specific thresholds for admission. All the estimated thresholds in Table 8 differ by demographic, implying that at least some colleges in each tier engage in AA.

In June 2023, the Supreme Court banned AA in college admissions, simultaneously

Figure I-1: URM Enrollment as a Function of Uncorrected Bias in the SAT



The figure shows URM enrollment if the SAT were banned as a function of bias in the SAT math and verbal exams. The simulation assumes that colleges do not correct for any bias against URM students. Under an assumption that the two exams are biased, status quo enrollment can only be rationalized by colleges having a higher preference for diversity. Banning the SAT is then simulated under this higher preference for diversity. Bias in both exams proceeds in increments of 0.1 sd, and 20 simulated data sets are used to compute URM enrollment for each bias increment.

outlawing both preferences for diversity and separate admissions thresholds by demographic (Students for Fair Admissions, Inc. v. President and Fellows of Harvard College, 600 U. S. 2023). In this section I use the model to simulate how banning AA would change enrollment patterns. Given the university problem in (7), each school's race-blind threshold is entirely determined by its capacity constraint.[43]

Table I-4 presents simulated attendance rates for URM and low-income students following the AA ban. Relative to a setting with AA (Table 11), banning AA reduces URM attendance at each school and by 6-8 pp overall depending on the policy regime.[44] Banning AA also reduces low-income enrollment, but by less. When AA is outlawed, banning the

[43] Chan and Eyster (2003) analyze a setting in which colleges have racial preferences but cannot use separate admissions thresholds by race. They show that the optimal admissions policy is a mixed strategy that depends on the capacity of the college and typically involves randomly admitting students with skills above a threshold. In this section, I remove racial preferences from colleges' objective function (7). Colleges compete to admit the most knowledgeable students, leading to a single threshold rule for all students.

[44] Bleemer (2022) finds that the AA ban in California in 1998 decreased URM enrollment at UC campuses by 5.8 pp conditional on application.

SAT and going SAT-for-All both have negligible effects on URM enrollment. Low-income enrollment, however, increases under both counterfactual policies relative to the status quo admissions policy.

The intuition for the null result of changing SAT policies on URM enrollment is the same regardless of whether colleges engage in AA. There are a small number of non-SAT takers who would gain admission without SAT requirements, but these marginal applicants are largely shut out by the rise in admissions thresholds as more students apply. Table I-5 shows that thresholds rise at each school, sometimes by quite large amounts, when banning the SAT and going SAT-for-All. While the ban on AA presents a new policy environment to analyze, the same mechanisms as before frustrate efforts to increase URM enrollment. Fundamentally, banning the SAT pits a larger and more diverse applicant pool at the expense of less information. If the pool of students who fail to take the SAT in the status quo contains a sizable number of highly-skilled URMs who can outcompete existing SAT-takers, then banning the SAT may raise URM enrollment. But, Figure 2 in the main text showed that this is unlikely, because the difference in skills between takers and non-takers is large. The negligible effects on URM enrollment in Table I-4 result from the unequal distribution of skills – between SAT-takers and non-takers and between URM and white and Asian students – among likely college applicants.

Table I-4: Access to College without Affirmative Action

	URM Attendance			*Low-Income Attendance*		
	Status Quo	No SAT	SAT-for-All	Status Quo	No SAT	SAT-for-All
Attend Tier 1 School	0.006	0.006	0.006	0.007	0.007	0.008
Attend Tier 2 School	0.034	0.033	0.029	0.033	0.035	0.032
Attend Tier 3 School	0.038	0.045	0.045	0.043	0.054	0.054
Attend Tier 4 School	0.016	0.016	0.016	0.018	0.019	0.019
Attend Tier 5 School	0.035	0.035	0.032	0.038	0.039	0.037
Attend Tier 6 School	0.134	0.128	0.132	0.149	0.161	0.165
Attend Any College	0.265	0.262	0.261	0.288	0.315	0.315

The table displays rates of attendance for URM and low-income students at each college tier following the ban on affirmative action under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit their scores with their applications. Low-income refers to students whose families earn less than the median ($52,500 per year). Simulated moments are computed using 200 simulated data sets. SOURCE: (ELS 2002)

