THE JOURNAL OF HIGHER EDUCATION
https://doi.org/10.1080/00221546.2021.1971488

Admitting Students in Context: Field Experiments on Information Dashboards in College Admissions

Michael N. Bastedo[a], D'Wayne Bell[b], Jessica S. Howell[c], Julian Hsu[d], Michael Hurwitz[c], Greg Perfetto[c], and Meredith Welch[e]

[a]Center for the Study of Higher and Postsecondary Education, Ann Arbor, Michigan, USA; [b]Graduate School of Education, Harvard University, Cambridge, Massachusetts, USA; [c]Research Division, College Board, Washington, District of Columbia, USA; [d]Customer Trust and Partner Services, Amazon.com, Bellevue, Washington, USA; [e]Department of Policy Analysis and Management, Cornell University, Ithaca, New York, USA

ABSTRACT
In this paper, we address whether the provision of contextual information about where students live and learn has the potential to expand postsecondary opportunities to students from disadvantaged neighborhoods and high schools. To examine this question, we describe the results of field experiments with admissions officers working in eight universities who re-read real applications from previous admissions cycles with a dashboard of contextual data about the applicant's neighborhood and high school. In this low-stakes context, admissions officers from institutions utilizing holistic admissions practices were more likely to recommend admitting low-SES applicants when provided with contextual data. The experiment also primed admissions readers to treat students from highly disadvantaged high school and neighborhood contexts more favorably relative to the results of the high-stakes official read, even when the dashboard was not shown to study participants. The results of this experiment suggest that contextualized data can improve equity in admissions, but fidelity to holistic admissions practices is crucial.

ARTICLE HISTORY
Received 24 September 2020
Accepted 19 August 2021

KEYWORDS
College admissions; contextual achievement; field experiment; holistic admissions; low-income students

Access to high-quality education is highly stratified across American high schools. Although there has been substantial progress in recent years, low-income students still have reduced access to the most rigorous high school curricula, such as Advanced Placement and International Baccalaureate (Bastedo et al., 2016; Kolluri, 2018; Perna et al., 2015; Rodriguez & McGuire, 2019). Many states have provided substantial incentives to increase access to rigorous coursework, yet enormous differences among high schools by socio-economic status remain (Conger et al., 2009; Jeong, 2009; Klugman, 2013). Rigorous high school curricula comprise only one element of high schools that are privileged by admissions officers. Underrepresented students are disadvantaged through other elements of the college application process, including early

CONTACT Michael N. Bastedo bastedo@umich.edu

admissions (Avery & Levin, 2010), legacy preferences (Card, 2017; Espenshade et al., 2004; Hurwitz, 2011), extracurricular activities (Weininger et al., 2015; Weis et al., 2014) — nearly every element of the holistic review process (Bastedo, 2016). These effects are exacerbated by racial, ethnic, and socioeconomic segregation of high school enrollments created by discriminatory federal, state, and local housing policies (Reardon & Bischoff, 2011), and overreliance on feeder high schools that are disproportionately wealthy and white (Salazar et al., 2021; Wolniak & Engberg, 2007).

A highly stratified and segregated high school system is a difficult context to create fair college admissions practices. To adapt to this system, admissions officers have adopted a holistic philosophy of college admission, where the credentials presented are evaluated in the context of the opportunities available in the applicant's family and high school (Bastedo et al., 2018; Gaertner & Hart, 2013; Lucido, 2015). Despite these practices, college enrollments remain highly stratified as well, with less than 5% of students at highly selective colleges coming from low-income backgrounds (Bastedo & Jaquette, 2011; Chetty et al., 2020) and a similar lack of progress has been made in the enrollment of students of color (Posselt et al., 2012; Saenz et al., 2007). Low-income applicants are also more likely than wealthier students to undermatch, enrolling in a less-resourced college than would be predicted by their credentials (Black et al., 2015; Roderick et al., 2011; Smith et al., 2013). This is a source of significant public frustration both among government officials and in the media (Leonhardt, 2017).

One avenue for exploring this lack of progress is examining holistic admissions practices directly, through experimental research with admissions officers. Recent work demonstrates that admissions officers' evaluations are strongly impacted by the quality of contextual information provided to them. When 311 admissions officers participated in a randomly controlled experimental simulation, those who had more detailed information on high school contexts were 26–28% more likely to admit a low-income student (Bastedo & Bowman, 2017). Admissions officers who espoused a holistic view of admissions practices were disproportionately likely to admit a low-income student (Bastedo et al., 2018), and those who worked at their alma mater were significantly less likely to do so (Bowman & Bastedo, 2018).

This research provided initial insights into the admissions decision-making process, but a significant limitation was the use of simulated applications that were artificially constructed by the researchers. It remains unknown whether the results in an experimental simulation would hold in real-world contexts with applications drawn from an admissions officer's own university. There are many ways in which admissions files vary among students and high schools, so a larger study using recent real-world applications would provide a stronger foundation for understanding the impact of contextual information on admissions decision making using real applications.

To address these limitations in the prior research, we recruited admissions staff at eight selective universities to participate in a randomly controlled experiment using actual applications to their own campuses. Admissions officers were assigned recent applications, some of which randomly included an Environmental Context Dashboard ("the Dashboard") in a treatment condition.[1] Based on nationally normed data, the Dashboard provides contextual information on both the high school and neighborhood environments, as well as an overall summative metric that averaged the high school and neighborhood challenge measures. Each application was previously read during the college's normal high-stakes review process, enabling us to examine changes in admissions outcomes from the official read, as well as between experimental control (original application portfolio) and treatment (original application portfolio plus the Dashboard) reads. Our main research questions are enumerated below.

RQ 1: Compared to the actual admissions recommendations and ratings in a high-stakes environment, are overall environmental context and contextual SAT scores more influential in shaping admissions recommendations and ratings in this study's experimental setting?

RQ 2: Does the contextual data in the Dashboard influence admissions decisions at familiar (feeder) high schools?

RQ 3: Does the Dashboard contextual information have a larger impact on admissions recommendations at colleges with a more holistic evaluation approach, which tend to have lower acceptance rates and rely heavily on non-academic factors in the admissions recommendations process?

From our experimental analyses, we gleaned two insights into how the Dashboard shifts admissions decisions. First, we found that a student's level of environmental challenge meaningfully shifted admissions recommendations in a manner that favored students from more challenging contexts, and that this impact was concentrated among colleges with a more holistic approach to college admissions. Our second experimental finding is that the Dashboard presentation of a student's SAT scores, in the context of her high school peers, shifted admissions recommendations in a manner that favored students who outperformed their high school peers by the widest margins. This effect was concentrated among students from non-feeder high schools that typically send few applicants to the sampled colleges. Collectively, these experimental findings suggest that greater equity in college enrollment can be achieved through the provision of high school and neighborhood contextual data.

An unexpected discovery from these experiments is that admissions officers were more likely to recommend admission for applicants from more disadvantaged contexts in the experimental read than they were in the previous high-stakes read, even when the Dashboard was unavailable to them. Participating in the experiment appeared to prime admissions officers to act on information already in the application that was indicative of the applicant's contextual backgrounds, such as prior knowledge of the high school, personal statements, and letters of recommendation. As a result, we find that providing additional contextual data can be a significant influence on admissions decision making, but we should not discount the importance of existing admissions officer knowledge, training and norming practices, and admissions philosophy and priorities.

Conceptual framework

Inequity in educational opportunities and college application behavior

Educational inequities leading up to college continue pose enormous challenges for American youth. Even after the formal dismantling of Jim Crow era laws and the Brown v. Board of Education decision of 1954 striking down "separate but equal" policies, school segregation by race has actually *accelerated* over the past several decades (Frankenberg et al., 2019). In addition to school segregation along racial/ethnic lines, there continues to be school segregation by income and family wealth (Owens, 2018; Reardon, 2011). Though Black-White achievement gaps persist, they have stabilized since 1980. By contrast, achievement gaps between high- and low-income students have continued to grow and are now about 66% larger than Black-White achievement gaps alone (Reardon, 2011).

The promise of postsecondary education is often thought of as an antidote to the accumulative challenges facing students of color and low-SES students. However, where low-SES students attend college makes a difference in many outcomes traditionally used to gauge student success, including bachelor's degree completion rates (Bowen & Bok, 1998; Goodman et al., 2017), future income (Dale & Krueger, 2011; Smith et al., 2020), contribution to home communities (Bowen & Bok, 1998), access to elite job markets (Rivera, 2011), and even long-term financial health (Chetty et al., 2020). There is also strong evidence that low-SES students are more likely to attend less selective colleges than those they have the ability to attend, with lower graduation rates and lower future income potential (Bastedo & Flaster, 2014; Bowen et al., 2009; Engberg & Wolniak, 2010). In addition, "micro-barriers" such as application fees, completing a college preparatory curriculum, and admissions essays may prevent students from applying to academic-match colleges (Holzman et al., 2019; Smith et al., 2015). Both low-income students and students of color assess signals of campus diversity, belonging, and inclusion in deciding where to apply to colleges (Holland, 2014; Jack, 2019; Slay, 2017).

There has been much attention garnered from successful efforts to induce students to apply to more selective colleges (e.g., Dynarski et al., 2021; Hoxby & Turner, 2013), though the scalability of such efforts remains an open question (Furquim & Glasener, 2017; Gurantz et al., 2021). Indeed, contemporary efforts to ensure that lower-income students are mimicking their highest-income peers in college application and enrollment have shown that remedying the socioeconomic stratification of students requires more than simple nudges (Oreopoulos & Petronijevic, 2019). Higher touch initiatives such as virtual advising (Gurantz et al., 2020) show some promise, but these efforts are obviously costlier than student mailers containing college suggestions and application fee waivers that were the mainstay of earlier efforts.

Correcting differential application patterns is complicated by the tendency of colleges to focus their recruitment efforts on more affluent high schools, many of which serve as traditional feeders (Salazar et al., 2021). Even if college representatives were to expand outreach efforts to uncharted territories, there is no guarantee that such initiatives would translate into meaningful shifts in the college enrollment behaviors of traditionally underserved students. First-generation students are less likely to attend college representative visits than non-first-generation students, and troublingly, they also appear to be more likely to be seduced by "instant admissions" visits by less-selective colleges offering students "admission on the spot" (Holland, 2014). Low-SES students also do not seem to be influenced by federal government information efforts such as the College Scorecard (Hurwitz & Smith, 2018).

Thus, there is an emerging consensus that information-based interventions and nudges seem unlikely to create substantial shifts in student behavior (Oreopoulos & Petronijevic, 2019). Under the best of circumstances, shifting the application behavior of students is no easy feat (McDonough, 1997; Weis et al., 2014), and the profound socioeconomic stratification that currently exists in the postsecondary sector cannot be remedied through efforts to transform the application behavior of underserved students. Far more work is needed to examine the challenges facing admissions professionals, and how data can be brought to bear to help them achieve shared goals of admitting and enrolling students who have faced greater environmental challenges in their neighborhoods and schools. This would require a shift in thinking from information interventions and nudges for applicants and families, to choice architectures that shapes the thinking and decisions of admissions officers (Thaler et al., 2013). This move also shifts away from a deficit-oriented perspective where low-SES applicants are a problem to be solved, and toward admissions officers with a mission to provide equitable access.

Reimagining the college admissions process

Over the past two decades, college application volume has increased markedly, and the 2020–21 cycle appears to be no exception. For fall 2021, the Common Application is reporting application increases of approximately 10% over the previous year (Jaschik, 2021), and surges are even larger at the nation's more selective colleges and universities, while community college enrollment has fallen dramatically (Anderson, 2021). Despite the overall increases to already saturated application pools, preliminary estimates are also showing declines in applications submitted by low-income and first-generation students. Application pools swollen by traditionally advantaged students means that, absent a concerted effort to recruit and admit disadvantaged students, already-existing gulfs in postsecondary educational opportunities will expand even further.

Even as the COVID-19 pandemic has squeezed university budgets, admissions leaders say they have increased their recruitment of traditionally disadvantaged students, including students of color and first-generation students (Jaschik & Lederman, 2020). This expressed commitment to equity in admissions suggests that admissions leaders are willing to act on the information provided through the Environmental Context Dashboard to expand postsecondary opportunities for disadvantaged students. One means by which admissions officers can level the playing field for low-SES applicants is through "whole context" holistic review, which evaluates applicants' credentials in the context of the opportunities available to them (Bastedo et al., 2018). Contextualized holistic review thus requires high-quality data on high schools and neighborhoods so that the achievements of individuals (provided on transcripts, resumes, and score reports) can be considered in light of similar students from their high schools and neighborhoods.

This is necessary due to well-known human biases that influence all decision processes, and college admissions is no exception. In particular, with respect to this study, the Environmental College Dashboard seeks to mitigate correspondence bias (sometimes called the fundamental attribution error), a well-known tendency for even experts to make attributions that overweight individual personality, motivation, or dispositions rather than the contexts or opportunities available to those individuals (Gilbert & Malone, 1995). These biases translate easily into the holistic admissions process, which seeks to evaluate credentials in context, but where admissions officers may struggle to do so without robust and salient data at the necessary moments in the reading process (Bastedo & Bowman, 2017).

These potential biases are likely exacerbated by the heavy workloads facing college admissions readers. During the height of the admissions season, many readers are evaluating upwards of 100 applications per week in addition to other assigned duties (Bowman & Bastedo, 2018), which increases the cognitive load of each admissions officer, which in turn increases the effects of known decision-making biases. To both increase data availability and reduce cognitive load, the Environmental Context Dashboard provides these needed data points on high schools and thereby reduces the likelihood of negative correspondence bias effects.

Data

The Environmental Context Dashboard

The Environmental Context Dashboard is an admissions tool developed by The College Board that draws upon various large-scale datasets to present contextual information about applicants' high school and neighborhood environments. The version of the Dashboard seen by study participants is presented in Appendix Figure A1, and Appendix A contains detailed information on how these measures are constructed. In this paper, we focus primarily on the "overall" adversity measure, which we subsequently refer to as environmental context, in the upper right corner of the Environmental Context Dashboard. We also focus on the SAT context measure, which appears directly below the overall environmental context measure, and presents the student's own SAT score alongside the interquartile range of scores earned by other students at the student's high school.

