INFERENCE BY COLLEGE ADMISSION DEPARTMENTS

Michael Conlin
Michigan State University
Marshall-Adams Hall
486 W Circle Dr. Rm 110
East Lansing, MI 48824
conlinmi@msu.edu
517-355-0285 (Office)
517-432-1068 (fax)
Corresponding Author

Stacy Dickert-Conlin
Michigan State University
Marshall-Adams Hall
486 W Circle Dr. Rm 110
East Lansing, MI 48824
dickertc@msu.edu

June 2017

Abstract: Theoretical, experimental and empirical research by economists and psychologists suggests biases in how people draw inferences. Eyster and Rabin (2005) review extensive experimental evidence that suggests people do not fully take into account how other people's actions depend on their private information. Using data from two colleges with optional SAT I policies, this paper quantifies the extent to which players underestimate this relationship. This policy provides applicants with a choice of whether to disclose their SATI scores to the college. Our empirical estimates indicate that colleges do underestimate the relationship between an applicant's action (not submitting) and type (SATI score).

(JEL Classifications: **D03, D84, D82, I23**)

Keywords: inference, voluntary disclosure

Acknowledgments: We thank Dan Benjamin for fruitful discussions at the start of this project and Jeff Wooldridge and Tim Vogelsang for valuable econometric advice. Comments from Ted O'Donoghue, Steve Coate, Eric Eyster and Philippe Jehiel improved the paper's exposition. The paper has also benefited from comments by seminar participants at Cornell University, Michigan State University, Oregon State University, Ohio State University and University of New South Wales.

INFERENCE BY COLLEGE ADMISSION DEPARTMENTS

I. Introduction

The college admissions process is filled with uncertainty: both students and colleges have private information and are making inferences about the other party's preferences or characteristics. With increasing competition among colleges to attract the best students, the stakes are high for accurate inference – as colleges attempt to identify and attract the students who will give the most value added to their institution. The economics literature typically assumes this inference is based on Bayes Rule.[1] However, Eyster and Rabin (2005) present a psychologically motivated equilibrium concept where players do not accurately map private information to actions and thus make inferences inconsistent with Bayes Rule.[2] The so-called *cursed equilibrium* allows for the possibility that players underestimate the relationship between other players' actions and their private information and is supported by experimental evidence in common value auctions, bilateral trades, and voting games.[3] If players do systematical underestimate the relationship between other players' actions and types, it has implications for all private information games.

This paper uses field data to quantify the degree to which players underestimate the relationship between other players' actions and their private information. This paper is the first to estimate their cursed parameter despite the fact that Eyster and Rabin's "primary motivation for defining cursed equilibrium is not based on learning or any other foundational justification, but rather on it pragmatic advantages as a powerful empirical tool to parsimoniously explain data in a variety of

---

[1] Since Harsanyi (1967) first developed the approach of modeling games of incomplete information, game theorists have developed numerous applicable equilibrium concepts. See Fudenberg and Tirole (1991), Kreps and Wilson (1982), Cho and Kreps (1987), and Banks and Sobel (1987) for details associated with the equilibrium refinements.
[2] See Camerer et al. (2004), Jehiel (2005), Jehiel and Koessler (2008), and Ross (1977) for other models where uninformed players make inferences that differ from Bayes Rule.
[3] Eyster and Rabin use the term cursed equilibrium because of its ability to explain the prevalence of the "winner's curse".

contexts (page 1633)". An inability to obtain data on players' private information and the difficulty of inferring players' beliefs from their observed actions may explain this lack of empirical research.[4]

We overcome these difficulties using college admissions data where the college's inference on the applicant's perceived quality influences the college's acceptance decision. We estimate this inference using application and admission data from two liberal arts schools (Colleges X and Y) that have an *optional SATI policy*. This policy allows applicants to choose whether or not to submit their standardized math and verbal SATI scores to the college – generating, for those applicants who do not submit, yet another dimension along which the college must make an inference. Through an agreement with the College Board, we obtained the SATI scores of the applicants who chose not to submit. This allows us to structurally estimate a model and quantify the degree to which the colleges underestimate the relationship between an applicant's decision not to submit and the applicant's actual SATI scores. Because we are unable to directly observe what SATI score the colleges infer for applicants who choose not to submit, we use the admission decisions to estimate this inference. We parameterize this inference in the manner proposed by Eyster and Rabin (2005).

For the sixteen percent of College X applicants and the twenty four percent of College Y applicants who chose not to submit their SATI scores, we find that both colleges underestimate the relationship between the applicants' decision not to submit and the applicant's actual SATI score. The results from our structural estimation suggest that the admissions departments overestimate the SATI scores of applicants who do not submit by 15 to 25 points on an 800 point scale.

The fact the colleges infer only slightly higher average scores reflects the fact that the other information in the college application, such as high school grade point average, class rank, SATII score, ACT score, gender, and race, are accurate predictors of an applicant's actual SATI score.

---

[4] The empirical literature on saliency and shrouding is related to how players make inferences. See, for example, Gabiax and Laibson (2006) Chetty, Looney, & Kroft (2009), Finkelstein (2009) and Brown, and Hossain and Morgan (2010).

Because the additional information the college obtains from an applicant's actual SATI score is not large, our simulations suggest that the composition of the student body would not change appreciably if the colleges correctly inferred SATI scores or if the colleges required all applicants to submit their SATI scores.[5]

II. Application/Admission Process and Data

2.A Application/Admission Process

The college admission process culminates with colleges making acceptance decisions and students accepting or declining those offers of acceptance. During the process, colleges and students exchange much information; colleges make inferences on which students are "best" for the school; and student make inferences about whether the school is their best option. The formal application process begins in November for students applying early decision and in January for students applying regular decision. When applying, students fill out an application to the school that includes basic information about demographics, high school experience, financial aid intent, extracurricular activities, and a personal essay. Students also provide information to the colleges through campus visits and interviews, as well as interactions with their guidance counselors and college admissions personnel. Students who apply in the early decision process sign a written agreement stating that they will attend if admitted and, because it is costly to renege on this contract, applying early decision provides a signal of an applicant's willingness to attend if accepted.[6] Finally, students typically provide a required standardized test score to the college.