Table I-5: Admissions Thresholds, with and without Affirmative Action

	Affirmative Action						No Affirmative Action		
	Status Quo		*No SAT*		*SAT-for-All*		*Status Quo*	*no SAT*	*SAT-for-All*
	WA	URM	WA	URM	WA	URM	All	All	All
Tier 1	2.50	1.81	2.55	1.85	2.60	1.93	2.40	2.45	2.50
Tier 2	1.43	0.86	1.65	1.22	1.73	1.22	1.29	1.55	1.60
Tier 3	0.75	0.02	1.01	0.36	1.02	0.34	0.55	0.83	0.84
Tier 4	2.03	1.62	2.11	1.87	2.15	1.78	1.96	2.06	2.08
Tier 5	1.29	1.09	1.51	1.48	1.59	1.42	1.22	1.49	1.53
Tier 6	0.47	-0.18	0.98	0.42	0.99	0.39	0.24	0.81	0.81

The table presents admissions thresholds at each college tier before and after the ban on affirmative action under three separate policies: the status quo, a policy where the SAT is banned, and a policy in which all students take the SAT and submit their scores with their applications. After the ban on affirmative action, all applicants to a school are subject to the same admissions threshold. WA refers to the population of white and Asian students. SOURCE: (ELS 2002)

J External Validation

J.1 Prior Literature

A small literature in economics has examined the effects of changing SAT requirements in admissions. Hurwitz et al. (2015) find that mandating the SAT in Maine caused a 2-3 pp increase in enrollment but no statistically significant increase in college completion. The authors' back-of-the-envelope calculations suggest that rates of completion would have increased by 1.3 pp, but such a change would be too small to detect in the sample. Hyman (2017) find that SAT-for-All raised four-year completion rates in Michigan by 0.5 pp, which is statistically significant given the much larger sample size. Half a percentage point is surely a lower bound for eight-year completion rates, which this paper finds increase by 1 pp in response to SAT-for-All.

In the mid 2000s a number of schools, primarily liberal arts colleges, began allowing applicants to voluntarily submit SAT scores, a policy that has been dubbed SAT-Optional admissions. Belasco, Rosinger, and Hearn (2015) create a panel of 180 selective liberal arts colleges between 1992 and 2010 and use difference-in-differences to compare several outcomes between schools that went test-optional and those still requiring exam scores. They find no effects of SAT-optional admissions on either URM enrollment or the proportion of Pell grant recipients (a proxy for low income). The policy instead caused a significant increase of approximately 14% in the number of applications received. These findings, which are confirmed in later analysis by Sweitzer, Blalock, and Sharma (2018) and Rosinger and Ford (2019), are consistent with the No-SAT policy evaluated in this paper. Saboe and Terrizzi (2019) extend the analysis to all colleges and universities using more recent data, and similarly find null results on racial and socioeconomic diversity. Bennett (2022) instead argues that the control group in this analysis should not include all colleges and universities, but instead only those schools with similar Barron's selectivity rankings who continued to maintain SAT requirements. In this modifed sample, Bennett finds that going SAT-Optional raised URM enrollment by 10% and the enrollment of Pell Grant recipients by about 3-4% but had not effect on applications. Rationalizing these findings with the earlier literature will require exploring the trends in enrollment and applications at schools excluded by Bennett but included in earlier analyses.

J.2 Statistics Following Covid-19 Pandemic

In response to the onset of the Covid-19 pandemic, a large majority of schools stopped requiring SAT scores from applicants. This makes it possible to use recent data to validate some of the model's predictions. This analysis is subject to several caveats. The first is that this paper uses data for a cohort of students that matriculated to college in 2004, sixteen years before the pandemic. Additionally, the pandemic altered many aspects of life related to the transition to college, not only admissions policies, and changes in patterns of college attendance are likely to reflect a combination of admissions changes and other Covid-induced factors. The model counterfactuals also feature an endogenous effort response, but it is likely that few students foresaw the onset of the pandemic and adjusted their behavior accordingly while in high school. Lastly, the current SAT-Optional policy differs from the ban on the SAT analyzed in the main text, although it still incorporates the same inherent tradeoff between allowing for a larger applicant pool but reducing the information available to select candidates.