The genesis of the Environmental Context Dashboard began in the 2014–2015 academic year with Bastedo and Bowman's (2017) experiment, which engaged 311 admissions professionals, and demonstrated that admissions recommendations favored low-SES applicants when high school contextual information was salient. The goal of the Dashboard was to provide systematic and consistent data on where students live and learn with the goal of simplifying and advancing holistic review processes. Additionally, the Dashboard streamlined review processes by obviating the independent collection of such data by the admissions readers. Since this experiment, the Dashboard has been rebranded by The College Board as Landscape. In the 2020–21 academic year, more than 180 colleges, universities, and scholarship organizations used Landscape to make high-stakes admissions decisions.

The key components of the Dashboard are Overall Environmental Context (called the "Overall Adversity Index" on the Dashboard) and contextualized SAT score.[2] The Overall Environmental Context is a summative metric that averages independent contextual measures calculated for the student's high school and their physical neighborhood. This measure summarizes more than a dozen data points from the American Community Survey, FBI, and the College Board data capturing neighborhood-level data on: 1) undermatch risk; 2) crime risk; 3) family stability; 4) educational attainment; 5) housing stability; 6) median family income. The data in this section are normed against a national population of College Board test takers presented as a percentile from 1 (lowest environmental challenge) to 100 (highest environmental challenge). The contextual Dashboard components are also shown on the panels in the lower quadrant of Appendix Figure A1. In these charts, the darker green shades represent the least environmental challenge and red areas represents more highest levels of environmental challenge.

The Dashboard also shows the student's SAT in context of prior students attending the high school (using the same population definition), obtained directly from College Board data. The score in context is defined as the difference between the student's SAT score (or concorded ACT score) and the 75th percentile score calculated for that high school. We chose the 75th percentile because of the selective nature of the participating colleges. These components contextualized students' academic and non-academic achievements. We include a more detailed examination of the Dashboard in Appendix A.

Participating universities and admissions outcomes

We recruited admissions staff from eight universities who were willing to reevaluate applications from past admissions cycles.[3] Each university was promised confidentiality in exchange for participating in the study. These eight universities span sectors, selectivity level, and size, as shown in Table 1. Their acceptance rates range from below 20% to above 40%, and yield on admitted students also varied from less than 30% to more than 50%. Of the eight colleges, three are public and five are private. Undergraduate enrollment ranges from fewer than 5,000 students to more than 20,000 students.

Consistent with the high level of variation in admissions practices nationally (Bowman & Bastedo, 2018; Clinedinst, 2019), the admissions processes differ across our participating universities. For example, at some universities, applications are reviewed only by individual readers, and some universities have committees review some or all applications. Additionally, some universities award summary ratings based on different components of the application; others have no such rating system. For our purposes, we condensed the variety of admissions measures into two main outcomes: whether the student was initially recommended for admission in the review process and total standardized admission ratings.[4] In this study, admissions officers at sampled colleges re-read between 848 and 4,698 historical applications, and nearly every reviewer read between 100 to 400 applications. This is a substantial number of applications per reader, but fewer than would be read in an entire admissions cycle.

At six of our eight participating universities, readers provided admission ratings based on academic and/or personal characteristics. The scales of these ratings systems differed across colleges, so we separately normalized the academic and personal admissions ratings at each university to have a mean of zero and a standard deviation of one. We then added these normalized metrics together to obtain the overall admissions rating, which we then also normalized at the college level. Variation in available outcomes changes our sample based on the outcomes of interest. For example, Table 1 shows that two of the large, public universities in our study did not provide any admission ratings because assigning such ratings was not their standard practice.

Table 1. Descriptive statistics of participating universities.

	University 1	University 2	University 3	University 4
From IPEDS Fall 2016 Survey				
Sector	Private	Private	Public	Private
Total Undergraduate Enrollment	5,000–9,999	1,000– 4,999	More than 20,000	5,000–9,999
Black and Hispanic Enrollment	More than 10%	More than 10%	Less than 10%	More than 10%
Acceptance Rate	Less than 20%	Less than 20%	20% to 40%	Less than 20%
Yield on Admitted Students	Greater than 50%	Greater than 50%	Less than 30%	Greater than 50%
College Entrance Examination Requirement	Required	Required	Required	Required
For Students Applying to these Universities for Fall 2016 Admission				
Average Overall Environmental Context	28.22	27.39	22.08	29.81
Average Contextual SAT Score Difference	118.95	218.69	−29.94	51.54
For Students Re-read in the Experimental Read Sample Applying for Either Fall 2017 or Fall 2018 Admission				
N of Students	2440	2600	1733	2067
Percent Read with the Dashboard in the Experimental Read	74.92%	50.00%	84.07%	83.99%
Provided Expermental Admission Decisions	x	x	x	x
Provided Experimental Admission Rating		x		x
Provided Experimental Personal Admission Rating				

	University 5	University 6	University 7	University 8
From IPEDS Fall 2016 Survey				
Sector	Public	Private	Private	Public
Total Undergraduate Enrollment	10,000–19,999	1,000–4,999	5,000–9,999	More than 20,000
Black and Hispanic Enrollment	Less than 10%	More than 10%	More than 10%	More than 10%
Acceptance Rate	Greater than 40%	Greater than 40%	Less than 20%	Greater than 40%
Yield on Admitted Students	Less than 30%	Less than 30%	Between 30%-50%	Between 30%-50%
College Entrance Examination Requirement	Required	Required	Required	Required
For Students Applying to these Universities for Fall 2016 Admission				
Average Overall Environmental Context	21.12	37.17	25.59	31.23
Average Contextual SAT Score Difference	42.52	74.01	173.96	69.29
For Students Re-read in the Experimental Read Sample Applying for Either Fall 2017 or Fall 2018 Admission				
N of Students	4698	2100	848	1800
Percent Read with the Dashboard in the Experimental Read	83.35%	85.00%	90.57%	80.00%
Provided Expermental Admission Decisions	x	x	x	
Provided Experimental Admission Rating		x	x	x
Provided Experimental Personal Admission Rating			x	x

Overall Environmental Context is a percentile, where a higher percentile represents a more disadvantaged environment. The contextual SAT difference is the student's SAT score subtracted from their high school's 75th percentile SAT. If the SAT score was on the 2400 scale, or if we only had ACT scores, we concorded them to the 1600-SAT scale. The 2400-SAT to 1600 SAT Concordance Table is available here: https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdf. The ACT-SAT Concordance Table is available here: https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdfhttps://collegereadiness.collegeboard.org/pdf/guide-2018-act-sat-concordance.pdf.
concordance.pdfhttps://collegereadiness.coll Admission ratings and decisions are for students we determined to be on the margin of admission for the Fall 2016 or 2015 admission cycle within a college. Official and experimental admission ratings are standardized by college, and we drop international applicants and applicants whose files are not read in the experiment.

These universities also provided historical applicant data, which we use to identify feeder high schools. Using universities' historical applicant data, we calculated the annual average number of applications from each high school to a university.[5] We classified high schools as "feeder high schools" to a particular university if the high school sent more than the median number applications and as "non-feeder high schools" in all other cases.[6] This allows us to determine whether, as we hypothesized, the Dashboard information has greater utility for students attending high schools with which admissions staff have the least familiarity.

As previously noted, Table 1 shows the diversity in participating universities. The variation in selectivity is also re-emphasized by the SAT contextual difference, which represents the average score difference between a student's individual SAT score and the 75th percentile of her high school on the 1600 scale (SAT Math and Evidence-Based Reading and Writing).[7] For example, on average, previous applicants to Universities 2 and 3 scored 219 points higher and 30 points lower, respectively, than the 75th percentile SAT of these students' high schools.

The commonality among all the participating universities is that the average challenge levels among sampled applicants, represented as the overall environmental context, are lower than the national median challenge levels of high school graduates taking College Board exams (PSAT/NMSQT, SAT, or Advanced Placement). In other words, previous applicants at these universities tend to come from less challenging environmental contexts than the typical high school graduate.[8] The average overall environmental context of applicants ranges from the 21st percentile at University 5 to the 37th percentile at University 6.

Empirical strategy

This research was executed as a low-stakes randomized-controlled trial evaluating historical college applications. In other words, this experiment is meant to simulate how admissions professionals would make decisions in a high-stakes context, though the experimental decisions have no impact on the actual composition of admitted or enrolling classes. The applications re-read by admissions professionals are from students who applied in previous admissions cycles and who have already decided on which college to attend.

Each reader was assigned a mix of "control applications" that contained the original application materials evaluated in the official read during the regular admissions cycle, and "treatment applications" that contained all of the original materials (transcripts, test scores, recommendation letters, student essays, etc.) plus the addition of the Dashboard. To ensure the integrity of the experimental conditions, we performed a block randomization, which specified that the ratio of control to treatment applications was identical across study participants. This block randomization was conducted using the *survey-select* command in the SAS software package, which is akin to a random number generator that randomizes applications into treatment and control units *within* a reader. In our randomization of applications to reviewers, we further specified that applications assigned to each reader were not reviewed by those readers during the official reviews.

After the initial round of randomization, we conducted balancing tests to ensure that key student covariates including overall environmental context, college entrance examination scores, race and gender were identical between the control and treatment groups, both overall and within readers. Per the advice of Lock Morgan and Rubin (2012), we re-randomized if statistically significant differences (at the 0.05 level) between the control and treatment groups emerged from about 5% or more of the balancing tests. This safeguarded us against the possibility of an unsuccessful randomization. We further elaborate on the covariate balancing below in our subsequent discussion of Table 2.

The share of applications that were assigned to the treatment/control condition was negotiated with the individual colleges, and, as Table 1 shows, ranged from about 50% to 90%. However, we were unwavering in our position that we block randomize applications — randomly assigning applications to readers and randomly assigning control/treatment conditions within readers. Random assignment of the Dashboard allowed us to evaluate how environmental contextual data and contextualized academic achievement influenced admission outcomes. Admission staff from participating universities overwhelmingly found the overall environmental context and SAT contextual differences to be most informative, so we focused our analyses on these two independent variables.[9]

Sampling strategy

For applicants with very high and very low probabilities of admission, the likelihood of additional contextual data changing the admissions decision is very low. Thus, to make our study more relevant to applicants for whom contextual data might influence reader decisions, we sampled from the set of applicants whose academic qualifications indicated eligibility for admission, but omitted students whose academic qualifications signaled either guaranteed admission or a very low probability of admission. This means that the results presented in this study are applicable to the subset of applicants where admissions decisions might be swayed by additional contextual information. We took two main approaches to identifying such applicants. First, we used logit models to estimate individual applicants' predicted probabilities of admission using historical admissions data, and, where available, the characteristics provided to us by colleges that are typically influential in admissions decisions, including race/ethnicity, gender, and standardized test scores. Using these predicted probabilities of admission, we selected applicants with predicted admission

Table 2. Covariate balance of dashboard assignment.

Joint F-Test of Individual Differences Below		Chi2=14.663	P-Value=0.876		
Covariate	Control Mean	Treatment — Control Difference	Standard Error	P-Value	N
Feeder High School	0.647	0.004	0.009	.649	18,229
High School SAT 25th Percentile	1005.365	2.161	2.155	.319	18,075
High School SAT 50th Percentile	1127.661	2.074	2.139	.335	18,075
High School SAT 75th Percentile	1250.279	2.083	2.082	.320	18,075
Average Contextual SAT Score Difference	15.242	−0.442	0.199	.029	17,814
SAT or Converted ACT Score	1378.442	2.290	2.887	.430	18,286
SAT/ACT Scores Available	0.983	0.003	0.002	.066	18,286
Prop. Took SAT	0.674	0.001	0.006	.873	18,286
Prop. Took ACT	0.500	0.001	0.006	.931	18,286
Prop. Female	0.502	−0.011	0.009	.210	18,286
Prop. White	0.464	−0.003	0.009	.773	18,286
Prop. Black	0.089	−0.001	0.005	.879	18,286
Prop. Hispanic	0.132	0.001	0.006	.867	18,286
Prop. Asian	0.265	0.003	0.007	.662	18,286
Prop. Native	0.013	0.000	0.002	.916	18,286
Prop. Multi Race	0.010	0.001	0.002	.391	18,286
Prop. Missing Ethnicity	0.028	−0.002	0.002	.348	18,286
Average Overall Environmental Context	25.608	−0.203	0.323	.531	18,286
Average High School Environmental Context	27.367	−0.332	0.320	.303	18,247
Average Neighborhood Environmental Context	23.756	0.028	0.433	.949	18,031
Prop. Officially Admitted	0.439	−0.011	0.008	.152	18,286
Prop. Officially Rejected	0.205	0.008	0.006	.189	18,286
Prop. Officially Waitlisted	0.356	0.003	0.006	.547	18,286
Official Total Admission Rating	−0.010	0.018	0.021	.405	11,845

All covariate balance tests include college fixed effects. The Total Admission Ratings are standardized at the college level as described in the text. Standard errors are clustered at the experimental reader level.

Environmental Context is a percentile, where a higher percentile represents a more disadvantaged environment.