[5] Using the same dataset as this paper, Conlin, Dickert-Conlin and Chapman (2013) find that these colleges' acceptance decisions are influenced by strategic considerations pertaining to the pool of students (i.e., those who submit and enroll) whose SAT scores are reported to ranking organizations such as *U.S. News & World Report*. At the time of the data, almost a quarter of the top 100 liberal arts colleges ranked by *U.S. News & World Report* had optional SATI policies. Relative to larger universities, these liberal arts colleges are able to collect and individually evaluate substantially more information on each applicant. Thus, applicants' SATI scores provide much less additional information to liberal art schools than to larger universities.

[6] The agreement is not legally binding and the applicant may be released if the college does not meet financial need. However reneging on an early decision acceptance may be costly if it is too late to reapply to alternative schools and could adversely affect the reputation of the high school guidance counselor who also signs the agreement (Avery et

In the Northeast, where the schools in our data are located, the SAT is the standardized test potential college students typically take.[7] A high school student takes the SAT exams in her junior and/or senior year during one of the seven annual test dates offered by the College Board. At the time of our data (the early 2000s), the SATI consisted of a two-part standardized verbal and math exam each scored on an 800 point scale.[8] Some students might also choose to take a subset of the 20 different SATII exams on subjects including English, history, mathematics, science and languages scored on an 800 point scale. Under the College Board's *Score Choice Option* for SATII exams in place during the time of our data, the College Board did not release SATII scores to colleges until the student saw and approved their release.[9]

Over time, an increasing number of schools have offered applicants the option to *not* report their two-part SATI score by checking a box on the application form that formally requests that these scores not be considered in the application process. Schools vary on whether they request alternative measures of college preparedness or simply consider all the other material provided by the student in making their admissions decision. The policy introduces more uncertainty over the non-submitting applicant's ability in the colleges' acceptance decision and generates another dimension over which colleges make inferences.

2.B Data Description and Summary Statistics

Our primary data come from the admissions databases from two schools in the Northeast, with applicant classes in the early 2000s, each with approximately 1800 students enrolled.[10] Through

al., 2003). Avery et al. (2003) report that often more than 40 percent of enrolled students applied early decision, when early decision is an option, highlighting the importance of the policy for schools.

[7] During the time of our data, the SAT was the most common standardized test taken along the East and West coasts, as well as in Texas. The ACT was most common in the Midwest.

[8] Students often take the SATI multiple times (Vigdor and Clotfelter, 2003). We do not have information on how many times the applicant took the different exams.

[9] This option was in place for SATII exams between 1993 and 2002 (College Board, unknown date). The College Board universally adopted the policy for all SAT scores, including the SATI, beginning in 2009 (College Board, 2009).

[10] We signed agreements with the colleges and College Board to allow us to use the data. This agreement stipulates that we cannot reveal the names of the colleges.

an agreement with the colleges and the College Board, we supplement the colleges' data with College Board data on SATI scores for all applicants, including those who chose not to submit.[11] Fewer than two percent of our sample of domestic applicants do not have an SATI score and we drop those few observations. These College Board data also include SATII scores and responses to the student descriptive questionnaire (SDQ), which includes self-reported data on family income as well as standardized measures of high school grades[12] and class rank. For College X, we have two years of applicant data, about five years into the school's optional SAT policy, during which 16 percent of the 6,557 applicants chose not to submit their SATI scores. For College Y, we have one year of applicant data, the first year that the school instituted the optional SAT policy, during which 24 percent of the 3,602 applicants from College Y chose not to submit their SATI scores. Summary statistics for Colleges X and Y are in Table 1.

Overall, applicants to both colleges often attend private high schools (around 40 percent), and come from high income families (more than 20 percent report family income in excess of $100,000). Demographically, more than 80 percent of applicants report their race as white and more than 70 percent are from the Northeast. Two thirds of applicants to College X are women and in both colleges, women are more likely to not submit their SATI scores.

Academically, applicants to College X and College Y are relatively well prepared. The average SATI scores of all applicants exceeds 1,200, relative to a mean for all persons taking the SATI during that time period of approximately 1,020 (College Board, 2002). Perhaps not surprisingly, students with lower SATI scores are more likely to take advantage of the optional SATI

[11] The colleges obtain SAT scores for a minority of applicants who did not submit their SATI scores, particularly those who ultimately enroll. We drop the international applicants from our analysis because the probability of obtaining a match with the College Board data is low without a Social Security number. Finally, we exclude applicants who withdrew from consideration before admission decisions were made, which is about 13 percent from College X and 5 percent from College Y.

[12] We use these self-reported GPAs because GPA scales as reported on applications are not standardized across high schools and therefore comparisons are extremely difficult (see Chaker, 2003). College Y did not record high school GPA for many applicants in their admissions data.

policy.[13] Likewise, applicants overall have impressive high school GPAs and class ranks; although these are generally slightly worse for applicants who did not submit their SATI scores.