At the time of writing, there were two publicly available data sources to externally validate the model. It is possible to construct tier-specific growth rates in applications between the 2019-2020 and 2020-2021 application season using data from the Integrated Postsecondary Education Data System (IPEDS).[45] The Common App instead has statistics on overall application growth for its 853 member colleges but does not break the statistics down by tier (Freeman et al. 2022). Table J-1 reveals that applications increased overall after the elimination of SAT requirements (and after the onset of the pandemic) by 6% in the IPEDS data and by 21% in the Common App data. The IPEDs statistics show dramatic increases in applications to elite schools (tiers 1 and 4). The corresponding statistics in the

[45] The 2020-2021 application season was the first season in which a majority of schools no longer required that applicants submit SAT scores.

Table J-1: Application Growth for 2021-2022 Academic Year

	Model	IPEDS	Common App
Tier 1	0.16	0.25	
Tier 2	0.17	0.05	
Tier 3	0.27	-0.06	
Tier 4	0.16	0.15	
Tier 5	0.20	0.09	
Tier 6	0.24	0.00	
Overall	0.21	0.06	0.21

The table displays statistics on application growth in 2020-2021 relative to 2019-2020 from IPEDS and the Common App compared with analogous statistics generated by the model in response to banning the SAT. SOURCE: Integrated Postsecondary Education Data System and (ELS 2002)

model for the policy that bans the SAT are 21% for overall application growth, and 16% for application growth at elite private and public universities.

The Common App breaks down statistics on application growth by URM status. Applications from URMs increased by 18%, while non-URM applications increased by 13%. The findings in this paper are 31% and 18%, respectively. The total increase in application volume to Common App schools between 2019-2020 and 2021-2022 was 21.3%. The corresponding number in the paper is 21%.

An interesting finding from the Common App statistics is that a racial gap in test score reporting has opened up. Before the change in admissions policies, URM applicants were 5% less likely to submit SAT scores with their applications. That figure has since increased to 15%. As long as test scores are not necessary, this paper predicts that URMs would be less likely to report them since they face greater financial and logistical barriers to taking the exam.

K A Model Including Noncognitive Skills

This section presents the estimates of a dynamic factor model of skill accumulation in high school with two latent skills. This model includes measurements of students' underlying noncognitive skills in addition to the cognitive measurements of grades and standardized tests analyzed in the main text.

The ELS 2002 creates subscales of noncognitive skills for students in the 10th and the 12th grades by taking the first principal component of a series of survey responses by the students and their teachers. There are eight of these noncognitive skill measurements, seven in the 10th grade and one in the 12th grade. The tenth grade measurements are English Self-Efficacy, Math Self-Efficacy, Writing Ability, Control Expectation, Action Control, Motivation, and Class Preparation. Students are additionally scored according to their Math Self-Efficacy in the 12th grade. All subscales except for the writing ability subscale are formed from ordered responses to individual survey questions asking students how much they agree with statements like "I'm certain I can understand the most difficult material presented in English texts" and "When studying, I try to work as hard as possible." The writing ability subscale is instead created from questions that ask each student's English teacher to rate things like the student's "ability to organize ideas logically and coherently."[46]

The noncognitive skill measurements do not appear in each year of high school, which makes it difficult to estimate a model that simultaneously allows for time-varying noncognitive and cognitive skills. The model in this section therefore allows cognitive skills to vary over time while noncognitive skills remain constant. The technology of skill formation is given by the following equation:

$$\begin{pmatrix} \log K_{i,t}^{(1)} \\ \log K_{i,t}^{(2)} \end{pmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \log K_{i,t-1}^{(1)} \\ \log K_{i,t-1}^{(2)} \end{pmatrix} + \begin{bmatrix} \beta_{11} & \dots & \beta_{1K} \\ 0 & \dots & 0 \end{bmatrix} \begin{pmatrix} I_{i,t}^{(1)} \\ \vdots \\ I_{i,t}^{(K)} \end{pmatrix} . \tag{K-1}$$

Equation (K-1) allows noncognitive skills in year $t-1$ to influence cognitive skills in year t through γ_{12} and investment variables to influence cognitive skills in the same year through $\beta_{11}, \dots, \beta_{1K}$. Noncognitive skills, however, are not affected by prior cognitive skills or contemporaneous investment.