Average Contextual SAT Score Difference is the student's SAT score subtracted from their high school's 75th percentile SAT. If the SAT score was on the 2400 scale, or if we only had ACT scores, we concorded them to the 1600-SAT scale. The 2400-SAT to 1600 SAT Concordance Table is available here: https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdf. The ACT-SAT Concordance t-concordance.pdfhttps://collegereadiness.coll Table is available here: https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdfhttps://collegereadiness.collegeboard.org/pdf/guide-2018-act-sat-concordance.pdf.

probabilities in an interval around approximately 0.50, with the size of the interval determined by the reading capacity of the admissions officers. At the most selective universities, there were no applicants with predicted probabilities of admission of 50%, so we identified borderline admission applicants by sampling exclusively from applicants waitlisted during the official read. After finalizing the sample, we randomly assigned applications to readers and also randomized which applications contained the Dashboard within readers.[10]

Preparing admissions officers to use the Dashboard

Prior to participating in the study, admissions officers received a brief orientation of the Dashboard, and where possible, were provided with the aggregate admissions rates of sampled students. We show an example of this orientation in Appendix A. The orientation highlights that the Dashboard is designed to help readers with the holistic review process, and to provide new information about environmental context. This orientation did not explicitly draw attention to any specific piece of information, but it did emphasize the importance of understanding applicants' high school and neighborhood contexts. Additionally, admission readers completed a pre-study survey that asked questions about their approaches to holistic admissions review and how environmental context factored into their recent official reviews.[11]

In the orientation documents, we noted that the applicants selected for the study were on the cusp of admission. We asked readers to evaluate these applications using the same criteria and standards from their official application review, but to view the Dashboard information as complementary to existing information in the individual application. Given the emphasis on environmental context throughout the orientation, admission readers may have been pre-disposed to infer applicants' environmental context through other application materials — for example, the high school profile, essays, recommendations or extracurricular activities.

Covariate balance

As described above, we randomized applications to reviewers and we also randomized which applications were accompanied by a Dashboard. Approximately 15% of the original sample of 21,450 applications were not read. There were two sources of sample attrition: 1) readers who were unable to complete reviewing all applications assigned to them and 2) readers who dropped out of the experiment altogether and read none of the applications assigned to them. Of the 15% attrition, approximately half was due to the each of the two sources above. The relatively small amount of attrition did not result in a biased sample. The attrition rate was 13% in the control group and 15% in the treatment group. Once college fixed effects were factored in, the difference in attrition rates was only 0.5 percentage points and was not statistically significant. Even under cautious assumptions, this is a tolerable threat of bias given the low attrition and small differential between the treatment and control conditions (Institute of Education Sciences [IES], 2020).

The absence of differential attrition alone is insufficient to prove that the randomization successfully created identical control and treatment groups. To accomplish this, we conducted a series of covariate balancing tests. Table 2 shows the average of the academic and socio-demographic characteristics for the control group (column 1) and the difference in these averages between the control and treatment groups (column 2).[12] The third and fourth columns express the standard errors, clustered at the reviewer level, and the p-values of the differences between the treatment and control groups.

The p-values in column 4 of Table 2 demonstrate that we successfully assigned students to two statistically equivalent groups. As an additional test of balance, we pool all our differences together to test the null hypothesis that all differences are jointly equal to zero. The p-value of 0.876 associated with the F-test of joint significance fails to reject this null hypothesis, offering further assurance that the treatment and control groups are balanced across student characteristics.[13]

Regression specifications

We use two elements within the Dashboard to measure applicant's environmental context and contextualized academic achievements. We use the continuous overall environmental context measure, on the 1–100 percentile scale, to represent applicants' environmental challenge from their neighborhood and high school, and a continuous measure of the difference between a student's SAT score and their high school's 75th percentile SAT (SAT contextual difference, hereafter). We also present specifications where neighborhood and high school are challenge treated separately in Appendix Table A1.

To determine whether the Dashboard orientation primed readers to prefer applicants from backgrounds with greater challenge or higher contextual academic achievement, we compare the relationship between each of the Dashboard elements (overall environmental context and SAT contextual difference) and admission outcomes, using the following logit regression equation:

$$Log(OddsAdmission_i) = \beta_0 + \beta_1 Context_i + X_i \tag{1}$$

In this model, $Context_i$ represents a continuous element of the Dashboard, either overall environmental context or SAT contextual difference, and the X_i represents a vector of student academic and demographic characteristics,[14] official reader fixed effects, and university fixed effects. The university fixed effects address the previously discussed variation in treatment assignment across universities, and we continue to cluster all standard errors at the reader level. Context exists for all applications in this study, but is only visible to readers in applications assigned to the *Treatment* condition.

The parameter of interest in Equation (1) is β_1, which expresses the relationship between the likelihood of admission and the Dashboard element, Z_i. We separately fit Equation (1) three times. First, we fit Equation (1) using all students from the official read. Second, we consider admissions outcomes for experimental control students, where the contextual information is not visible to the reader. Finally, we consider admissions outcomes for experimental treatment students. In the regressions where the outcome represents the official admissions and the experimental control decisions, the Dashboard attributes can be thought of as concealed student-level characteristics. Though these data exist for all students, they are only visible if the student is in the experimental treatment group.

When we compare our estimates from fitting Equation (1) across all three groups, we can identify whether study participants were primed to shift admissions decisions from standard practice as a result of study participation and the Dashboard orientation. We adopt a different approach to isolate the effects of the Dashboard within the experiment. We leverage the random assignment of the Dashboard within readers to identify the impact of revealing contextual information on admissions outcomes by fitting Equation (2) below. We only fit Equation (2) for the experimental outcomes. This allows us to identify the impact of the Dashboard on admissions decisions above and beyond the priming effect of the Dashboard.

$$Log(OddsAdmission_i) = \beta_0 + \beta_1 Context_i + \beta_2 Treatment_i + \beta_3 Treatment_i * Context_i + \beta_4 X_i \tag{2}$$

In Equation (2), the outcome represents the experimental admission decision and vector X_i includes the admission decision from the official read in addition to the same applicant-level covariates as Equation (1). In this model, if the student was in the experimental treatment and her application was accompanied by the Dashboard, the binary *Treatment* variable equals 1, or else the *Treatment* variable equals 0. Parameter β_2 on the binary Treatment variable indicates whether the likelihood of admission differs between the experimental treatment and control conditions. Finally, the parameter of interest in Equation (2), β_3, on the multiplication of the binary *Treatment* variable and the continuous *Context* variable, indicates the extent to which the relationship between the odds of admission and the continuous Dashboard elements differs between the experimental treatment and control students. In other words, parameter β_3 indicates whether the Dashboard data increase the likelihood of admission for students from high-challenge environments and for students whose SAT scores outperformed those of their high school peers. Equation (2) is the central model of this study and is applied when answering all three research questions enumerated above.

For the outcomes in this study that are not binary, such as total admissions ratings, we modify Equation (2) and replace $Log(OddsAdmission_i)$ with the continuous standardized admission ratings. We then fit the data with standard OLS regression models.

Analysis and results

Research question 1: Compared to the actual admissions recommendations and ratings in a high-stakes environment, are overall environmental context and contextual SAT scores more influential in shaping admissions recommendations and ratings in this study's experimental setting?

In this section, we demonstrate three sets of findings. First, we show that, in both the official and the experimental read, applicants from more challenging environments were more likely to gain admission, controlling for student characteristics. Second, we present evidence that readers were more sensitive to students' environmental context and contextualized academic achievement in the experimental read than in the official read, regardless of whether students were in the treatment or control groups. Finally, we show that the Dashboard-revealed overall environmental context and contextualized academic achievement, influenced both admissions ratings and admissions recommendations.

Table 3 shows parameter estimates from fitting Equations (1) and (2) to data from the official and experimental reads. Columns (i), (ii), and (iii) present the results from fitting Equation (1), and the final column shows the results from fitting Equation (2). We show results for overall environmental context in the top panel and for contextual SAT score difference in the bottom panel.[15]

The log-odds estimates in the top panel of Table 3 show that applicants facing higher levels of overall environmental context (higher challenge) are more likely to gain admission controlling for student socio-demographic characteristics. For example, column (i) shows that, with each percentile point increase in an applicant's overall environmental context, her odds of admission are multiplied by 1.006 in the official read. The odds-ratios estimates of 1.011 and 1.009 in columns (ii) and (iii) show that readers were more sensitive to applicants' environmental context in the experimental read than in the official read, both for the experimental control (column ii) and experimental treatment (column iii) groups.

Table 3. Evidence of priming — admission based on environmental context and contextual SAT score difference.

Reporting log-odds ratios	(i) Admitted in Official Read	(ii) Admitted in Experiment, Control	(iii) Admitted in Experiment, Treatment	(iv) Admitted in Experiment, Treatment/ Control
Overall Environmental Context	1.006***	1.011***	1.009***	1.009***
	[0.002]	[0.002]	[0.002]	[0.003]
Treatment				0.960
				[0.079]
Overall Environmental Context X Treatment				1.000
				[0.003]
N	13849	3538	12,769	16,458
Psuedo-R2	0.315	0.254	0.308	0.459
Average Overall Environmental Context	25.562	25.353	25.804	25.696
Average Admission	0.481	0.406	0.427	0.423
Contextual SAT Score Difference	1.004	1.014***	1.013***	1.013***
	[0.003]	[0.005]	[0.003]	[0.004]
Treatment				0.913
				[0.062]
Contextual SAT Score Difference X Treatment				1.005
				[0.004]
N	13655	3499	12,585	16,229
Psuedo-R2	0.315	0.253	0.308	0.460
Average Contextual SAT Score Difference/10	11.252	16.040	11.235	12.266
Average Admission	0.482	0.407	0.427	0.422

****p*-value < .01, ***p*-value < .05, **p*-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

We find similar results when we replace overall environmental context with SAT contextual difference in the bottom panel of Table 3. As shown in column (i), for every 10-point increase in SAT contextual difference, an applicant's odds of admission are multiplied by 1.004 in the official read. The odds-ratio estimates of 1.014 and 1.013 in columns (ii) and (iii) show that readers were also more sensitive to applicants' relative academic performance in the experimental read than in the official read.

Taken together, these results suggest that when readers were given the Dashboard orientation, they were primed to pay attention to applicants' environmental context when making admissions recommendations in a manner that favored students facing higher levels of environmental challenge. If readers similarly valued environmental context factors in the experimental read as they did in the official read, we would expect the log-odds ratios in columns (i) and (ii) to be the same. We formally test whether the parameter estimates on overall environmental context and SAT contextual difference in column (i) differ from those in column (ii), and find clear evidence that the readers in the control arm of the experiment were weighing environmental context and SAT context more in the decision process than in the official read. The *p*-values below 0.05 show that the estimates in column (i) are different from those in Column (ii) (see Appendix Table A3). We conclude from this finding that admissions readers were able to infer contextual information from other application materials. To rule out the possibility that differences in the composition of readers between the official read and the experimental read are driving the demonstrated priming results shown in Table 3, we separately generate this table for the subset of applications evaluated by readers participating in both the experimental read and the official read (Appendix Table A2). This restriction halves the study sample, yet the priming story remains intact. In fact, using this sub-sample, the differences in parameter estimates on overall environmental context between the experimental control and the official read are even slightly more pronounced.[16] In the final column of Table 3, we show the results from fitting Equation (2). Here, we are most interested in the estimate of β_3 on the interaction terms (overall environmental context x Treatment and SAT contextual difference x Treatment). These estimates reveal whether the presence of the Dashboard for reviewers played a role in shifting admissions decisions beyond the priming documented in the first three columns. In the top panel, the log-odds ratio of 1.000 on the interaction term of overall environmental context and treatment suggests that revealing overall environmental context to readers on the Dashboard did not shift admissions decisions. In contrast, the log-odds ratio of 1.005 on the interaction term of SAT contextual difference and treatment offers suggestive evidence that the Dashboard heightened the likelihood of admission for applicants with higher relative academic performance on the SAT, but this estimate does not reach statistical significance at conventional levels.

In Figure 1, we illustrate the estimates from columns (i) through (iii) as predicted probabilities of admission, holding covariates at their sample means. For the official read (column i), the experimental read control group (column ii), and the experimental read treatment group (column iii), we show how an applicant's probability of admission changes across a range of values of overall environmental context or SAT contextual difference. There are three findings that this figure relays. First, across almost the entire range of context, the probability of admission is higher in the official read than it was in the experimental read. Though it is impossible to identify the cause of this phenomenon, one possibility is that readers have their own institutional acceptance rates in mind when making recommendations, and they are using this heuristic to guide the share of applications recommended for admission in the experiment. As an example of how to interpret the curves shown in Figure 1, if the average applicant's overall environmental context increased from 1 to 99, then her probability of admission in the

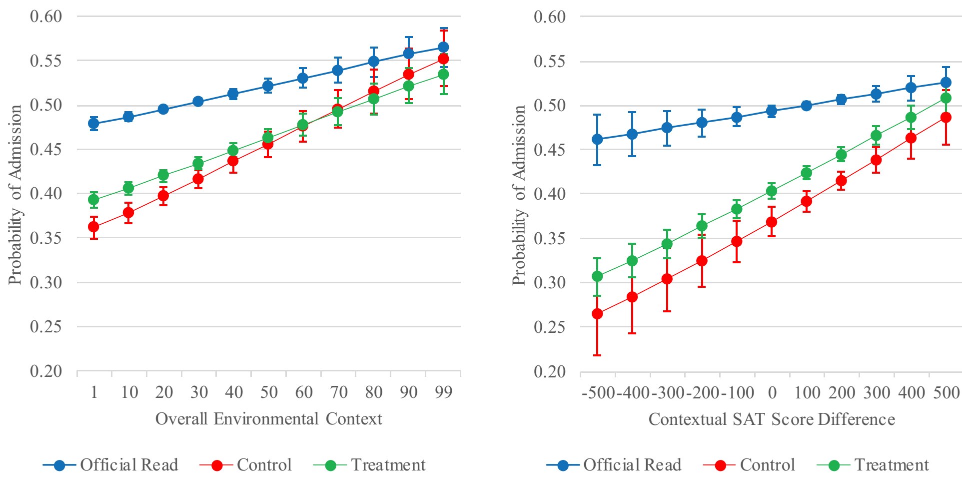


Figure 1. Predicted probabilities of admission across contextual information.
Notes: These figures use Stata's margins command to show students' predicted outcomes across contextual information. 95% confidence intervals are shown as error bars around the average predicted outcome.

official read increases 9 percentage points — from 47% to 56%. More generally, this line shows that each 10 percentile point increase in overall environmental context is associated with a 1 percentage point increase in admissions probability in the official read. Similarly, if the average student's SAT contextual difference increased from −200 to 200, then her probability of admission in the official read increases from 47% to 50%. The second key insight from Figure 1 is what we identify as a priming effect by which readers are putting their thumb on the scale for students facing higher levels of environmental challenge or students who outperformed their peers on the SAT by the widest margins. This is exemplified by the steeper slope for both the treatment and control group in the experimental read compared to the official read. The third key insight is that the overall environmental context or SAT contextual difference through the Dashboard in the experimental setting does not shift the likelihood of admission, on average. This is reflected by overlapping confidence intervals around the estimated probabilities for the lines representing the treatment and control groups in Figure 1. As we discuss in greater depth in the conclusion, the priming impact observed in this experiment may emerge because study participants are not forced to grapple with real-world consequences of admitting larger numbers of low-income students. These consequences might mean an augmented financial aid budget and more academic supports or remediation.