Both Colleges X and Y follow SATI optional policies and include on their applications a place for applicants to indicate that they do not want the college to consider their SATI score. Note that the SATI was the required standardized test before the optional policy went into place and the students are choosing not to submit their SATI under the optional policy. For applicants who choose not to submit their SATI score, the colleges require an alternative standardized test score or scores, which includes tests that are more curriculum based than the SATI. College X requires *all* applicant to submit either their ACT scores or three SATII scores, so their optional test policy just removes the requirement of submitting an SATI score. Table 1 shows that 73 percent of all applicants submit an SATII score and only 17 percent submit an ACT score. College Y does not require additional test scores for applicants who submit an SATI score, although applicants can elect to submit scores from their SATII exams, ACT exam, and/or Advanced Placement (AP) exams (subject tests typically taken following a year-long high school class).[14] Applicants at College Y who choose not to submit their SATI scores must submit at least one of these alternative scores.[15] Table 1 shows that even among applicants who submit their SATI scores and are therefore not obligated to supply other test scores, 33.8 percent provide SATII scores, but none submit an ACT score. Not surprisingly, applicants who withhold an SATI score are more likely to provide SATII and ACT scores.

Finally, only a small percent of students apply early decision, seven percent of College X applicants and ten percent of College Y applicants. However, consistent with the policy's intentions,

[13] Grossman and Hart (1980), Grossman (1981) and Milgrom (1981) present models of voluntary disclosure where, when disclosure is costless, all players reveal their private information (i.e. disclose) to avoid the uninformed party from assuming that their decision to withhold information implies something worse than their actual private information. Perhaps the fact that applicants have very minimal experience/knowledge of how their SATI submission decision will affect the college's acceptance decision, as well as the complexity associated with the colleges' inference, explains why such a large proportion of applicants at both colleges choose not to submit.
[14] See https://apstudent.collegeboard.org/apcourse accessed 3/30/17.
[15] Based on the data we have, a few applicants appear not to have satisfied these requirements.

in numbers not shown in Table 1, more than 95 percent of students accepted in the early decision process enroll and early decision applicants comprise approximately 28 percent of the freshman classes.

Turning to the acceptance and enrollment decisions of colleges and students, we see that admission departments' internal ratings and acceptance rates are unconditionally higher for students who submit their SATI at College X, relative to those who do not submit their SATI, and unconditionally higher for students who do not submit their SATI at College Y, relative to those who do not. For College Y, this is surprising, given that observable characteristics are generally worse for students who do not submit their SATIs. In both colleges, students who *did not* submit their SATI scores are more likely to enroll than students who did and their SATI scores are lower conditional on acceptance and enrollment.

III. College's Acceptance Decision

3.A Model

In this section, we specify a college's objective function to explain their acceptance decisions. The college's acceptance decision provides a measurable outcome for whether the colleges underestimate the relationship between the applicant's decision not to submit and her actual SAT score.

In our model, the college considers how its acceptance decisions affect the expected "quality" of the current and future student body. An individual student's quality depends on the ability level of the student prior to attending the college and how much value added the college provides through academic quality and demographic diversity, which in turn depends on the composition of the student body. The applicant provides the college with a great deal of information that the college can use to judge these qualities. In the case of an applicant who chooses not to submit an SATI score, the college's expectation of the applicant's quality depends on its inference about her SATI score. We assume that the college does not make inferences over SATII or ACT scores if the student does not provide them because most applicants do not take any specific SATII exam or the ACT. Instead, the colleges infer that applicants who did not report an SATII score or an ACT score did not take those exams.

Current acceptance decisions influence future student body quality because they affect the reputation and external ranking of the school. A college's ranking, by organizations such as *U.S. News and World Report*, depends on the quality of the enrolled students, based on measured outcomes like high school performance and standardized test scores. An optional SATI policy potentially influences these ratings because only the scores of students who submit SATI scores determine the average SATI score reported to these ranking organizations. These policies generate a potential wedge between the ability of the enrolled students, as measured by their SATI scores, and the "reported" ability. Because enrolled students with low SATI scores are less likely to submit, the reported average SATI score will be greater than the average SATI score of all enrolled students. During the time of our data, the college's yield rate, defined as the fraction of accepted applicants that choose to enroll, also affected rankings.

With this setting, the college's objective function depends on the perceived ability of the incoming students, the "reported" ability of these students, the demographic characteristics of the student body, and the yield rate. Specifically, we assume that the college accepts applicant *i* if the expected payoff from accepting exceeds the payoff from rejecting applicant *i*:

$$
\begin{aligned}
&P_i^e\left[\Pi^P\left(SAT_{+i}^P, \boldsymbol{Z_{+i}^P}\right) + \varepsilon_{qi} + \Pi^R\left(SAT_{+i}^R, \boldsymbol{Z_{+i}^R}\right) + \Pi^D\left(\boldsymbol{Z_{+i}^D}\right)\right] \\
&\qquad + (1 - P_i^e)\left[\Pi^P\left(SAT_{-i}^P, \boldsymbol{Z_{-i}^P}\right) + \Pi^R\left(SAT_{-i}^R, \boldsymbol{Z_{-i}^R}\right) + \Pi^D\left(\boldsymbol{Z_{-i}^D}\right)\right] + f(YR_{ai}) > \\
&\qquad\qquad \Pi^P\left(SAT_{-i}^P, \boldsymbol{Z_{-i}^P}\right) + \Pi^R\left(SAT_{-i}^R, \boldsymbol{Z_{-i}^R}\right) + \Pi^D\left(\boldsymbol{Z_{-i}^D}\right) + f(YR_{ri}) \qquad [1]
\end{aligned}
$$

where P_i^e denotes the college's expectation that an accepted applicant *i* enrolls.[16] The term $\Pi^P(\cdot)$ captures how the college's payoff is affected by the portion of the incoming students' perceived ability that depends on observable variables. Likewise, $\Pi^R(\cdot)$ and $\Pi^D(\cdot)$ capture the college's payoff from acceptance that depends on reported ability (R) and demographic (D) characteristics. If the college accepts applicant *i* and she enrolls, applicant *i*'s characteristics are included in incoming class, $+i$. If, on the other hand, applicant *i* does not enroll or is not accepted, her characteristics are not included in the incoming class, *-i*. If an applicant chooses not to submit under the optional SATI policy, SAT_{+i}^P depends on the college's inference of that SATI score and the average reported SATI score of the incoming class is the same whether the applicant enrolls, SAT_{+i}^R, or not, SAT_{-i}^R. The ***Z*** vectors denote other observable variables that affect the perceived ability, reported ability and demographics of the expected incoming class.