As in section 3, the skill measurements are noisy representations of the underlying latent skills as follows:

$$\underbrace{\mathbf{y}_{i,t}}_{M_t \times 1} = \underbrace{\boldsymbol{\mu}_t^R}_{M_t \times 1} + \underbrace{\boldsymbol{\alpha}_t^R}_{M_t \times 2} \underbrace{\log K_{i,t}}_{2 \times 1} + \underbrace{\boldsymbol{\varepsilon}_{i,t}^R}_{M_t \times 1} , \tag{K-2}$$

where $\boldsymbol{\alpha}_t^R$ now has dimension $M_t \times 2$ (M_t is the number of measurements in year t) to account for the fact that some measurements load onto both the cognitive and noncognitive skill dimensions. I assume that the standardized tests in the data load only onto the first dimension of skill (which I refer to as cognitive), grade point averages (GPAs) load onto both dimensions of skill, and the noncognitive skill measurements load only onto the second dimension of skill.

[46]Depending on the question, the answers are Almost Never, Sometimes, Often, Almost Always; or Poor, Fair, Good, Very Good, Outstanding.

Because the noncognitive measurements are assumed to load only onto the second latent skill, the model is overidentified, and it is possible to allow for correlation between the two latent skills in the initial period.[47] I parameterize the covariance between cognitive and noncognitive skills as a function of the same variables that determine the mean and variance of initial skills (URM, gender, grade retention, a single parent dummy, mother's education, and household income). The exact structure of $\boldsymbol{\alpha}_t^R$, which determines which measurements load onto each skill, can be seen in Table K-4, where an estimate of 0 and a standard error of $(-)$ means that the measurement does not load onto that skill.

I estimate the two-skill dynamic factor model on the same sample as the main results in section 6. Because complete teacher surveys are only available for 45% of these students, I first use multiple imputation to impute a full set of responses to all questions before conducting principal components analysis and estimating the dynamic factor model, with the first principal component of each set of measurements forming a unique noncognitive skill measurement. The results of the estimation, in Tables K-1, K-2, K-3, K-4, and K-5, represent the results from the data set combining imputed and complete data.[48]

Table K-1 shows estimates of parameters governing the distribution of ninth-grade skills. The findings for cognitive skills are similar in sign and magnitude to those for the one-factor model in Table 4: URM students begin high school with a cognitive skill disadvantage of over half a standard deviation, students who had been retained prior to high school enter at an even larger disadvantage, and initial cognitive skills are sharply increasing in parental education. Noncognitive skills show a similar pattern, with URM and retained students experiencing a disadvantage in these skills and noncognitive skills increasing in mother's education and family income (albeitly more slowly than for cognitive skills). A noteworthy difference, however, is that girls enter high school with a very slight disadvantage in cognitive skills but an advantage in noncognitive skills of 0.11 sd.

Table K-2 presents estimates of the technology of skill formation. Studying and lagged knowledge are similarly productive as in the one-skill model (Table 6). Noncognitive skills, at least as identified by the noncognitive skill measurements in the ELS 2002, do not appear to contribute to the development of cognitive skills in high school.[49] The

[47] With fewer overidentifying restrictions, it would be necessary to normalize the covariance matrix of ninth grade skills to be diagonal.

[48] 45% of individuals in the sample from the main text have complete data on all survey responses that comprise the noncognitive subscales. However, only 15.1% of observations are imputed, because most students with incomplete data lack only a few observations. The imputation method is an ordered logistic regression that uses student and teacher ordered responses as predictors only. Therefore, a student's response to how hard they work when studying can be used to predict their (missing) response to a question about how prepared they are for class, but the student's demographic group, family income, or grades are not used as part of the imputation model.

[49] This contrasts with a body of research showing how noncognitive skills in childhood and early adoles-

Table K-1: Parameters Governing Initial Distribution of Skills

	Cognitive Skill		Noncognitive Skill	
	Mean	Log Variance	Mean	Log Variance
URM	-0.56	-0.03	-0.73	0.31
	(0.03)	(0.05)	(0.22)	(0.14)
Female	-0.06	-0.07	0.11	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Retain	-0.67	-0.15	-0.42	0.06
	(0.04)	(0.03)	(0.05)	(0.03)
Single Parent	-0.10	-0.01	-0.10	0.06
	(0.02)	(0.02)	(0.03)	(0.02)
Mother: High School	0.18	-0.05	0.02	0.01
	(0.04)	(0.03)	(0.05)	(0.03)
Mother: Some College	0.34	-0.07	0.07	-0.01
	(0.04)	(0.03)	(0.05)	(0.03)
Mother: Bachelors	0.60	-0.03	0.27	0.00
	(0.05)	(0.03)	(0.05)	(0.03)
Mother: Postgraduate	0.77	-0.03	0.35	-0.06
	(0.06)	(0.04)	(0.06)	(0.04)
HH Income	0.08	-0.01	0.04	-0.01
	(0.01)	(0.00)	(0.01)	(0.00)