Table 4. Evidence of priming — admissions ratings based on overall environmental context and contextual SAT score differences.

Reporting OLS Regression Coefficients	(i) Official Read Total Ratings	(ii) Experimental Read Total Ratings, Control	(iii) Experimental Read Total Ratings, Treatment	(iv) Experimental Read Total Ratings, Treatment/ Control
Overall Environmental Context	0.001*	0.003**	0.003***	0.003***
	[0.001]	[0.001]	[0.001]	[0.001]
Treatment				0.033
				[0.031]
Overall Environmental Context X Treatment				0.000
				[0.001]
N	11840	2993	8847	11,840
R2	0.158	0.146	0.129	0.275
Average Overall Environmental Context	26.952	26.439	27.126	26.952
Average Admision Rating	0.002	−0.144	0.000	−0.036
Contextual SAT Score Difference	0.000	0.003*	0.003	0.003*
	[0.001]	[0.002]	[0.002]	[0.001]
Treatment				0.029
				[0.034]
Contextual SAT Score Difference X Treatment				0.000
				[0.001]
N	11522	2912	8610	11,522
R2	0.163	0.155	0.129	0.272
Average Contextual SAT Score Difference/10	16.989	19.459	16.153	16.989
Average Admision Rating	0.009	−0.141	0.001	−0.035

***p-value < .01, **p-value < .05, *p-value < .10.
Standard errors are clustered at the experimental reader level.
Five out of the eight universities we partnered with provided admission ratings for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes additional controls of experimental reader fixed effects and original admission outcome. We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Total ratings

When we look at applicants' admission ratings as an outcome, we find similar patterns. In Table 4 we show the parameter estimates from fitting Equations (1) and (2) on applicants' standardized admission ratings. The structure of Table 4 follows that of Table 3, except here we interpret the parameter estimates as increases in admission ratings rather than as log-odds ratios of admission.

The parameter estimates in the top panel of Table 4 show that applicants from environmental contexts with higher levels of environmental challenge are more likely to receive higher admission ratings. This is true in both the high-stakes and the experimental setting. For example, in column (i), we find that for each percentile point increase in overall environmental context, the average applicant received a 0.001 standard deviation higher admission rating in the official read, controlling for our standard demographic and academic controls. The parameter estimates in columns (ii) and (iii) indicate that for each percentile point increase in overall environmental context, the average applicant, in both the experimental treatment and experimental control groups, received a 0.003 standard deviation higher admission rating in the experimental read. We found that the parameter estimates on environmental context column (i) differ at the 0.05 level of significance from those in column (ii) and column (iii) (Appendix Table A5). These larger estimates show that readers were more sensitive to applicants' environmental challenge in the experimental read than in the official read, again suggesting that readers were primed to consider applicants' contexts.

We find similar results when we examine the relationship between SAT contextual difference and admission ratings, shown in the bottom panel of Table 4. As illustrated in column (i), we find that there is no relationship between SAT contextual difference and total admission ratings in the official read. The experimental read tells a different story. Columns (ii) and (iii) show that for each point increase in SAT contextual difference, the average applicant's admission rating increases 0.003 standard deviations in the experimental read.[17]

In the final column of Table 4, we estimate Equation (2) using OLS to evaluate whether revealing contextual information using the Dashboard provides an additional boost in admission ratings for applicants randomly assigned into the experimental treatment group, beyond what we find in column(i). Similar to column (iv) of Table 3, we are interested in the parameter estimate, β_3 on the interaction term between context and treatment. The top panel of Table 4 reveals a parameter estimate of 0.003 on the overall environmental context and 0.000 on the interaction between overall environmental context and treatment. This indicates that applicants from more challenging environments received higher admission ratings on average in the experiment, regardless of whether the reader had access to the Dashboard (treatment students) or not (control students). The exact same story plays out when we focus on the interaction term of SAT contextual difference and the treatment indicator in the bottom panel of Table 4.[18]

We again illustrate estimates from the first three columns of Table 4 as predicted standard deviation increases in total admission ratings (Figure 2). As with our results for admission decisions, the changing relationships between the Dashboard elements and admission ratings across the official read and experimental read are exemplified by the steeper slopes for both the experimental treatment and control groups in the experimental read. Additionally, the increased weight placed on contextual information in the experimental read is indistinguishable between those that had access to the Dashboard and those that did not. This is reflected by the parallel lines reflecting the experimental control and treatment groups. Once again, this finding reinforces that, perhaps through non-Dashboard materials contained in the application, study participants were able to identify students from challenging environments as

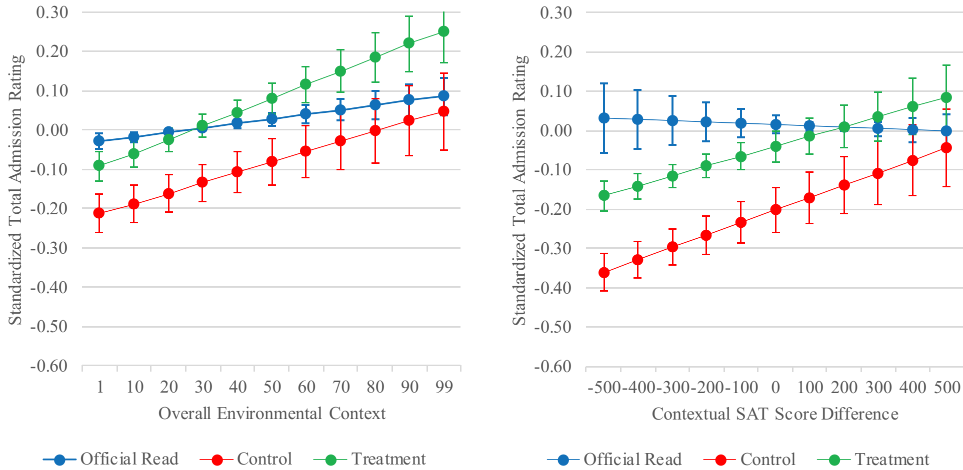


Figure 2. Predicted admission ratings across contextual information.
Notes: These figures use Stata's margins command to show students' predicted outcomes across contextual information. 95% confidence intervals are shown as error bars around the average predicted outcome.

well as students who outperformed peers at their high school on the SAT, and were likely primed to assign these students higher ratings in the application review.

Research question 2: Does the contextual data in the dashboard influence admissions decisions at unfamiliar (non-feeder) high schools?

Assessing an applicant's context through traditional materials alone may be challenging when the application comes from unfamiliar high schools or geographic regions (Hill & Winston, 2010). Colleges hoping for more applicants from these non-feeder high schools may find the Dashboard information particularly useful during the review process. We turn our analysis to the experimental read data to understand how the availability of the Dashboard and contextual information informs admission for students who applied from feeder and non-feeder high schools.

When we examine the impact of the Dashboard separately for applicants from feeder high schools and non-feeder high schools, we find evidence that providing contextual information about an applicant's relative academic performance shifts admissions decisions in favor of applicants with higher relative academic performance (SAT scores in context) in instances where admissions staff may be less familiar with an applicant's academic context. In Table 5, we show parameter estimates on context when we fit Equation (2) to the experimental read data separately for applicants from non-feeder high schools and feeder high schools. We show parameter estimates for SAT contextual

Table 5. Experimental admission recommendation across feeder high schools vs. non-feeder high schools.

Reporting log-odds ratios

	(i)	(ii)	(iii)	(iv)
	Non-Feeder HS	Feeder HS	Non-Feeder HS	Feeder HS
Treatment	0.781**	0.910	1.118	0.944
	[0.094]	[0.066]	[0.163]	[0.077]
Overall Environmental Context			1.012***	1.008***
			[0.004]	[0.003]
Overall Environmental Context X Treatment			0.997	0.998
			[0.004]	[0.003]
Average Overall Environmental Context			32.962	22.108
Contextual SAT Score Difference	0.999	1.010**		
	[0.006]	[0.005]		
Contextual SAT Score Difference X Treatment	1.013**	1.000		
	[0.006]	[0.004]		
Average Contextual SAT Score Difference/10	17.720	9.513		
N	5041	10,906	5200	10,947
Pseudo-R2	0.284	0.340	0.284	0.339
Average Experimental Admission	0.435	0.422	0.434	0.422

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

For each high school, we calculated the average number of applications sent to each college using historical application data provided by colleges. Feeder high schools exceed this median number.

difference in columns (i) and (ii) and parameter estimates for overall environmental context in columns (iii) and (iv). Similar to column (iv) of Tables 3 and 4, we are again interested in the parameter estimate on the interaction terms of both contextual elements and the treatment indicator, which has a value of 1 if the Dashboard accompanied the application, and 0 otherwise.

The statistically significant log-odds ratio of 1.013 in column (i) on the interaction of contextual SAT score differences and treatment indicates that for applicants from non-feeder high schools, the Dashboard induced readers to give more favorable admissions recommendations to applicants who exceed their high school's peer SAT scores by the widest margins. The corresponding estimate of 1.000 on the interaction of contextual SAT score differences and treatment in column (ii) indicates that no such differences in slopes between the control and treatment groups exist for applicants from feeder high schools. Since readers are likely familiar with the typical academic qualifications of students from feeder schools, it is unsurprising that the presence of the Dashboard did not alter recommended admissions decisions for these students.

We show the difference in these interaction terms visually in Figure 3. The Dashboard (treatment) and No Dashboard (control) lines are parallel for applicants from feeder high schools, but they are decidedly not parallel for

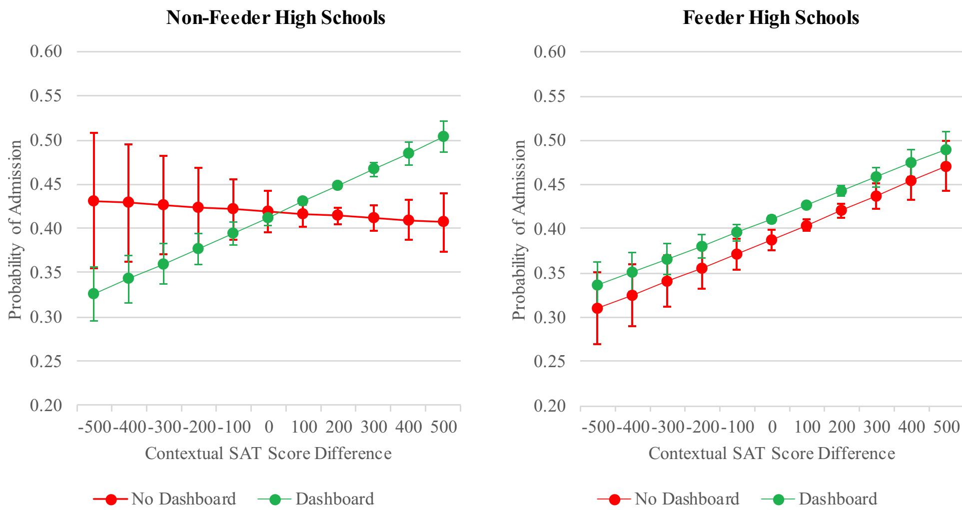


Figure 3. Predicted probabilities of admission across contextual information, by feeder high school status.

Notes: These figures use Stata's margins command to show students' predicted outcomes across contextual information. 95% confidence intervals are shown as error bars around the average predicted outcome.

non-feeder high schools. These intersecting lines confirm that when readers are less familiar with an applicant's high school, the Dashboard tips the admissions scales in favor of students with stronger relative academic performance. For applicants from non-feeder high schools, a 100-point increase in SAT context is associated with a roughly 2 percentage point increase in the probability of admission. Since 100 points on the SAT is equivalent to 0.5 standard deviation (σ) units, another way of expressing this metric is that a 1 σ increase in SAT contextual scores increases the probability of admission by about 4 percentage points.[19]

When we replace SAT contextual difference with overall environmental context, we find that revealing environmental context information in the experimental read does not impact the odds of admission for applicants from non-feeder (column iii) or feeder (column iv) high schools (Table 5). The log-odds ratios of 1.012 and 1.008 on overall environmental context show that both applicants from non-feeder and feeder high schools from more disadvantaged environments are more likely to be admitted when not read with the Dashboard. The log-odds ratios of 0.997 and 0.998 on the interaction terms are not statistically different from 1.000, and therefore indicate that the positive relationship between environmental context and the probability of admission is not distinguishable between the control and treatment groups.

Research question 3: Does the Dashboard contextual information have a larger impact on admissions recommendations at colleges with a more holistic evaluation approach?

When we disaggregate our analyses of the experimental data by universities, we find differing relationships between applicants' environmental context and admission decisions. As only seven universities provided admission decisions in the experimental read — the eighth offered only ratings — it is difficult to definitively group sampled universities into categories representative of all universities. None of the participating universities maintained open-enrollment admissions policies, but there are two clearly defined groups, those with lower admissions rates that employ a less formula-driven and more holistic approach to admissions (universities 1, 2, 4 and 7) and another group where the admissions processes appeared more formula-driven and less holistic. Among our sampled universities, the unpredictability of admissions decisions is linked to their selectivity, and colleges in the more holistic category routinely reject students with excellent academic preparation. We bifurcated the sampled universities to test the hypothesis that the Dashboard had more potential to shift admissions decisions in colleges less reliant on formula-driven approaches in their decision processes.