The payoffs for the college are similar when applicant *i* is accepted but chooses not to enroll as when applicant *i* is rejected, *except* for the college's expected yield rate. We denote the expected yield rate if applicant *i* is accepted (rejected) by YR_{ai} (YR_{ri}) and how the yield rate affects the college's expected payoff by $f(\cdot)$. The variable ε_{qi} represents aspects of applicant *i*'s perceived quality which the college observes but we do not, including factors such as how well the applicant performed in her interview, discussions with her high school guidance counselor, and recommendation letters.[17]

A higher SATI score for applicant *i* has two effects on the college's expected payoff. First, it increases applicant *i*'s perceived ability and, if applicant *i* submits, the college's reported ability. Second, a higher score may decrease the probability an accepted applicant *i* enrolls, which decreases the college's expected yield rate. We expect the first effect to dominate, which results in the

[16] Estimates of a model that endogenizes the applicant's decisions and allows the probability of enrolling to depend on whether the applicant applies early decision or submits are contained in Table 6.
[17] ε_{qi} may also represent other characteristics of applicant *i* that influences the reported ability or the demographics of the incoming class which the college observes and we do not.

probability of acceptance increasing with actual or inferred SATI score. Our empirical results suggest that this is the case with SATI score having a small effect on the enrollment decision (See Table 2).[18] If this is indeed the case, a cutoff SATI score exists for applicants who submit their SATI score, such that, conditional on the applicant's observables and unobservables, the college accepts the applicant only if her SATI score is above the cutoff. Likewise, for applicants who chose not to submit their SATI score, a cutoff exists where the college accepts the applicant only if the inferred SATI score is above the cutoff, conditional on the applicant's observables and unobservables.

3.B Likelihood Function

To construct a likelihood function for applicants who chose to submit and chose not to submit, we first impose the following functional form restrictions:

$$f(YR_{ai}) = \beta^{YR} YR_{ai}\,,\ f(YR_{ri}) = \beta^{YR} YR_{ri} \qquad [2]$$

$$\Pi^P\left(SAT^P_{+i}, \boldsymbol{Z^P_{+i}}\right) = \beta^P_{SAT} SAT^P_{+i} + \boldsymbol{\beta^P} \cdot \boldsymbol{Z^P_{+i}}\,,\ \Pi^P\left(SAT^P_{-i}, \boldsymbol{Z^P_{-i}}\right) = \beta^P_{SAT} SAT^P_{-i} + \boldsymbol{\beta^P} \cdot \boldsymbol{Z^P_{-i}}$$

$$\Pi^R\left(SAT^R_{+i}, \boldsymbol{Z^R_{+i}}\right) = \beta^R_{SAT} SAT^R_{+i} + \boldsymbol{\beta^R} \cdot \boldsymbol{Z^R_{+i}}\,,\ \Pi^R\left(SAT^R_{-i}, \boldsymbol{Z^R_{-i}}\right) = \beta^R_{SAT} SAT^R_{-i} + \boldsymbol{\beta^R} \cdot \boldsymbol{Z^R_{-i}}$$

$$\Pi^D(\boldsymbol{Z^D_{+i}}) = \boldsymbol{\beta^D} \cdot \boldsymbol{Z^D_{+i}} \ and\ \Pi^D(\boldsymbol{Z^D_{-i}}) = \boldsymbol{\beta^D} \cdot \boldsymbol{Z^D_{-i}}$$

where $\left(\beta^{YR}, \beta^P_{SAT}, \boldsymbol{\beta^P}, \beta^R_{SAT}, \boldsymbol{\beta^R}, \boldsymbol{\beta^D}\right)$ are coefficients and $\left(\boldsymbol{Z^P_{+i}}, \boldsymbol{Z^P_{-i}}, \boldsymbol{Z^R_{+i}}, \boldsymbol{Z^R_{-i}}, \boldsymbol{Z^D_{+i}}, \boldsymbol{Z^D_{-i}}\right)$ are vectors of the characteristics of the college's expected student body with $(+i)$ and without $(-i)$ applicant *i*. More specifically, $\boldsymbol{Z^P_{+i}}$ and $\boldsymbol{Z^P_{-i}}$ include average <u>perceived</u> SATI, average SATII, average ACT, fraction in different gpa/class rank brackets, and fraction who attended private high school; $\boldsymbol{Z^R_{+i}}$ and $\boldsymbol{Z^R_{-i}}$ include average reported SATI, SATII, ACT, and fraction in different gpa/class rank brackets;

[18] The estimates in Table 2 indicate that a higher SATI score actually increases the probability of enrollment at College Y. Unlike College X, College Y is less likely to be a "backup" school for applicants preferring to attend an elite Ivy League School.

and $\boldsymbol{Z^D_{+i}}$ and $\boldsymbol{Z^D_{-i}}$ include fraction female, fraction legacy, fractions in different income brackets, fraction non-white and fraction from the Northeast.[19]