The table presents estimates of parameters governing the initial distribution of cognitive and noncognitive skills in the ninth grade. The mean and variance of both skills have been normalized to 0 and 1, respectively, for individuals whose covariates are all equal to 0. HH Income is measured in hundreds of dollars per week. High school dropout is the omitted education category. SOURCE: (ELS 2002)

yearly change in cognitive skills, given by the constant, is similar to the model that does not account for noncognitive skills.

Table K-3 presents estimates of the intercepts in the measurement equations, μ_t^R, for all measurements in the data. The NCES exams in Math in the 10th and 12th grades and ninth grade GPA have been normalized to be unbiased. Even in a model when grades are allowed to load onto both cognitive and noncognitive skills, the inferences about bias on the SAT and GPA are qualitatively the same as in the single skill model in Table 5. There does not appear to be any evidence that the SAT is biased against URM students, and there appears to be some bias against URM students in twelfth grade GPA. Table K-3 provides evidence against bias in students' and teachers' subjective assessments of URM students. On nearly all of the noncognitive skill measures, URM students score higher than one would expect given their underlying latent noncognitive skill (the one exception is the class preparation subscale).

cence have strong effects on the development of cognitive skills (Cunha, Heckman, and Schennach 2010; Agostinelli and Wiswall 2020). Both the difference in the stage of development (high school vs childhood) and differences in the set of measurements used to identify noncognitive skills may explain the discrepancy.

Table K-2: Technology of Skill Formation

	URM	White & Asian
Knowledge(-1)	1.03	1.05
	(0.01)	(0.00)
Noncognitive(-1)	0.00	0.00
	(0.00)	(0.00)
Study, 10 hours/wk	0.07	0.06
	(0.01)	(0.01)
Private School	0.04	0.02
	(0.01)	(0.01)
Free Lunch	-0.10	-0.15
	(0.02)	(0.02)
Student Teacher Ratio	0.00	0.00
	(0.00)	(0.00)
Mother: High School	-0.01	-0.01
	(0.01)	(0.01)
Mother: Some College	-0.02	-0.01
	(0.01)	(0.01)
Mother: Bachelors	-0.03	-0.02
	(0.02)	(0.01)
Mother: Postgraduate	-0.03	-0.02
	(0.02)	(0.02)
Constant	0.18	0.17
	(0.03)	(0.02)

The table displays estimates of parameters governing the technology of skill formation. Study refers to the effect of studying 10 hours per week on next year's skills. Free lunch is measured on a scale from 0 to 1. High school dropout is the omitted education category. SOURCE: (ELS 2002)

Table K-4 displays estimates of the signal-to-noise ratios for all of the measurements in the data. Unlike in Table 5, there are separate signal-to-noise ratios for each measurement's loading onto the cognitive and noncognitive skill dimensions. The signal-to-noise ratios for measurements that load only onto cognitive skills are computed by dividing that measurement's loading onto the first skill dimension by the standard deviation of the shock of that measurement. For measurements that load only onto noncognitive skills, I divide that measurement's loading onto the second skill dimension by the standard deviation of the shock of that measurement. There are two separate signal-to-noise ratios for measurements that load onto both skills (GPAs). Table K-4 shows that GPAs are less informative signals of both cognitive and noncognitive skills for URM students than for white and Asian students, results that echo the findings from the one-skill model in Table 5. The SAT Math exam has a lower signal-to-noise ratio for URM students, but all other standardized exams have higher signal-to-noise ratios for URM students than for white and Asian students. Finally, the noncognitive skill measurements are noisier for URM students than for white and Asian students. These findings support the analysis in