In Table 6, we show parameter estimates when we fit Equation (2) to the experimental read data separately for universities using more and less holistic admissions approaches.[20] We show parameter estimates for SAT contextual difference in columns (i) and (ii) and parameter estimates for overall environmental context in columns (iii) and (iv). Similar to our previous approach across feeder and non-feeder high schools, we are again interested in the estimates on the interaction terms of both Dashboard elements and the treatment indicator.

At both sets of institutions, applicants who outperform their peers' SAT scores have more favorable odds of admission. This is revealed through the statistically significant parameter estimates on SAT contextual differences in Table 6, with log-odds ratio estimates of 1.016 (column (i)) and 1.010 (column (ii)), and is true even after accounting for the applicant's own SAT scores. We cannot, however, conclude that the slope of the relationship between scores in context and admissions likelihood differs between the treatment and control groups as the log-odds ratios on the interaction of contextual SAT scores and the treatment indicator fails to reach statistical significance for both the less holistic and more holistic colleges

When we replace SAT contextual difference with overall environmental context in columns (iii) and (iv) of Table 6, we find that the log-odds ratio of 0.996 in column (iii) on the interaction term of overall environmental context and treatment is not statistically different from 1.000. This means that, at universities using less holistic approaches to admission, the Dashboard-revealed data on environmental context did not shift admissions recommendations. By contrast, the log-odds ratio of 1.005 on the interaction term in column (iv) for universities using

Table 6. Experimental admission recommendation across less and more holistic admissions approaches.

Reporting log-odds ratios

	(i) Less Holistic	(ii) More Holistic	(iii) Less Holistic	(iv) More Holistic
Treatment	0.837**	0.935	0.933	0.924
	[0.061]	[0.142]	[0.107]	[0.079]
Overall Environmental Context			1.009***	1.011***
			[0.003]	[0.003]
Overall Environmental Context X Treatment			0.996	1.005*
			[0.003]	[0.003]
Average Overall Environmental Context			27.042	23.884
Contextual SAT Score Difference	1.016**	1.010**		
	[0.007]	[0.005]		
Contextual SAT Score Difference X Treatment	0.996	1.005		
	[0.006]	[0.006]		
Average Contextual SAT Score Difference/10	3.394	21.675		
N	8349	7880	8531	7927
Pseudo-R2	0.364	0.234	0.361	0.241
Average Experimental Admission	0.515	0.323	0.515	0.323

****p*-value < .01, ***p*-value < .05, **p*-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Based on characteristics in Table 1, Universities 3, 5, and 6 are the More Holistic Universities. Universities 1, 2, 4, and 7 are the Less Holistic Universities.

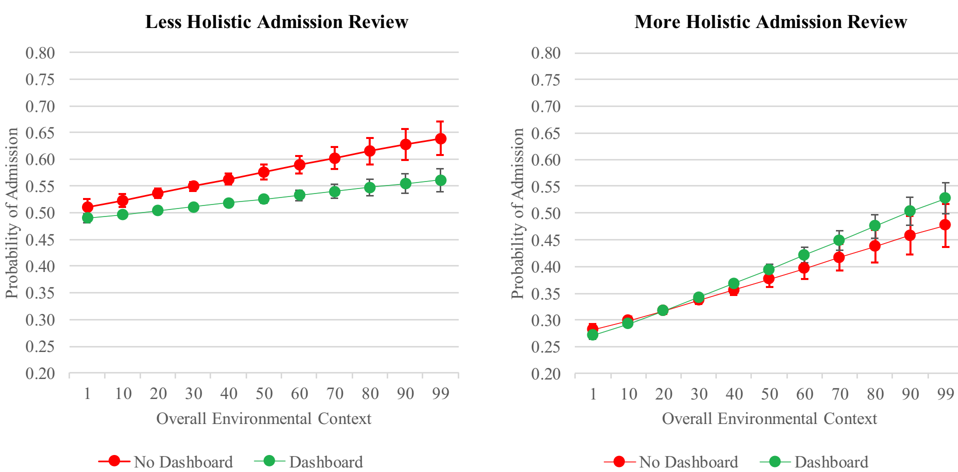


Figure 4. Predicted probabilities of admission across contextual information, by type of admission review.

Notes: These figures use Stata's margins command to show students' predicted outcomes across contextual information. 95% confidence intervals are shown as error bars around the average predicted outcome.

more holistic approaches reaches statistical significance at the 0.10 level, indicating that the Dashboard-revealed environmental context encouraged more favorable admissions decisions for students from more challenging environments. We demonstrate this impact visually in Figure 4. At less holistic institutions, the relationships between overall environmental disadvantage and the probability of admission are similar between the Dashboard and No Dashboard groups. At colleges with more holistic practices, the relationship between admissions probability and environmental context is steeper in the Dashboard group than in the No Dashboard group.

Limitations

Our results provide some indication of how more robust environmental contextual data and SAT scores in context may benefit students from more challenging contexts or students who outperform their high school peers on the SAT. However, there are several empirical limitations preventing us from making definitive statements on how these data might change admissions at a broad range of universities. The eight universities that participated are not representative of all higher education institutions — all are at least moderately selective and more resourced than the average university in the United States. However, it is likely that the results of our paper can be extrapolated to the broader set of colleges that practice holistic admissions. Moreover, our results are most relevant for applicants near the cusp of admission and do not speak to the role of context for the most and least admissible applicants.

As admissions officers were reading applications from previous admissions cycles, they did not have to worry about the potential effects of admitting more low-income students, particularly considering the revenue effects of replacing students from higher-income families with those from more challenging environments who would need additional financial assistance. Admissions officers may also believe that low-income applicants are less likely to persist and graduate, and might therefore might be less likely to admit these applicants in real-world scenarios. Future research will examine how the composition of admitted and enrolled students changes as a direct result of this tool's introduction in the official admissions process, where decisions obviously have real consequences for the student and the institution.

Our experimental design prevents us from making causal claims about how individual pieces of contextual information in isolation influence admission decisions. We randomized applications to be read with the full contextual information provided by the Dashboard in its entirety, or no contextual information at all. We also did not manipulate applicants' contextual information itself. Our analysis focused on the two most valued components of the Dashboard among readers — overall environmental context and contextual SAT performance. In the future, it will be important to more clearly identify which aspects of the applicant's context are the most relevant, useful, and actionable as additional admissions considerations, particularly in the increasingly dynamic landscape of highly selective college admissions.

Discussion and conclusion

Our results suggest that the Environmental Context Dashboard, and contextual information more broadly, can provide a meaningful benefit to both applicants and admissions staff. We find that the Dashboard shifted admissions decisions above and beyond any priming effects, particularly for applicants from high schools with which admissions readers have less familiarity. We also find that admissions officers at colleges with more holistic admissions processes were more likely to admit students who outperformed their high school peers when provided with the Dashboard.

There is evidence that providing a brief orientation and background document about the Dashboard induced admissions readers to consider student contextual information in admissions decisions, regardless of whether or not this information was revealed through the Dashboard. This suggests that admissions staff were primed to infer information about an applicant's background from traditional application materials, and that emphasizing such information when reviewing applications may also increase the likelihood of admission for applicants who substantially outperformed high school peers on the SAT or who came from more challenging environments.

We add to emerging research suggesting that contextualized data can be a significant boon for equity in college admissions practices. Data from admissions reform in Colorado suggests that including environmental context data in admissions review led to a substantial increase in enrollment by both low-income students and students of color (Gaertner & Hart, 2013). Similarly, an experimental simulation with admissions officers from selective colleges from across the country found that admissions officers were 26–28% more likely to admit a low-SES applicant when provided with more robust contextual information (Bastedo & Bowman, 2017). This research responds to the limitations of previous studies and provides additional causal evidence that data quality and presentation are important considerations in designing equitable admissions practices.

There are important additional considerations. This research provides evidence that changes in data quality or presentation are inadequate without a commitment to holistic admissions practices. Though 95% of admissions officers report using holistic admissions, only 29% of admissions officers read applications in light of the opportunities available in the family, neighborhood, and high school (Bastedo et al., 2018). Thus, a Dashboard that provides high quality contextual data is only likely to be influential among those admissions officers who espouse a holistic approach, which was confirmed in this study.

Training and norming practices in admissions offices are another crucial consideration. In this experiment, admissions officers were oriented to the data elements in the Dashboard and how they could be useful, but they were encouraged to treat applications as they would in an official review process. Although no college appeared to employ an unwavering formula-based approach to admissions, some reported on a later survey that they could not accommodate the data in their more formula-driven process. Training and norming are crucial in an admissions office to ensure that there is high reliability in admissions decisions across admissions officers (Rideout, 2018). We need to know more about how admissions decisions may change once admissions officers were trained by senior leadership on how to use the Dashboard in their particular offices. This will also tell us more about how the data elements are interpreted by admissions officers, and which elements are perceived to be more useful than others.

This research has shaped the development of a new iteration of the Dashboard, Landscape, which builds upon the earlier prototype used in this research.[21] Responding to public feedback, Landscape reduces negative-frame language, like "adversity" and "disadvantage," and is now easily integrated into admissions office data systems. The in-person and online training provided to admissions officers on Landscape now include an overview of research on decision-making bias and the role of context in holistic review; a review of data sources, definitions, and methodology underlying the tool; best practice guides to support varied use cases; and presentations for admissions leaders to use in training their staff and informing their campus community.

There are many remaining questions related to this research. In particular, how does the Dashboard allow for a more streamlined and accurate presentation of data, especially compared with more ad-hoc methods of gathering contextual information? Despite focusing on admissions decisions and ratings, an important ancillary question is whether providing the Dashboard simplified the evaluation process for readers by assembling information that they might have gleaned from documents in their application files. It is also unknown whether differential effects may be found in universities that use other decision-making strategies, such as committee-based evaluation (Romero Da Silva, 2017). Finally, we do not know the degree to which these contextualized elements are connected to student success, such as grades, retention, and graduation. All of these are important considerations for future work in this area.

Admissions decisions are crucial to shaping the composition of the enrolling class, but there are many other considerations in the full enrollment management process. Many of the key practices and policies in enrollment management — such as early decision, merit scholarships, rankings, and many others — actively disadvantage low-income students (Bastedo, 2016). In this study, we have determined through our priming analyses, that, in many instances, admissions professionals have the ability to identify students from more challenging environments even without the provision of the Dashboard. We also show that the key contextual data have the potential to add value by either confirming or reshaping the priors of admissions professionals.

As we have already noted, the results presented in the paper reflect real-world applications, not real-world decisions. There are clear financial implications that come along with admitting students, and the study participants were not forced to contend with the financial consequences of their decisions. For all but a tiny sliver of postsecondary institutions, revenue considerations become highly important at the end of the decision process, and more equitable admissions practices need to be matched by state, federal, and institutional funding policies that facilitate the enrollment of low-income students. Similarly, admissions professionals in an experimental context might be more willing to recommend students with lower SAT scores who happened to outperform their peers by wide margins than in a high-stakes situation, where consideration about remediation or college completion might emerge.

Finally, we feel compelled to acknowledge how this work intersects with the potentially long-lasting changes to the college admissions process introduced by the COVID-19 pandemic. We have already seen dramatic shifts in both application behavior, with surges in applications to the most-selective colleges, and sharp declines in enrollment at less-selective colleges and community colleges. Colleges have greatly expanded test-optional reviewing practices, but it is unknown to what extent these will continue. The full impact of COVID-19 on the college admissions landscape remains to be seen, but widespread instructional disruption and shifts in college admissions policies will likely influence the application portfolios of students, including the presentation of high school grades, college entrance examination scores and extracurriculars. These shifts in admissions practices and application behavior may amplify the importance of the data provided by contextual dashboards, as admissions professionals contend with making fair and equitable choices in a more challenging decision-making environment.

Notes

1. In the pilot phase of this work, the Dashboard was called the Environmental Context Dashboard (ECD) by the College Board. After further development and expanded use in live admissions cycles by more than 100 colleges and universities, the ECD became Landscape in August 2019. The description of the ECD in this paper follows the pilot dashboard in use during our experimental study. We provide a brief overview of changes between ECD and Landscape in the discussion at the end of the paper.
2. The phrase "Overall Adversity Index" was changed after controversy and criticism in the national media over the purported use of "adversity scores" from across the ideological spectrum (e.g., Patel, 2019; Will, 2019). Although there was a great deal of misunderstanding of the adversity index — raw SAT scores were never adjusted or modified, for example — the College Board acknowledged that the adversity language was problematic and removed it from the Dashboard in its next iteration (Hartocollis, 2019).

3. The admissions files used were for students applying to be part of the Fall 2016 or Fall 2017 freshmen cohort.
4. For the sake of consistency, we considered the first reader's admissions recommendation. This is our admissions outcome of interest, and the one used to create the experimental group.
5. We exclude the most recent cohort of applications from which we draw files for the experimental review.
6. For five universities, this was about three applications from a given high school each year. For others it was 1.3, 6, and 7.
7. We concorded SAT and ACT scores to the 1600-scale. The concordance tables are available at https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdf.
8. We include SAT and PSAT takers among high school graduates, which includes non-college-going students.
9. Since the Dashboard pilot, we have engaged with several other institutions that agreed to be surveyed. Readers were asked to rank different components of the Dashboard. Overall environmental context was ranked first for 43% of all respondents, contextual SAT scores was ranked first for 21%. The third most commonly ranked first component was neighborhood environmental context, with 12%. The twelve universities we surveyed are similar in composition to the universities who participated. In fact, Universities 4 and 6 are among the twelve.
10. One institution did not provide official reader information.
11. We provided historical reports to participating universities that analyzed environmental context across previous admissions cycles. However, this information was not explicitly provided as part of the Dashboard orientation and was not disseminated to the individual admission readers participating in the experiment.
12. Averages are calculated within institutions, using university fixed effects.
13. The original sample consisted of 21,450 students across all 8 colleges. The final analytic sample contained 18,246 applications, leading to a non-response rate of 15%. In the covariate balance table, we show that the analytic sample is perfectly balanced suggesting that the modest attrition did not result in meaningful differences between the treatment and control groups.
14. We control for whether students took the 2400-scale SAT, took the ACT, their SAT (or equated ACT) score, whether they are missing an SAT or ACT score, gender, and reported ethnicity (white, black, Hispanic, Asian, Native American/Pacific Islander/Hawaiian, or missing).
15. For the purpose of this analysis, SAT scores are on a 40–160 scale. This allows for easier interpretation of the parameter estimates, since SAT scores can only change in 10-point increments.
16. We offer some caution regarding the interpretation of parameter estimates on the interaction term between overall environmental context and treatment in Appendix A Tables A2 and A4. All applications in this study were reviewed in a high-stakes environment. However, due to annual turnover in admissions offices, some of the study participants may not have been employed at their respective universities when the study's applications were reviewed in a high-stakes context. Similarly, some of the staff who read the study's applications in a high-stakes context may have departed from their respective universities or were otherwise unable to participate in the study. Preserving only the applicants evaluated by readers in the official and experimental reads leads to imbalances on key covariates between the treatment and control groups.