While the college's objective function incorporates group characteristics, empirically we would like to relate the probability the college accepts applicant *i* to applicant *i*'s individual characteristics, Z_i^j , rather than the group characteristics, $\boldsymbol{Z^j_{-i}}$ and $\boldsymbol{Z^j_{+i}}$. Like the group characteristics, Z_i^j includes perceived ability ($\boldsymbol{Z_i^P}$), reported ability ($\boldsymbol{Z_i^R}$), and demographic ($\boldsymbol{Z_i^D}$) characteristics. Appendix A shows how, after substituting the functional form assumptions of [2] into [1], the college acceptance decision can be represented by changes in the expected characteristics of the student body if the applicant enrolls in the college ($Z^j_{+i} - Z^j_{-i}$ where $j \in P, R, D$) and the expected probability applicant *i* enrolls (P_i^e). We can express these differences, $Z^j_{+i} - Z^j_{-i}$, as a function of characteristic *j* of applicant *i*, Z_i^j, if we assume that the college's expectations of Z^j_{-i} ($j \in P, R, D$), SAT^k_{-i} ($k \in P, R$), YR_{ri}, student body size (N^{sb}), number of enrolled students that submit (N^{sub}), and number of applicants accepted (N^a) do not vary across applicants. We believe this assumption is reasonable because the expected size of the student body is relatively large making it unlikely that a single applicant could change the student body characteristics appreciably.[20]

With these assumptions and the functional form assumptions in [2], we can represent the probability of acceptance for an applicant <u>who submits</u> as a function of her observables and the college's expectation of her enrolling. Specifically, the probability applicant *i* is accepted if she <u>submits</u> is:

$$P_s^a = Prob\left\{\varepsilon_{qi} > \beta^{int} - \beta_{SAT}^{P'} SAT_i - \boldsymbol{\beta^{P'}} \cdot \boldsymbol{Z_i^P} - \beta_{SAT}^{R'} SAT_i - \boldsymbol{\beta^{R'}} \cdot \boldsymbol{Z_i^R} - \boldsymbol{\beta^{D'}} \cdot \boldsymbol{Z_i^D} + \beta^{YR'} / P_i^e\right\} \quad [3]$$

[19] For College Y, we exclude the ACT indicator and ACT score from vector Z_i because so few applicants submitted ACT scores. The likelihood function does not converge when these ACT covariates were included because of the collinearity.

[20] In addition, it would be empirically intractable to take into account how these characteristics of the expected student body differ across applicants when making acceptance decisions.

where:

$$\beta_{SAT}^{P'} = \beta_{SAT}^{P}/(N^{sb}+1), \boldsymbol{\beta}^{\boldsymbol{P}'} = \beta^{P}/(N^{sb}+1), \beta_{SAT}^{R'} = \beta_{SAT}^{R}/(N^{sb}+1),$$

$$\boldsymbol{\beta}^{\boldsymbol{R}'} = \boldsymbol{\beta}^{\boldsymbol{R}}/(N^{sub}+1), \boldsymbol{\beta}^{\boldsymbol{D}'} = \boldsymbol{\beta}^{\boldsymbol{D}}/(N^{sb}+1), \beta^{YR'} = \beta^{YR}/(N^{a}+1)$$

and β^{int} is a function of Z_{-i}^{j} (for all *j*), N^{sb}, N^{sub} and N^{a}. For those covariates in both the perceived and reported vectors, only the sum of the corresponding coefficients is identified. For example, the estimated coefficient associated with SATI score for an applicant who submits would be $\beta_{SAT}^{P'} + \beta_{SAT}^{R'}$.

The college's expectation of enrollment will likely depend not only on the characteristics of the applicant but also the applicant's decision of whether to submit an SATI score and whether to apply early decision. Because of the college's vast admissions experience, we expect how the applicant's characteristics and decisions affect the college's expectations of enrollment to be consistent with the actual relationship to enrollment. Therefore, we assume that $P^{e} = P_{i}^{e} = exp(\boldsymbol{\beta}^{\boldsymbol{e}} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{e}})/[1 + exp(\boldsymbol{\beta}^{\boldsymbol{e}} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{e}})]$ where $\boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{e}}$ includes a constant, SATI score, an SATI submit indicator, SATII indicator, average SATII score, ACT indicator, ACT score, private high school, female, high school gpa indicators, class rank indicators, income indicators, legacy, white, Northeast region indicator and an early decision indicator. Table 1 indicates that, especially for College X, accepted applicants who do not submit are more likely to enroll than those who submit. A supplemental appendix available from the authors presents a model that endogenizes the applicant's decisions to apply early, submit and enroll. Table 6 reports the estimates from this model, which are qualitatively similar to those presented in Table 2.

The acceptance probability of an applicant who does <u>not submit</u> is similar to Equation [3] except her perceived quality depends on the SATI score the college infers and her actual SATI score does not influence the average student body SATI score the college reports to the ranking organizations. We construct a likelihood function for applicants who do not submit SATI scores by first parameterizing the college's inference. Based on Eyster and Rabin (2005), we assume the inferred score is a weighted average of the college's belief of the SATI score if the college correctly updates its beliefs about the score based on applicant *i*'s decision to withhold it, $SAT_{i,cond}$, and the college's belief of applicant *i*'s SATI score if the college does not condition on applicant *i*'s decision to withhold, $SAT_{i,un}$. Following Eyster and Rabin's notation, $\chi \in [0,1]$ is the weight the college places on the applicant's "unconditional" SATI score when making this inference and, therefore, the college's belief of applicant *i*'s SAT score if she does not submit is $\chi SAT_{i,un} + (1-\chi)SAT_{i,cond}$. If the college <u>does not</u> condition on applicant *i*'s decision to withhold her score, we assume that the college's belief of her SATI score is based on all the other information the college has on the applicant and how that information is correlated with the SATI scores for applicants who do submit their scores. Specifically, $SAT_{i,un} = \boldsymbol{\alpha} \cdot \boldsymbol{Z_i} | submit + \varepsilon_{i,un}$ where $\varepsilon_{i,un}$ is a random component capturing information the college uses to infer SATI scores but that we do not observe. If the college does not condition on applicant *i* choosing not to submit, the probability the college infers that applicant *i* obtained a score of SAT_i is:

$$P_{sat} = Prob\{\varepsilon_{i,un} = SAT_i - \boldsymbol{\alpha} \cdot \boldsymbol{Z_i}\}$$

where the estimates of $\boldsymbol{\alpha}$, $\widehat{\boldsymbol{\alpha}}$, are based on those applicants who do submit.