Table K-3: Estimates of $\boldsymbol{\mu}_t^R$

	URM	White & Asian	Difference
GPA, 9th grade	-0.15	-0.15	0
	(0.03)	(0.03)	(-)
GPA, 10th grade	-0.23	-0.25	0.01
	(0.04)	(0.03)	(0.03)
GPA, 11th grade	-0.35	-0.33	-0.02
	(0.03)	(0.03)	(0.03)
GPA, 12th grade	-0.47	-0.33	-0.14
	(0.03)	(0.02)	(0.03)
SAT Math	-0.97	-1.02	0.05
	(0.03)	(0.04)	(0.04)
SAT Verbal	-0.88	-0.88	0.00
	(0.03)	(0.03)	(0.03)
NCES Math, 10th grade	-0.40	-0.40	0
	(0.03)	(0.03)	(-)
NCES Read, 10th grade	-0.34	-0.32	-0.03
	(0.03)	(0.03)	(0.02)
NCES Math, 10th grade	-0.40	-0.40	0
	(0.03)	(0.03)	(-)
Eng Self-Efficacy	0.28	-0.17	0.45
	(0.12)	(0.03)	(0.12)
Math Self-Efficacy	0.22	-0.15	0.37
	(0.11)	(0.03)	(0.10)
Writing Ability	-0.09	-0.01	-0.08
	(0.05)	(0.02)	(0.05)
Control Expectation	0.33	-0.20	0.54
	(0.14)	(0.04)	(0.13)
Action Control	0.36	-0.21	0.57
	(0.14)	(0.04)	(0.14)
Motivation	0.31	-0.18	0.49
	(0.12)	(0.03)	(0.12)
Preparation	-0.06	0.00	-0.05
	(0.04)	(0.02)	(0.05)
Math Self-Efficacy (12th Grade)	0.05	-0.06	0.11
	(0.05)	(0.02)	(0.05)

The table displays estimates of $\boldsymbol{\mu}_t^R$ in equation (K-2). SOURCE: (ELS 2002)

the main text and suggest that colleges will struggle to identify highly skilled individuals (whether in terms of cognitive or noncognitive skills) in the absence of SAT scores.

The dynamic factor model that I estimate in this section allows for an individual-specific correlation between cognitive and noncognitive skills. This correlation varies across individuals, because it is parameterized as a function of the same covariates that govern the initial mean and variance of skills (Table K-1). Figure K-1 depicts the density of estimated correlation coefficients across individuals in the sample. The mean is 0.52

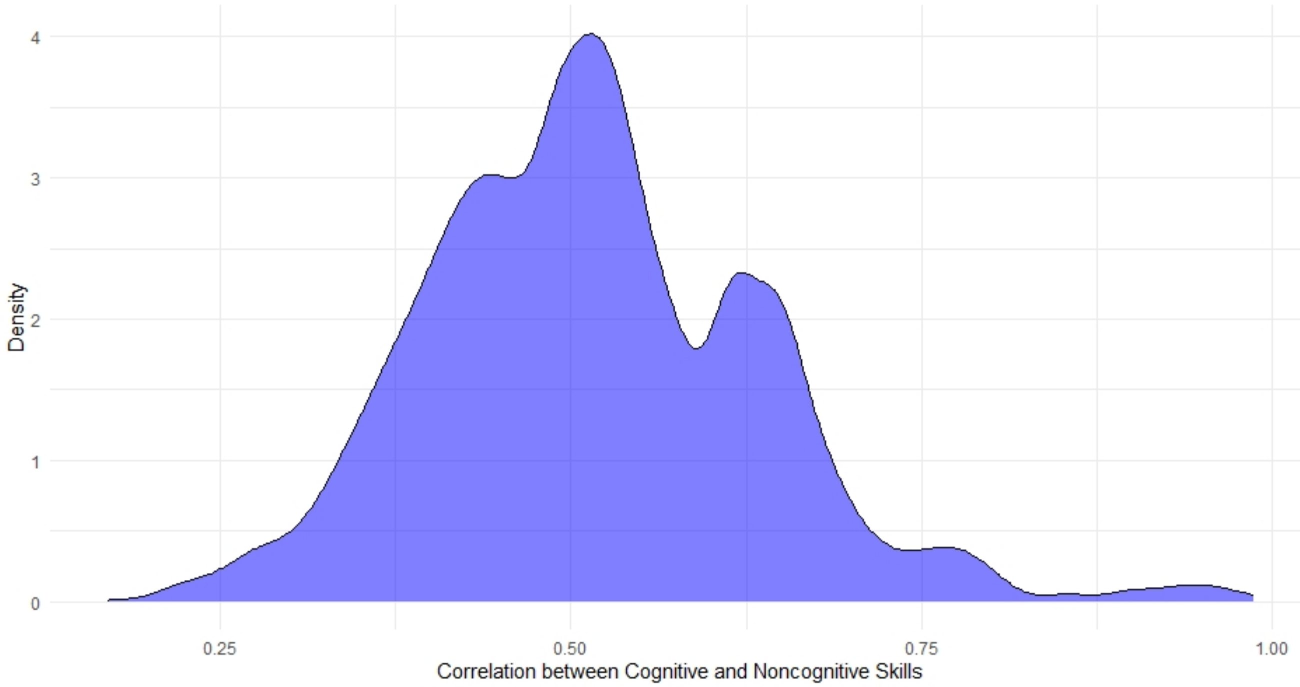
Table K-4: Signal-to-Noise Ratios ($\frac{\alpha_{t,j}}{\sigma_{t,j}}$)