17. We test the sensitivity of our priming conclusions to sub-samples containing only applications reviewed by evaluators participating in both the experimental and official evaluations. Here, the evidence of priming is suggestive in nature. That is, the t-statistics of differences between columns (i) and (ii) do not reach statistical significance (Appendix Table A5).
18. Just as we did for Table 3, we estimate the regressions in Table 4 on the subset of students read by reviewers participating in both the official and experimental reviews. These results are in Appendix Table A5.
19. The SAT standard deviation appears in this report: https://reports.collegeboard.org/pdf/2020-total-group-sat-suite-assessments-annual-report.pdf
20. Universities 3, 5, and 6 from Table 1 are classified as being less holistic, and Universities 1, 2, 4, and 7 are classified as being more holistic.
21. A more detailed description of Landscape is available at https://pages.collegeboard.org/landscape.

Disclosure statement

No potential conflict of interest was reported by the author(s). Julian Hsu conducted this research prior to joining Amazon.com.

ORCID

Michael N. Bastedo http://orcid.org/0000-0003-3358-2564

References

Anderson, N. (2021, January 29). Applications surge after big-name colleges halt SAT and ACT testing rules. *The Washington Post*. Retrieved February 5, 2021, from https://www.washingtonpost.com/local/education/harvard-uva-sat-act-requirement-college-applications/2021/01/29/90566562-6176-11eb-9430-e7c77b5b0297_story.html

Avery, C., & Levin, J. (2010). Early admissions at selective colleges. *American Economic Review*, *100*(5), 2125–2156. https://doi.org/10.1257/aer.100.5.2125

Bastedo, M. N. (2016). Enrollment management and the low-income student: How holistic admissions and market competition can impede equity. In A. P. Kelly, J. S. Howell, & C. Sattin-Bajaj (Eds.), *Matching students to opportunity: Expanding college choice, access, and quality* (pp. 121–134). Harvard Education Press.

Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher*, *46*(2), 67–77. https://doi.org/10.3102/0013189X17699373

Bastedo, M. N., Bowman, N. A., Glasener, K. M., & Kelly, J. L. (2018). What are we talking about when we talk about holistic review? Selective college admissions and its effects on low-SES students. *The Journal of Higher Education*, *89*(5), 782–805. https://doi.org/10.1080/00221546.2018.1442633

Bastedo, M. N., & Flaster, A. (2014). Conceptual and methodological problems in research on college undermatch. *Educational Researcher*, *43*(2), 93–99.

Bastedo, M. N., Howard, J. E., & Flaster, A. (2016). Holistic admissions after affirmative action: Does 'maximizing' the high school curriculum matter? *Educational Evaluation and Policy Analysis*, *38*(2), 389–409. https://doi.org/10.3102/0162373716642407

Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low-income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis*, *33*(3), 318–339. https://doi.org/10.3102/0162373711406718

Black, S. E., Cortes, K. E., & Lincove, J. A. (2015). Academic undermatching of high-achieving minority students: Evidence from race-neutral and holistic admissions policies. *American Economic Review*, *105*(5), 604–610. https://doi.org/10.1257/aer.p20151114

Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions*. Princeton University Press.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton University Press.

Bowman, N. A., & Bastedo, M. N. (2018). What role may admissions office diversity and practices play in equitable decisions? *Research in Higher Education*, *59*(4), 430–447. https://doi.org/10.1007/s11162-017-9468-9

Card, D. (2017, December 15). Report of David Card, Ph.D. Expert report submitted for *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College*, United States District Court for the District of Massachusetts. Retrieved October 2, 2018, from https://projects.iq.harvard.edu/files/diverse-education/files/expert_report_as_filed_d._mass._14-cv-14176_dckt_000419_033_filed_2018-06-15.pdf .

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the United States. *Quarterly Journal of Economics*, *135*(3), 1567–1633. https://doi.org/10.1093/qje/qjaa005

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review*, *106*(4), 855–902. https://doi.org/10.1257/aer.20150572

Clinedinst, M. (2019). *State of college admission 2019*. National Association for College Admission Counseling.

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management*, *28*(4), 555–576. https://doi.org/10.1002/pam.20455

Dale, S. B., & Krueger, A. B. (2011). *Estimating the return to college selectivity over the career using administrative earnings data* (No. w17159). National Bureau of Economic Research.

Dynarski, S., Libassi, C. J., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721–1756. https://doi.org/10.1257/aer.20200451

Engberg, M. E., & Wolniak, G. C. (2010). Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*, *51*(2), 132–153. https://doi.org/10.1007/s11162-009-9150-y

Espenshade, T. J., Chung, C. Y., & Walling, J. L. (2004). Admission preferences for minority students, athletes, and legacies at elite universities. *Social Science Quarterly*, *85*(5), 1422–1446. https://doi.org/10.1111/j.0038-4941.2004.00284.x

Frankenberg, E., Ee, J., Ayscue, J. B., & Orfield, G. (2019). *Harming our common future: America's segregated schools 65 years after Brown*. UCLA Civil Rights Project.

Furquim, F., & Glasener, K. M. (2017). A quest for equity? Measuring the effect of QuestBridge on economic diversity at selective institutions. *Research in Higher Education*, *58*(6), 646–671. https://doi.org/10.1007/s11162-016-9443-x

Gaertner, M. N., & Hart, M. (2013). Considering class: College access and diversity. *Harvard Law & Policy Review*, *7*, 367–403.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*(1), 21. https://doi.org/10.1037/0033-2909.117.1.21

Goodman, J., Hurwitz, M., & Smith, J. (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics*, *35*(3), 829–867. https://doi.org/10.1086/690818

Gurantz, O., Howell, J., Hurwitz, M., Larson, C., Pender, M., & White, B. (2021). A national-level informational experiment to promote enrollment in selective colleges. *Journal of Policy Analysis and Management*, *40*(2), 453–479. https://doi.org/10.1002/pam.22262

Gurantz, O., Pender, M., Mabel, Z., Larson, C., & Bettinger, E. (2020). Virtual advising for high-achieving high school students. *Economics of Education Review*, *75*, 1–9. https://doi.org/10.1016/j.econedurev.2020.101974

Hartocollis, A. (2019, August 27). SAT "adversity score" is abandoned in the wake of criticism. *The New York Times* (pp. B5).

Hill, C. B., & Winston, G. C. (2010). Low-income students and highly selective private colleges: Geography, searching, and recruiting. *Economics of Education Review*, *29*(4), 495–503. https://doi.org/10.1016/j.econedurev.2009.12.004

Holland, M. M. (2014). Navigating the road to college: Race and class variation in the college application process. *Sociology Compass*, *8*(10), 1191–1205. https://doi.org/10.1111/soc4.12203

Holzman, B., Klasik, D., & Baker, R. (2019). Gaps in the college application gauntlet. *Research in Higher Education*, *61*(7), 795–822. https://doi.org/10.1007/s11162-019-09566-8

Hoxby, C. M., & Turner, S. (2013). *Expanding college opportunities for high-achieving, low income students* (No. 12–014). Stanford Institute for Economic Policy Research.

Hurwitz, M. (2011). The impact of legacy status on undergraduate admissions at elite colleges and universities. *Economics of Education Review*, *30*(3), 480–492. https://doi.org/10.1016/j.econedurev.2010.12.002

Hurwitz, M., & Smith, J. (2018). Student responsiveness to earnings data in the College Scorecard. *Economic Inquiry*, *56*(2), 1220–1243. https://doi.org/10.1111/ecin.12530

Institute of Education Sciences (IES). (2020). *What works clearinghouse standards handbook 4.1*. IES.

Jack, A. A. (2019). *The privileged poor: How elite colleges are failing disadvantaged students*. Harvard University Press.

Jaschik, S. (2021, January 26). *Alarm bells on first-generation, low-income applicants*. Inside Higher Ed. Retrieved February 5, 2021, from https://www.insidehighered.com/admissions/article/2021/01/26/common-apps-new-data-show-overall-gains-applications-not-first .

Jaschik, S., & Lederman, D. (2020). *2020 survey of college and university admissions directors: A mess of a year*. Inside Higher Ed. Retrieved April 5, 2021, from https://www.insidehighered.com/news/survey/2020-survey-admissions-leaders .

Jeong, D. W. (2009). Student participation and performance on Advanced Placement exams: Do state-sponsored incentives make a difference? *Educational Evaluation and Policy Analysis*, *31*(4), 346–366. https://doi.org/10.3102/0162373709342466

Klugman, J. (2013). The Advanced Placement arms race and the reproduction of educational inequality. *Teachers College Record*, *115*(5), 1–34.

Kolluri, S. (2018). Advanced placement: The dual challenge of equal access and effectiveness. *Review of Educational Research*, *88*(5), 671–711. https://doi.org/10.3102/0034654318787268

Leonhardt, D. (2017). The assault on public colleges – And the American dream. *The New York Times*. Retrieved September 25, 2018, from https://www.nytimes.com/2017/05/25/opinion/sunday/the-assault-on-colleges-and-the-american-dream.html .

Lock Morgan, K., & Rubin, D. (2012). Rerandomization to improve covariate balance in experiments. *The Annals of Statistics*, *40*(2), 1263–1282.

Lucido, J. A. (2015). How admissions decisions get made. In D. Hossler & B. Bontrager (Eds.), *Handbook of strategic enrollment management* (pp. 147–170). Jossey-Bass.

McDonough, P. M. (1997). *Choosing colleges: How social class and schools structure opportunity*. State University of New York Press.
Oreopoulos, P., & Petronijevic, U. (2019). *The remarkable unresponsiveness of college students to nudging and what we can learn from it* (No. w26059). National Bureau of Economic Research.
Owens, A. (2018). Income segregation between school districts and inequality in students' achievement. *Sociology of Education*, *91*(1), 1–27. https://doi.org/10.1177/0038040717741180
Patel, L. (2019, May 22). *The SAT's new adversity score is a poor fix for a problematic test*. The Conversation. Retrieved May 26, 2020, from https://theconversation.com/the-sats-new-adversity-score-is-a-poor-fix-for-a-problematic-test-117363
Perna, L. W., May, H., Yee, A., Ransom, T., Rodriguez, A., & Fester, R. (2015). Unequal access to rigorous high school curricula: An exploration of the opportunity to benefit from the International Baccalaureate Diploma Programme (IBDP). *Educational Policy*, *29*(2), 402–425. https://doi.org/10.1177/0895904813492383
Posselt, J. R., Jaquette, O., Bielby, R., & Bastedo, M. N. (2012). Access without equity: Longitudinal analyses of institutional stratification by race and ethnicity, 1972–2004. *American Educational Research Journal*, *49*(6), 1074–1111. https://doi.org/10.3102/0002831212439456
Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. J. Murnane & G. J. Duncan (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). Russell Sage Foundation.
Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092–1153. https://doi.org/10.1086/657114
Rideout, B. M. (2018). Variance and inter-rater reliability in holistic admissions review. *College and University*, *93*(3), 2–13.
Rivera, L. (2011). Ivies, extracurriculars, and exclusion: Elite employers' use of educational credentials. *Research in Social Stratification and Mobility*, *29*(1), 71–90. https://doi.org/10.1016/j.rssm.2010.12.001
Roderick, M., Coca, V., & Nagaoka, J. (2011). Potholes on the road to college: High school effects in shaping urban students' participation in college application, four-year college enrollment, and college match. *Sociology of Education*, *84*(3), 178–211. https://doi.org/10.1177/0038040711411280
Rodriguez, A., & McGuire, K. M. (2019). More classes, more access? Understanding the effects of course offerings on Black-White gaps in Advanced Placement course-taking. *The Review of Higher Education*, *42*(2), 641–679. https://doi.org/10.1353/rhe.2019.0010
Romero Da Silva, Y. M. (2017). *Change in the admissions evaluation process: A study of the adoption of committee-based evaluation at selective colleges and universities* [Doctoral dissertation]. University of Pennsylvania.
Saenz, V. B., Oseguera, L., & Hurtado, S. (2007). Losing ground? Exploring racial/ethnic shifts in enrollment at selective institutions. In G. Orfield, P. Marin, S. M. Flores, & L. M. Garces (Eds.), *Charting the future of college affirmative action: Legal victories, continuing attacks, and new research* (pp. 173–204). The Civil Rights Project at UCLA.
Salazar, K. G., Jaquette, O., & Han, C. (2021, in press). Coming soon to a neighborhood near you? Off-campus recruiting by public research universities. *American Educational Research Journal*. https://doi.org/10.3102/00028312211001810
Slay, K. (2017). *Choosing colleges in a post-affirmative action era: Black students' perceptions of institutional diversity and campus climate* [Doctoral dissertation]. University of Michigan.