If the college <u>does</u> condition on applicant *i* not submitting, we assume that they have unbiased expectations of the applicant's actual SATI scores, due to the college's extensive information on and experience with applicants. While unbiased, the colleges are unable to perfectly infer SATI scores for applicants who do not submit and, therefore, we allow $SAT_{i,cond}$ to equal applicant *i*'s actual SATI score plus a random component, $\varepsilon_{i,cond}$.

Due to the random components associated with $SAT_{i,un}$ and $SAT_{i,cond}$, the probability of acceptance if applicant *i* does not submit is:

$$P_{ns}^{a} = Prob\left\{ \begin{matrix} \varepsilon_{qi} + \left(\beta_{SAT}^{P'}\right)(\chi)\varepsilon_{i,un} + \left(\beta_{SAT}^{P'}\right)(1-\chi)\varepsilon_{i,cond} > \\ \beta^{int} - \beta_{SAT}^{P'}[\chi(\boldsymbol{\alpha}\cdot\boldsymbol{Z_i}) + (1-\chi)SAT_i] - \boldsymbol{\beta^{P'}}\cdot\boldsymbol{Z_i^P} - \boldsymbol{\beta^{R'}}\cdot\boldsymbol{Z_i^R} - \boldsymbol{\beta^{D}}\cdot\boldsymbol{Z_i^D} + \beta^{YR'}/P_i^e \end{matrix} \right\} \quad [4]$$

where $P_i^e = exp(\boldsymbol{\beta^e}\cdot\boldsymbol{Z_i^e})/[1+exp(\boldsymbol{\beta^e}\cdot\boldsymbol{Z_i^e})]$. We assume that the population distributions are independent, $\varepsilon_{qi} \backsim N(0,1)$, $\varepsilon_{i,un} = N(0,\sigma_u^2)$ and $\varepsilon_{i,cond} = N(0,\sigma_c^2)$.

Given these functional form, distributional, and parameterization assumptions, we obtain the following likelihood function where $y_a = 1$ if applicant *i* is accepted, $y_s = 1$ if applicant *i* submits and $y_e = 1$ if applicant *i* enrolls (equal 0 otherwise).

$$\Pi[P_s^a]^{y_a y_s}[1-P_s^a]^{(1-y_a)y_s}[P_{ns}^a]^{y_a(1-y_s)}[1-P_{ns}^a]^{(1-y_a)(1-y_s)}\left[P_{sat}\right]^{y_s}[P^e]^{y_a y_e}[1-P^e]^{y_a(1-y_e)} \quad [5]$$

By maximizing the above likelihood function, our goal is to estimate a vector of structural parameters $\boldsymbol{\theta} = \left\{\chi, \beta_{SAT}^{P'}, \beta_{SAT}^{R'}, \boldsymbol{\alpha}, \sigma_u, \sigma_c, \beta^{int}, \boldsymbol{\beta^{P'}}, \boldsymbol{\beta^{R'}}, \boldsymbol{\beta^{D'}}, \boldsymbol{\beta^{e}} \text{ and } \beta^{YR'}\right\}$.

3.C Identification

The part of the likelihood function related to the acceptance decision is a standard probit model except fitted values are used when constructing the college's SATI inference for those who do not submit ($\widehat{\boldsymbol{\alpha}}$), fitted values are used when constructing the college's expectation of enrollment ($\widehat{\boldsymbol{\beta^e}}$), and the variance of the error term is not one for applicants who do not submit. Specifically, we can write the college's acceptance decision for those who submit and those who do not by combining the two probabilities in equations [3] and [4],

$$y_a = 1 \text{ if } y_a^* > 0;\ y_a = 0 \text{ if } y_a^* < 0 \text{ and}$$

$$y_a^* = \beta^{int} - \left(\beta_{SAT}^{P'} + \beta_{SAT}^{R'}\right)(S_i)SAT_i - \beta_{SAT}^{P'}(1-S_i)[\chi(\widehat{\boldsymbol{\alpha}}\cdot\boldsymbol{Z_i}) + (1-\chi)SAT_i] \quad [6]$$

$$-\boldsymbol{\beta^{P'}}\cdot\boldsymbol{Z_i^P} - \boldsymbol{\beta^{R'}}\cdot\boldsymbol{Z_i^R} - \boldsymbol{\beta^{D'}}\cdot\boldsymbol{Z_i^D} + \beta^{YR'}\left[1+\exp\left(\widehat{\boldsymbol{\beta^e}}\cdot\boldsymbol{Z_i^e}\right)\right]/\exp\left(\widehat{\boldsymbol{\beta^e}}\cdot\boldsymbol{Z_i^e}\right) + v_i$$

where $S_i = 1$ if applicant *i* submits her SATI score (zero otherwise) and ν_i is assumed to be uncorrelated with the regressors. In addition, v_i is standard normally distributed for applicants who submit [v_i equals ε_{qi}] and normally distributed with a variance of $1 + \left[\left(\beta_{SAT}^{P'}\right)(\chi)\sigma_u\right]^2 +$ $\left[\left(\beta_{SAT}^{P'}\right)(\chi)\sigma_c\right]^2$ for applicants who do not submit $\left[v_i \text{ equals } \varepsilon_{qi} + \left(\beta_{SAT}^{P'}\right)(\chi)\varepsilon_{i,un} + \left(\beta_{SAT}^{P'}\right)(1-\chi)\varepsilon_{i,cond}\right]$.

There are two primary identification issues to discuss. The first involves how each particular applicant characteristic affects the college's acceptance decision and how this effect varies depending on whether the applicant submits an SATI score. The other involves the variation in the difference between the predicted and actual SATI scores, for those who do not submit (which identifies χ), and whether this variation is uncorrelated with v_i.