	URM		White & Asian		Difference	
	$K^{(1)}_{i,12}$	$K^{(2)}_{i,12}$	$K^{(1)}_{i,12}$	$K^{(2)}_{i,12}$	$K^{(1)}_{i,12}$	$K^{(2)}_{i,12}$
GPA, 9th grade	0.68	0.13	0.86	0.23	-0.17	-0.09
	(0.05)	(0.06)	(0.03)	(0.02)	(0.05)	(0.06)
GPA, 10th grade	0.65	0.17	0.81	0.27	-0.17	-0.10
	(0.04)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
GPA, 11th grade	0.56	0.13	0.69	0.19	-0.13	-0.07
	(0.04)	(0.02)	(0.03)	(0.02)	(0.04)	(0.03)
GPA, 12th grade	0.43	0.08	0.50	0.16	-0.08	-0.07
	(0.03)	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)
SAT Math	1.73	0	1.95	0	-0.23	0
	(0.1)	(-)	(0.06)	(-)	(0.1)	(-)
SAT Verbal	1.27	0	1.22	0	0.05	0
	(0.08)	(-)	(0.04)	(-)	(0.08)	(-)
NCES Math, 10th grade	2.54	0	2.04	0	0.49	0
	(0.15)	(-)	(0.06)	(-)	(0.13)	(-)
NCES Reading, 10th grade	1.15	0	1.07	0	0.08	0
	(0.07)	(-)	(0.03)	(-)	(0.06)	(-)
NCES Math, 12th grade	2.41	0	2.18	0	0.23	0
	(0.14)	(-)	(0.07)	(-)	(0.14)	(-)
English Self-Efficacy	0	0.81	0	0.95	0	-0.13
	(-)	(0.12)	(-)	(0.03)	(-)	(0.12)
Math Self-Efficacy	0	0.68	0	0.90	0	-0.22
	(-)	(0.10)	(-)	(0.03)	(-)	(0.11)
Writing Ability	0	0.20	0	0.39	0	-0.19
	(-)	(0.03)	(-)	(0.02)	(-)	(0.04)
Control Expectation	0	1.39	0	1.76	0	-0.38
	(-)	(0.20)	(-)	(0.06)	(-)	(0.20)
Action Control	0	1.49	0	1.65	0	-0.17
	(-)	(0.21)	(-)	(0.05)	(-)	(0.21)
Motivation	0	0.93	0	1.06	0	-0.13
	(-)	(0.13)	(-)	(0.04)	(-)	(0.13)
Preparation	0	0.13	0	0.22	0	-0.09
	(-)	(0.03)	(-)	(0.01)	(-)	(0.03)
Math Self-Efficacy, 12th Grade	0	0.25	0	0.37	0	-0.13
	(-)	(0.04)	(-)	(0.03)	(-)	(0.05)

The table displays estimates of signal-to-noise ratios, $\frac{\alpha_{t,j}}{\sigma_{t,j}}$, for all measurements in the data. GPAs load onto both skill dimensions, standardized tests load only onto the first skill dimension, and the third set of measurements load only onto the second skill dimension. SOURCE: (ELS 2002)

and the standard deviation is 0.12. This means that cognitive and noncognitive skills are quite highly correlated for the average individual in the sample. A high correlation between the two skills explains why noncognitive skills, as measured by the noncognitive subscales in the ELS 2002, do not appear to have significant effects on college completion