Smith, J., Goodman, J., & Hurwitz, J. (2020). *The economic impact of access to public four-year colleges* (NBER Working Paper No. 27177). National Bureau of Economic Research.
Smith, J., Hurwitz, M., & Howell, J. (2015). Screening mechanisms and student responses in the college market. *Economics of Education Review*, *44*, 17–28. https://doi.org/10.1016/j.econedurev.2014.10.005
Smith, J., Pender, M., & Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review*, *32*, 247–261. https://doi.org/10.1016/j.econedurev.2012.11.001
Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2013). Choice architecture. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 428–439). Princeton University Press.
Weininger, E. B., Lareau, A., & Conley, D. (2015). What money doesn't buy: Class resources and children's participation in organized extracurricular activities. *Social Forces*, *94*(2), 479–503. https://doi.org/10.1093/sf/sov071
Weis, L., Cipollone, K., & Jenkins, H. (2014). *Class warfare*. University of Chicago Press.
Will, G. (2019, June 7). The SAT's new 'adversity index' is another step down the path of identity politics. *The Washington Post*.
Wolniak, G. C., & Engberg, M. E. (2007). The effects of high school feeder networks on college enrollment. *The Review of Higher Education*, *31*(1), 27–53. https://doi.org/10.1353/rhe.2007.0054

Appendix A: Environmental Context Pilot Pre-Read and Orientation

As part of our continuing commitment to providing resources, data, and new and innovative tools to the higher education community, The College Board has been working with admissions practitioners from a wide range of colleges to better understand the holistic review process. The goal of this ongoing project is to explore and document recent and emerging needs and trends in how colleges select students and to partner with the admissions community on the development of new sources of information and practical tools designed to meet these new and emerging challenges. An important part of this work is to better understand how practitioners combine the many distinct sources of quantitative and qualitative information available to them as they make admission decisions.

One topic that has emerged as being potentially important is the need for additional contextual information about students' environments, particularly for those students who come from areas or attend high schools where the admission officer does not have direct personal experience. Working with colleges over the past year we have developed a prototype "Environmental Context Dashboard" that attempts to capture key elements of an applicant's environment that might suggest adverse influences or other obstacles not otherwise apparent from the student's application. In particular, we have assembled data in this Dashboard that attempts to quantify three areas of influence:

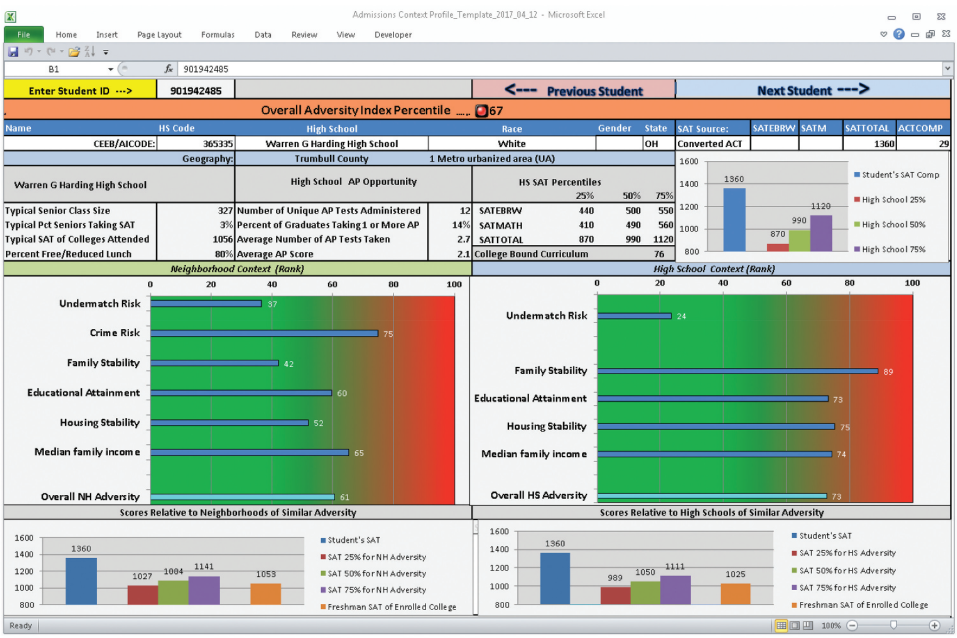
(1) **High School Environment** — Measures related to access to AP courses, the socioeconomic distribution of the high school's student body, relative academic performance, and "undermatching."
(2) **Social Environment —** Measures related to family income, the proportion of single parent families, the average educational level, and the percentage of non-native speakers.

(3) **Neighborhood Environment —** Measures related to the socio-cultural milieu the applicant is exposed to as they move between school and home as determined by location, including housing values, vacancy rates, poverty measures, and crime risk.

It is important to recognize that the data we have incorporated into the Dashboard, while systematically and consistently measured, do not necessarily represent the student's personal experience, but rather, suggest the environment to which they were likely exposed. As such, it does not substitute for first-hand knowledge of the applicant or specific information that is conveyed by the applicant's written narrative. It does however provide an additional lens through which to view the student's application and might help to highlight or further explain the detail found in the application — particularly for those high schools or neighborhoods that are less familiar to the file reader.

The Environmental Context Dashboard

The Environmental Context Dashboard organizes the metrics that we have assembled, derived, or estimated, and integrates these data with basic information about the student (SAT scores, location, and high school). The student's location and high school serve as the basis for the contextual information on the student's environment. The resulting Excel-based Dashboard for a particular applicant is depicted below.



We have pre-populated the dashboard with your institution's 2017 applicants, along with the related contextual information for each applicant. The information relevant to a particular application can be displayed by entering the applicant's ID in second column of the first row, and then pressing enter. The second two rows of the dashboard contain **all of the student specific information**; the remainder of the dashboard provides the context on the applicant's high school or neighborhood.

Applicant Information

The fourth row provides the labels for the data contained in the third row. Listed in order are:

- HS Code — The CEEB code for the applicant's high school as provided by your Admission Office
- High School — name of the applicant's High School based on the CEEB code
- State — applicant's home state based on the address provided by your Admission Office
- Gender — applicant's gender as provided by your Admission Office
- Race — applicant's race if provided by your Admission Office
- SAT Scores — The applicant's SATEBRW, SATM, SATOTAL Scores. This source data was provided by your Admission Office, as either legacy SAT scores, new SAT scores, or as ACT scores. Regardless of the source the scores were converted to the new SAT scale for comparison purposes. However, the original source is noted.

Location

Immediately under the high school name we indicate general the location of the student including the county and indicator of whether the student lives in a population center, a suburb or rural environment.

High School Information

The five rows underneath the applicant attributes contain contextual information that is specific to that applicant's particular High School. Block headings are colored with a gray background and include:

- High School Name (all data based on a three year average)
  - The average senior class size
  - Average percent of seniors taking the SAT
  - Average freshman SAT at colleges attended by SAT-taking graduates of the applicant's high school
  - Percent of students with Free and Reduced Lunch*
- High School AP Opportunity (three year average)
  - Number of unique AP Exams taken by students from applicant's high school
  - Percent of senior class who took at least one AP Exam
  - Average number of AP Exams taken by graduates with at least one AP
  - Average AP scores across all AP-takers and Exams
- High School Percentiles (three year average) — This block lists the 25th, 50th, and 75th SATV, SATM, SATC (V + M) score percentiles for recent graduates from applicant's high school
- Chart — applicant's SAT (M + V) and the 25th, 50th and 75th SAT score percentiles at the applicant's high school
- An indicator of the relative strength of the curriculum taken by college going students from that high school (1–100)

**Please note that individual data elements in this block may be empty or show missing values. In such cases the missing data was not available for that specific high school. In particular, the Free and Reduced Lunch information is only available for public high schools that report that information to the US Department of Education.*

Neighborhood & High School Bar Graphs

Below the high school data are two horizontal bar graphs that contain derived contextual metrics for the applicant's neighborhood (left) and high school (right). During the development stage of the Environmental Context Dashboard, we engaged with the literature on the importance of neighborhoods (see Chetty et al., 2016) as well as the lived experiences of admissions professionals to identify the data elements that shape the overall challenge measures in Environmental Context Dashboard. These include enumerated below. The neighborhood context presents data aggregated from population-based sources and historical participants in College Board programs such as SAT, PSAT and AP. The data are aggregated across previous students from each neighborhood. The neighborhoods were adapted from College Board Segment Analysis Service and represent small (total population of 4–5 thousand) physically-contiguous geographical areas similar to census tracts. The High School Context is similarly based on historical participants in College Board programs such as SAT, PSAT and AP, with the data being aggregated for past students at that particular high school. The horizontal bars illustrate the **percentile rank** for each attribute based on the national population, with 50 being the national average and higher scores indicating more "adverse" environments. The percentiles for each applicant's Neighborhood and High School Context is shown for the following dimensions or areas:

(1) **Undermatch Risk —** Academic undermatch occurs when a student's academic credentials substantially exceed the credentials of their peers enrolled in the same postsecondary institution. For each neighborhood and high school, we aggregate the difference between the historical SAT scores of individual students from that neighborhood or high school and the average freshman SAT scores of the colleges those students attend. This average difference indicates the degree to which the typical student from a given high school or neighborhood is at risk for undermatching in the college enrollment process, which research has demonstrated to negatively impact a range of educational and occupational outcomes.
(2) **Crime Risk (Neighborhood only) —** The Crime Risk is a geodemographic measure that represents the likelihood of being a victim of a crime — *not* the likelihood of committing a crime. The Crime Risk measure is derived from data that includes the FBI Uniform Crime Reports (UCR) and other risk related data,
(3) **Family Stability** — Family stability is a combined measure based on the proportion of intact families, single-parent families, and children living under the poverty line within each neighborhood, or across the neighborhoods of past students attending that high school. It is primarily based on U.S. Census derived population data.
(4) **Educational Attainment** — Educational attainment is a combined measure that looks at the pattern of educational attainment demonstrated by young adults in the community. It is based largely on population statistics and reflects the overall educational level of recent high school graduates in the student's environment.
(5) **Housing Stability** — Housing stability is a composite measure that includes vacancy rates, rental vs. home ownership, and mobility/housing turnover, again based on aggregate population statistics.
(6) **Median Family Income —** Median family income is based on weighted data from the Census/American Community Survey, and reflects the general SES of the environment.
(7) **Overall Context —** Overall context is a weighted average of the individual metrics listed above.

The data used to develop the environmental measures for the Dashboard is independent of race or ethnicity and the indicators can be considered to be race neutral

Below the horizontal bar graphs are vertical bar graphs for the neighborhood and high school that depict the applicant's SAT score relative to others who share the applicant's overall percentile of neighborhood and high school adversity as well as the Average Freshmen SAT of entering students at the colleges that these respective groups of students attended.

Overall Adversity Index — The overall environmental context measure is indicated in the second row (and is the average of the High School and Neighborhood levels). This index ranges from 0–100, with higher scores being relatively more adverse, and is likewise color coded from green to red.

Environmental Dashboard Pilot

As part of this pilot study you will be asked to read a set of fall 2017 (or fall 2016) applications that were submitted to your university. These folders have been selected to represent a range of geographies and applicant characteristics, with a special focus on students are in the middle of the pool, and less of emphasis on applicants near the top or bottom of your 2017 applicant pool. Additionally, your list was personalized to ensure that it only contains students whose applications you did not read during the actual cycle.

We are asking that you read these folders and evaluate each applicant for admission using the same holistic criteria and standards you employed during this year's recently completed review cycle. Since your list primarily includes applicants "in the middle" the expected admit rate may not match the overall admit rate for the whole applicant pool. Therefore, we will also provide you the actual 2017 admit rate for the specific pool students you will be reviewing, to help you calibrate your decision process. In general, you should strive to recommend admissions for approximately the same percentage of your pool that were actually recommended for admission at your institution this year, but of course, who individually to recommend for admissions, or denial, is based on your overall read and professional judgment.

The Environmental Context Dashboard is intended to compliment the normal reading process by providing additional context about applicants' educational and neighborhood environments. It is not intended to override or substitute for known characteristics of the applicant. Rather, it might be thought of as a lens through which to view the application, and in this way might highlight certain characteristics, aid in interpreting the student's qualifications in light of their opportunities, and generally help the admission reader better understand the applicant's unique path to preparing for college.

In order to aid in this interpretive function, it is recommended that the information provided by the dashboard be reviewed prior to the reading the application. You might review back to specific components of the dashboard as they related to elements of the application that seem incongruent or inconsistent, but the main focus should be on the broad picture presented by the Dashboard, and you should think of it as comprehensive profile with the overall pattern being more important than any of the individual numbers.

In order to help us calibrate the new ratings we ask that you review and evaluate a group of files, as you did previously this cycle, without referring to the Environmental Dashboard. Finally, we will ask you to complete two brief online questionnaires, one before your start reading folders and one after you complete your reads, in order to learn more about your experience and to provide any feedback on the Dashboard data or design.

Finally, in order to retrieve the data for a specific student please type the students ID in the second column of the first row (Enter Student ID ---> 901942485). This number must be typed — it cannot be pasted. Also — since we do not provide the name of applicant. It is important that you verify the correct ID and whether the high school, gender, race and State matches the application data before proceeding.

Table A1. Separately using neighborhood and high school environmental context to estimate experimental admission decisions (log-odds ratios).

		Across Students' High School Feeder Status		Across Type of Admission Review	
	Entire Sample	Feeder High Schools	Non-Feeder High Schools	Less Holistic	More Holistic
High School Environmental Context	1.008***	1.009***	1.004	1.002	1.002
	[0.002]	[0.003]	[0.003]	[0.004]	[0.003]
Treatment	0.942	1.017	0.881	0.919	0.892
	[0.083]	[0.166]	[0.096]	[0.134]	[0.095]
High School Environmental Context X Treatment	1.001	1.001	1.003	0.999	1.005*
	[0.002]	[0.004]	[0.003]	[0.004]	[0.003]
N	16413	5161	10,945	8506	7907
Psuedo-R2	0.458	0.425	0.485	0.493	0.325
Mean High School Environmental Context	27.046	34.480	23.352	28.067	25.948
Average Admission	0.422	0.433	0.422	0.515	0.323
Neighborhood Environmental Context	1.007***	1.007**	1.005*	0.999	1.002
	[0.002]	[0.003]	[0.003]	[0.003]	[0.002]
Treatment	0.972	1.044	0.914	0.917	0.936
	[0.076]	[0.154]	[0.088]	[0.122]	[0.088]
Neighborhood Environmental Context X Treatment	1.000	1.000	1.002	1.000	1.004
	[0.002]	[0.004]	[0.003]	[0.003]	[0.003]
N	16203	5081	10,811	8316	7887
Psuedo-R2	0.454	0.419	0.482	0.488	0.324
Mean Neighborhood Environmental Context	24.311	31.441	20.874	26.520	21.981
Average Admission	0.422	0.433	0.422	0.517	0.323

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

For each high school, we calculated the average number of applications sent to each college using historical application data provided by colleges. Feeder high schools exceed this median number.