In the model, if the applicant submits an SATI score, the effect of her personal characteristics on acceptance is entirely captured by the coefficient estimates representing the colleges' payoff $\left(\boldsymbol{\beta}^{\boldsymbol{P}'}, \beta_{SAT}^{P'}, \boldsymbol{\beta}^{\boldsymbol{R}'}, \beta_{SAT}^{R'}, \boldsymbol{\beta}^{\boldsymbol{D}'}, \boldsymbol{\beta}^{\boldsymbol{e}}, \beta^{YR'}\right)$. If an applicant does not submit her SATI score, her personal characteristics play two roles in influencing the probability of acceptance. The first is directly through the effect on the college's payoff, similar to applicants who submit an SATI score. In addition, each characteristic can indirectly influence the college's inference of her SATI score if the college believes that the characteristic is highly correlated with SATI score. By estimating $\hat{\boldsymbol{\alpha}}$ using information on applicants who submit, we assume that the college's inference of these correlations is based on these applicants. Suppose SATI scores are positively correlated with SATII scores and high school performance but uncorrelated with legacy status (whether a close relative attended the school) among students who submit their scores. In this case, we assume the college will infer a relatively high SATI score for an applicant, with a high SATII score and an A+ on high school grades, who chose not to submit. Ultimately, the college's acceptance decision should be based more on SATII scores and high school performance and less on legacy status than for applicants who submit.[21]

The variation that identifies the main parameter of interest, χ, is made evident by noting that we can rewrite the latent acceptance decision in [6] as:

$$y_a^* = \beta^{int} - \left(\beta_{SAT}^{P'} + \beta_{SAT}^{R'}\right)(S_i)SAT_i - \beta_{SAT}^{P'}(1-S_i)SAT_i - \beta_{SAT}^{P'}(1-S_i)(\chi)(\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{Z_i} - SAT_{\boldsymbol{i}})$$
$$-\boldsymbol{\beta}^{\boldsymbol{P}'} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{P}} - \boldsymbol{\beta}^{\boldsymbol{R}'} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{R}} - \boldsymbol{\beta}^{\boldsymbol{D}'} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{D}} + \beta^{YR'}\left[1 + \exp\left(\widehat{\boldsymbol{\beta}^{e}} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{e}}\right)\right] / \exp\left(\widehat{\boldsymbol{\beta}^{e}} \cdot \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{e}}\right) + v_i \quad [7]$$

The identification of χ comes from variation in $\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{Z_i} - SAT_{\boldsymbol{i}}$: the difference between predicted and actual SATI scores for those who do not submit.[22]

One threat to this identification strategy for χ is the correlation between $\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{Z_i} - SAT_{\boldsymbol{i}}$ and v_i, and specifically the random component of the applicant's perceived ability level, ε_{qi}. For example, colleges may observe applicant attributes that we do not and these attributes are correlated with the college's perception of the applicant's quality and the difference between applicant's predicted and actual SATI scores. Given our detailed admission data, these qualities are likely to be non-academic characteristics that would make the applicant and college a better match. Section IV addresses these concerns of unobserved match quality by using measures of the colleges' internal *non-academic* rankings as covariates and subsequent observed college performance measures among enrolled students.

3.D Parameter Estimates

Table 2 reports the estimates of the structural parameters obtained from maximizing the likelihood function in equation [5], with standard errors in parentheses. Our primary parameter of interest is χ, which represents the degree to which the college underestimates the relationship

[21] The results do not change appreciably when $\boldsymbol{\alpha}$ is estimated using information from all applicants (See Table 6).
[22] Note that variation based on the decision to submit does not identify χ.

between an applicant's decision not to submit and her actual SATI score. The estimates of 0.191 and 0.339 suggest that, while the admission departments at the colleges do condition on the fact that an applicant's decision to not submit is related to her actual SATI score, they underestimate this relationship when making inference on the applicant's SATI score. To provide some context to these estimates, the average SATI score for the 1017 College X applicants who do not submit is 1139 and, based on the $\boldsymbol{\alpha}$ estimates, the average predicted SATI score ($SAT_{i,un}$) for these applicants is 1215. Therefore, the estimate of 0.191 suggests that the admission department infers an average SATI score of .191*1215+(1-.191)*1139=1154 for those applicants who do not submit. Among the 868 College Y applicants who did not submit, their average SATI score is 1228, the average predicted SATI score ($SAT_{i,un}$) is 1265, and the college infers an average score of 1241. The null hypothesis that the college does not underestimate this relationship (i.e., χ =0) can be rejected at the five percent significance level for College X and at the one percent significance level for College Y. These lower inferred scores are consistent with, in Eyster and Rabin's terminology, the colleges being *partially* cursed but not *fully* cursed because these inferred scores are less than the actual scores but greater than those predicted if the college did not condition at all on the applicants' decisions not to submit.

Other estimates of interest include how an applicant's actual or perceived SATI score affects the probability of acceptance. The positive coefficient estimates associated with submitted SATI score and inferred SATI score indicate that a 100 point increase in either score increases the probability of acceptance by approximately 0.18 for College X and 0.15 for College Y. These are large increases, relative to the mean acceptance rates of 0.40 to 0.45. Note that we cannot statistically reject the hypothesis that the coefficient associated with a submitted SATI score (estimate of $\beta_{SAT}^{P'} + \beta_{SAT}^{R'}$) is the same as the coefficient associated with inferred SATI score (estimate of $\beta_{SAT}^{P'}$).[23] One might have thought that the coefficient associated with the submitted SATI score

[23] Also note that our estimates of $\beta_{SAT}^{P'} + \beta_{SAT}^{R'}$ and $\beta_{SAT}^{P'}$ are imprecisely estimated.

would be larger or, equivalently, the estimate of $\beta_{SAT}^{R'}$ would be positive because a submitted SATI score affects both the expected quality of the enrolled applicant and the average SATI score the college reports to the ranking organizations, while the inferred SATI score only affects the expected quality. There are potentially a number of explanations for why our point estimates of $\beta_{SAT}^{P'} + \beta_{SAT}^{R'}$ and $\beta_{SAT}^{P'}$ are similar. Perhaps most importantly, our model assumes that the college's payoff is linear with respect to the submitted and inferred SATI score. Because inferred scores are often significantly less than submitted scores and the change in a college's payoff is likely different when an applicant's submitted or inferred SATI score increases from 1,100 to 1,150 compared to an increase from 1,250 to 1,300, a non-linear payoff function (where the college has diminishing marginal benefit associated with a higher SATI score) could explain why the estimates are similar.