Figure K-1: Correlation Coefficients between Cognitive and Noncognitive Skills



The figure shows the distribution of correlation coefficients between cognitive and noncognitive skills across individuals in the ELS 2002. The correlation coefficients are allowed to vary by URM status, gender, grade retention status, a single parent dummy, mother's education, and household income. The mean correlation coefficient is 0.52 and the standard deviation is 0.12.

at any college (Table K-5). The point estimates for the effects of cognitive skills on college completion in Table K-5 are similar to the estimates from the one-factor model in Table 10. As before, there is no statistically significant college completion penalty for URM students after controlling for where students matriculate to college and their skills at the time of matriculation.

The findings from the two-skill dynamic factor model are reassuring for two reasons. First, the finding that noncognitive skills do not matter for college completion conditional on cognitive skills suggests that the main findings from the paper on how banning the SAT would affect completion rates are not biased by omitting these skills. Given the null results found here, it is not obvious that colleges would want to place substantial weight on noncognitive skills as part of their admissions policies, since this would entail less weight on cognitive skills and thus lower rates of completion for matriculating students. It is entirely possible that there are other noncognitive skills, which are not well-measured in the ELS 2002, that affect college completion. However, incorporating them in the analysis would require additional information that is not available in the data sources used in this paper.

Table K-5: College Completion Model

	Estimate	Standard Error
Tier 1	-0.26	0.42
Tier 2	0.07	0.16
Tier 3	-0.13	0.10
Tier 4	0.15	0.23
Tier 5	-0.02	0.15
Tier 6	-0.33	0.07
$\log K_{i,12}^{(1)} \times$ Tier 1	0.31	0.16
$\log K_{i,12}^{(1)} \times$ Tier 2	0.27	0.08
$\log K_{i,12}^{(1)} \times$ Tier 3	0.26	0.06
$\log K_{i,12}^{(1)} \times$ Tier 4	0.18	0.10
$\log K_{i,12}^{(1)} \times$ Tier 5	0.27	0.07
$\log K_{i,12}^{(1)} \times$ Tier 6	0.35	0.04
$\log K_{i,12}^{(2)} \times$ Tier 1	0.00	0.10
$\log K_{i,12}^{(2)} \times$ Tier 2	0.00	0.06
$\log K_{i,12}^{(2)} \times$ Tier 3	0.05	0.05
$\log K_{i,12}^{(2)} \times$ Tier 4	0.01	0.07
$\log K_{i,12}^{(2)} \times$ Tier 5	-0.01	0.06
$\log K_{i,12}^{(2)} \times$ Tier 6	0.04	0.03
URM	-0.01	0.06
HH Income	0.03	0.01
Mother has College Degree	0.07	0.04

The table displays parameters of the college completion model with two skill dimensions. The constant and slopes with respect to cognitive and noncognitive skills are allowed to vary by college tier. Income is measured in hundreds of dollars per week. SOURCE: (ELS 2002)

This section has also shown that inferences on the bias and informativeness of grades and the SAT are robust to the inclusion of noncognitive skills. Both the one-skill and two-skill models provide little evidence that the SAT is biased against URM students. Both models also show that grades are less informative signals of all students' cognitive skills than are standardized tests and that grades are less informative for URM students than they are for white and Asian students. An absence of bias on the exam coupled with the relatively low informational content of grades suggests that banning the SAT will harm highly skilled URM applicants holding university preferences for race constant. In addition, since URM students enter high school at a substantial disadvantage in noncognitive skills, there seems to be little evidence that a model that gives universities preferences over two dimensions of skill together with demographics would generate substantially different conclusions on the effects of banning the SAT than the main model in the paper.

It is possible that the ELS 2002 lacks the "right" noncognitive measurements to analyze their influence on college-going and college completion. If future data sets have high-quality measures of extracurricular skill (say, athletic or musical skill), an important extension of the model would allow universities to have preferences over cognitive skills, diversity, and talent in sports, art, or music. As the ability to invest in these extracurriculars is positively correlated with income, a greater reliance on these skills in admissions may provide students from richer households with another way of gaining admission in the absence of the SAT.