Based on characteristics in Table 1, Universities 3, 5, and 6 are the More Holistic Universities. Universities 1, 2, 4, and 7 are the Less Holistic Universities.

Table A2. Admissions decisions regressions for students evaluated by readers participating in both the official and experimental reads (log-odds ratios).

	(i)	(ii)	(iii)	(iv)
Treatment	1.001	1.006**	1.006**	1.009***
	[0.002]	[0.002]	[0.002]	[0.002]
Overall Environmental Context				1.085
				[0.090]
Overall Environmental Context X Treatment				0.996**
				[0.002]
N	6205	6075	6075	8462
Psuedo-R2	0.191	0.215	0.215	0.202
Average Overall Environmental Context	28.560	26.703	28.638	28.086
Average Admission	0.489	0.374	0.372	0.370
Treatment	0.994*	1.007**	1.007**	1.004
	[0.004]	[0.003]	[0.003]	[0.004]
Contextual SAT Score Difference				0.908
				[0.065]
Contextual SAT Score Difference X Treatment				1.004
				[0.004]
N	6053	5930	5930	8282
Psuedo-R2	0.191	0.214	0.214	0.199
Average Contextual SAT Score Difference	9.629	17.908	9.698	11.917
Average Admission	0.494	0.375	0.370	0.369

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A3. Are estimates of contextual information on admissions decisions on the official read statistically different from those from the experimental read?

Across All Students		
Comparing Estimates on Overall Environmental Context		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
Chi-Squared Statistic	14.65743	16.32563
P-Value	.00066	.00029
Comparing Estimates on Contextual SAT Score Difference		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
Chi-Squared Statistic	8.45550	1.72388
P-Value	.01459	.42234
For Students Evaluated by Readers Participating in both the Official and Experimental Reads		
Comparing Estimates on Overall Environmental Context		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
Chi-Squared Statistic	24.27969	6.53417
P-Value	.00001	.03812
Comparing Estimates on Contextual SAT Score Difference		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
Chi-Squared Statistic	6.00398	0.44860
P-Value	.04969	.79907

Chi-squared tests come from stacking the data from Column (i) on Columns (ii) or (iii), and testing whether the coefficient estimates on an indicator of whether the data come from the experimental read and an interaction between that indicator and the contextual information are jointly different from zero. These regressions condition for the same student covariates noted in Table 3 and Appendix Table A4.

Table A4. Admissions ratings regressions for students evaluated by readers participating in both the official and experimental reads (OLS).

	(i)	(ii)	(iii)	(iv)
Treatment	0.000	0.000	0.000	0.003***
	[0.001]	[0.001]	[0.001]	[0.001]
Overall Environmental Context				0.060*
				[0.033]
Overall Environmental Context X Treatment				−0.002***
				[0.001]
N	6765	4644	4644	6763
R2	0.131	0.163	0.163	0.174
Average Overall Environmental Context	29.285	27.097	30.289	29.289
Average Admisison Rating	0.000	−0.192	−0.077	−0.113
Treatment	−0.003**	−0.003*	−0.003*	0.000
	[0.002]	[0.002]	[0.002]	[0.002]
Contextual SAT Score Difference				0.023
				[0.050]
Contextual SAT Score Difference X Treatment				−0.002
				[0.002]
N	6722	4614	4614	6720
R2	0.135	0.164	0.164	0.173
Average Contextual SAT Score Difference	16.890	20.529	15.230	16.891
Average Admisison Rating	0.001	−0.192	−0.076	−0.112

***p-value < .01, **p-value < .05, *p-value < .10.
Standard errors are clustered at the experimental reader level.
Five out of the eight universities we partnered with provided admission ratings for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes additional controls of experimental reader fixed effects. We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A5. Are estimates of contextual information on admission ratings on the official read statistically different from those from the experimental read?

Across All Students		
Comparing Estimates on Environmental Context		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
F-Statistic	11.16039	7.33561
P-Value	.00007	.00135
Comparing Estimates on Contextual SAT Score Difference		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
F-Statistic	3.77498	1.17304
P-Value	.02820	.31599
For Students Evaluated by Readers Participating in both the Official and Experimental Reads		
Comparing Estimates on Environmental Context		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
F-Statistic	12.69393	9.42187
P-Value	.00005	.00040
Comparing Estimates on Contextual SAT Score Difference		
	Column (i) vs Column (ii)	Column (i) vs Column (iii)
F-Statistic	3.51703	2.35274
P-Value	.03849	.10724

F-tests come from stacking the data from Column (i) on Columns (ii) or (iii), and testing whether the coefficient estimates on an indicator of whether the data comes from the experimental read and an interaction between that indicator and the contextual information are jointly different from zero. These regressions condition for the same student covariates noted in Table 4.

Table A6. Admissions decisions regressions (log-odds ratios).

	(i)	(ii)	(iii)	(iv)
Treatment	1.051	1.007	1.007	0.96
	[0.062]	[0.079]	[0.079]	[0.079]
Overall Environmental Context	0.997	1.010***	1.010***	1.009***
	[0.002]	[0.002]	[0.002]	[0.003]
Overall Environmental Context X Treatment	0.996**	0.999	0.999	1.000
	[0.002]	[0.002]	[0.002]	[0.003]
N	16486	16,486	16,486	16,458
Psuedo-R2	0.091	0.428	0.428	0.459
Average Overall Environmental Context	25.718	25.718	25.718	25.696
Average Admission	0.422	0.422	0.422	0.423
Treatment	0.827***	0.917	0.917	0.913
	[0.047]	[0.062]	[0.062]	[0.062]
Contextual SAT Score Difference	1.042***	1.011***	1.011***	1.013***
	[0.007]	[0.004]	[0.004]	[0.004]
Contextual SAT Score Difference X Treatment	1.011***	1.005	1.005	1.005
	[0.003]	[0.004]	[0.004]	[0.004]
N	16272	16,272	16,272	16,229
Psuedo-R2	0.150	0.429	0.429	0.460
Average Contextual SAT Score Difference	12.276	12.276	12.276	12.266
Average Admission	0.422	0.422	0.422	0.422
College Fixed Effects	x	x	x	x
Demographics		x	x	x
Official Admission			x	x
Official and Experimental Reader Fixed Effects				x

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A7. Admissions ratings regressions (OLS).

	(i)	(ii)	(iii)	(iv)
Treatment	−0.080	−0.080	−0.080	−0.107
	[0.068]	[0.068]	[0.068]	[0.077]
Overall Environmental Context	−0.008***	−0.006**	−0.006**	−0.006**
	[0.002]	[0.002]	[0.002]	[0.002]
Overall Environmental Context X Treatment	0.005**	0.005**	0.005**	0.005**
	[0.002]	[0.002]	[0.002]	[0.002]
N	3283	3283	3283	3283
R2	0.013	0.061	0.061	0.118
Average Overall Environmental Context	25.734	25.734	25.734	25.734
Average Admision Rating	0.060	0.060	0.060	0.060
Treatment	−0.017	0.012	0.012	0.013
	[0.070]	[0.072]	[0.072]	[0.076]
Contextual SAT Score Difference	0.000	−0.002	−0.002	−0.002
	[0.003]	[0.003]	[0.003]	[0.003]
Contextual SAT Score Difference X Treatment	0.003	0.002	0.002	0.001
	[0.003]	[0.003]	[0.003]	[0.003]
N	3281	3281	3281	3281
R2	0.006	0.059	0.059	0.115
Average Contextual SAT Score Difference	19.997	19.997	19.997	19.997
Average Admision Rating	0.059	0.059	0.059	0.059
College Fixed Effects	x	x	x	x
Demographics		x	x	x
Official Admission			x	x
Official and Experimental Reader Fixed Effects				x

****p*-value < .01, ***p*-value < .05, **p*-value < .10.

Standard errors are clustered at the experimental reader level.

Five out of the eight universities we partnered with provided admission ratings for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes additional controls of experimental reader fixed effects and original admission outcome. We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A8. Experimental admission recommendation for colleges with admission ratings (log-odds ratios).

	(i)	(ii)	(iii)	(iv)
Treatment	1.000	1.002	1.002	0.922
	[0.178]	[0.178]	[0.178]	[0.173]
Overall Environmental Context	0.991*	0.998	0.998	0.998
	[0.005]	[0.005]	[0.005]	[0.005]
Overall Environmental Context X Treatment	1.008	1.008	1.008	1.009
	[0.006]	[0.006]	[0.006]	[0.006]
N	3283	3283	3283	3265
Psuedo-R2	0.304	0.379	0.379	0.448
Average Overall Environmental Context	25.734	25.734	25.734	25.656
Average Admission	0.240	0.240	0.240	0.240
Treatment	1.427*	1.538*	1.538*	1.445
	[0.296]	[0.401]	[0.401]	[0.482]
Contextual SAT Score Difference	1.018***	1.008	1.008	1.004
	[0.006]	[0.007]	[0.007]	[0.010]
Contextual SAT Score Difference X Treatment	0.993	0.989	0.989	0.989
	[0.010]	[0.012]	[0.012]	[0.015]
N	3281	3281	3281	3263
Psuedo-R2	0.306	0.378	0.378	0.447
Average Contextual SAT Score Difference	19.997	19.997	19.997	19.981
Average Admission	0.240	0.240	0.240	0.240
College Fixed Effects	x	x	x	x
Demographics		x	x	x
Official Admission			x	x
Official and Experimental Reader Fixed Effects				x

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level. Some students are dropped in the analyses if there is no variation in outcomes after controlling for covariates.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A9. Evidence of priming using OLS instead of a logistic regression — admission based on environmental context and contextual SAT score difference.

	(i)	(ii)	(iii)	(iv)
	Admitted in Official Read	Admitted in Experiment, Control	Admitted in Experiment, Treatment	Admitted in Experiment, Treatment/ Control
Overall Environmental Context	0.001**	0.002***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]
Treatment				−0.002
				[0.010]
Overall Environmental Context X Treatment				0.000
				[0.000]
N	16486	3696	12,790	16,486
R2	0.429	0.312	0.341	0.527
Overall Environmental Context	25.718	25.339	25.828	25.718
Average Admission	0.422	0.405	0.427	0.422
Contextual SAT Score Difference	0.000	0.002**	0.002***	0.001***
	[0.000]	[0.001]	[0.000]	[0.000]
Treatment				−0.007
				[0.008]
Contextual SAT Score Difference X Treatment				0.000
				[0.000]
N	16272	3653	12,619	16,272
R2	0.430	0.310	0.341	0.530
Contextual SAT Score Difference/10	12.276	15.797	11.257	12.276
Average Admission	0.422	0.406	0.427	0.422

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Table A10. Experimental admission recommendation across feeder high schools vs. non-feeder high schools, OLS results.

	(i)	(ii)	(iii)	(iv)
	Non-Feeder HS	Feeder HS	Non-Feeder HS	Feeder HS
Treatment	−0.037*	−0.009	0.013	−0.002
	[0.022]	[0.010]	[0.025]	[0.013]
Overall Environmental Context			0.002***	0.001***
			[0.001]	[0.000]
Overall Environmental Context X Treatment			0.000	0.000
			[0.001]	[0.000]
Overall Environmental Context Mean			32.869	22.125
Contextual SAT Score Difference	0.000	0.001**		
	[0.001]	[0.001]		
Contextual SAT Score Difference X Treatment	0.002*	0.000		
	[0.001]	[0.001]		
Average Contextual SAT Score Difference/10	17.941	9.544		
N	5304	10,964	5471	11,000
Pseudo-R2	0.346	0.368	0.346	0.368
Average Experimental Admission	0.423	0.422	0.423	0.422

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

For each high school, we calculated the average number of applications sent to each college using historical application data provided by colleges. Feeder high schools exceed this median number.

Table A11. Experimental admission recommendation across less and more holistic admissions approaches, OLS results.

	(i)	(ii)	(iii)	(iv)
	Less Holistic	More Holistic	Less Holistic	More Holistic
Treatment	−0.026**	0.002	−0.005	−0.009
	[0.011]	[0.021]	[0.016]	[0.015]
Overall Environmental Context			0.002***	0.002***
			[0.000]	[0.001]
Overall Environmental Context X Treatment			−0.001**	0.001
			[0.000]	[0.000]
Overall Environmental Context Mean			27.294	24.028
Contextual SAT Score Difference	0.002***	0.002*		
	[0.001]	[0.001]		
Contextual SAT Score Difference X Treatment	0.000	0.000		
	[0.001]	[0.001]		
Average Contextual SAT Score Difference/10	3.394	21.657		
N	8358	7914	8531	7955
Pseudo-R2	0.384	0.260	0.380	0.264
Average Experimental Admission	0.516	0.323	0.515	0.323

***p-value < .01, **p-value < .05, *p-value < .10.

Standard errors are clustered at the experimental reader level.

Seven out of the eight universities we partnered with provided admission recommendations for students. Average Contextual SAT Score Difference is the difference between a student's own SAT score and the student's high school's 75th SAT percentile. Columns (i) through (iii) control for college fixed effects, student demographics (SAT score, whether the student is an ACT taker, gender, and ethnicity), and original reader fixed effects. Column (iv) also includes controls for experimental reader fixed effects and official admission outcome.

We divide all Contextual SAT Score Differences by 10, so a difference in 1 point is equivalent to an actual increase of 10 SAT points on the 1600 point scale.

Based on characteristics in Table 1, Universities 3, 5, and 6 are the More Holistic Universities. Universities 1, 2, 4, and 7 are the Less Holistic Universities.