As for the other coefficient estimates associated with the colleges' objective functions, they suggest admission departments are interested in the quality and the diversity of the student body. For example, applicants with high SAT test scores and superior high school performance (gpa and class rank) at a private high school are more likely to be admitted. The fact that a much larger fraction of College X applicants are female explains why female candidates are less likely to be accepted by College X but not by College Y. As expected, legacies are also more likely to be accepted by both colleges. Finally, the desire to achieve racial diversity in the student body may explain why both colleges are much more likely to accept a non-white applicant. While more likely to accept a non-white applicant, College X is less likely to accept an applicant with family income less than \$100,000.

The estimates of $\boldsymbol{\alpha}$ reported in Table 2 indicate that while high school gpa and class rank are positively correlated with SATI score, the relationship between SATI score and average SATII score is stronger. For applicants who submit an SATI score, a one standard deviation increase in the average SATII score (approximately 70 points) increases the applicant's expected SATI score by over 100 points for both colleges. For College X applicants who submit an ACT score and an SATI score, these scores are also positively correlated.[24] Table 2 also indicates that being a legacy is not highly correlated with the SATI score of applicants who submit; however, being a white male is positively and statistically significantly correlated with SATI scores. Recall that these correlations may indirectly affect the probability of acceptance for students who do not submit an SATI score because they influence the college's inference about the SATI score.

The negative estimates of $\beta^{YR'}$ in Table 2 indicate that the probability of acceptance increases with the college's expectation that the applicant will enroll if accepted. This is expected if the colleges are concerned about their yield rate, perhaps because of ranking concerns, or if the college observes positive applicant attributes, which we do not, and these attributes are correlated with the applicant enrollment decision such as match quality. Furthermore, the estimates of $\boldsymbol{\beta}^{\boldsymbol{e}}$ indicate that applicants who do not apply early decision, attend a private high school and performed better academically in high school are less likely to enroll if accepted. The effect of being lower income, having a high SATI score and being white on the probability of enrollment differs across the colleges. Most interesting is the fact that while submitting an SATI score increases enrollment probability, the increase is neither economically and statistically significant. While Table 1 indicates that the probability of enrolling conditional on acceptance is much higher for College X applicants who choose not to submit, these applicants' high school academic performances are worse than those applicants who submit.

Table 2 uses the college's acceptance decision as the measurable outcome of their inference. The college's data also include an internal measure of academic ranking on a seven point scale,

[24] Because some students may take the ACT rather than the SAT regardless of the optional SAT policy, we estimate our model without the ACT as a predictor of the SATI and after dropping applicants who submit an ACT score. In addition, we estimate our model after dropping applicants from the Midwest where the ACT is the predominant college admission test taken by high school students. In results available from the authors, our parameter estimates do not change appreciably.

which presumably also incorporates the inference process.[25] Individual admissions counselors assign the academic ranking, then the entire admission's staff meets to discuss individual cases and make acceptance decisions.[26] Therefore, we estimate a linear model for the academic ranking with the same functional form as the acceptance decision and an error term which has a normal distribution.[27] The statistically significant estimates of χ from this model are contained in Table 3 and are 0.358 for College X and 0.286 for College Y. These χ estimates are similar to those in Table 2 and suggest that the colleges underestimate the relationship between an applicant's decision not to submit and their actual SATI score when assigning academic rankings, as well as when making acceptance decisions.

IV. Identification Assumptions and Robustness

As mentioned in Section III, our primary threat to identification is that the model's random components, ν_i, are correlated with the difference between predicted and actual SATI scores, $\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{Z_i} - SAT_i$. Suppose, for example, the college observes positive attributes of applicants who do not submit and whose predicted SATI scores are much larger than their actual SATI scores and these attributes are not observable to us. In this case, these applicants should more likely be accepted and our estimate of χ would be biased upward. We test this identification assumption with additional data on the college's internal applicant non-academic rating and subsequent student performance measures.

[25] The colleges also provide a measure of an applicant's non-academic ranking which we include as a covariate in one of our robustness tests.

[26] One reason for not using the academic rating as our primary outcome for the college is that the acceptance decision more broadly characterizes the aggregated inference by the admissions committee, rather than that of a single admissions person who we cannot identify.

[27] Specifically, the academic ranking equals $\gamma^{int} - \left(\gamma_{SAT}^{P'} + \gamma_{SAT}^{R'}\right)(S_i)SAT_i - \gamma_{SAT}^{P'}(1 - S_i)[\chi(\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{Z_i}) + (1 - \chi)SAT_i] - \boldsymbol{\gamma}^{\boldsymbol{P}'} \cdot \boldsymbol{Z}_i^{\boldsymbol{P}} - \gamma^{\boldsymbol{R}'} \cdot \boldsymbol{Z}_i^{\boldsymbol{R}} - \boldsymbol{\gamma}^{\boldsymbol{D}'} \cdot \boldsymbol{Z}_i^{\boldsymbol{D}} + \gamma^{\boldsymbol{e}} ED_i + \varepsilon_R$ where γ's are parameters to be estimated, ED_i is an indicator for early decision and $\varepsilon_R \backsim N(0, \sigma_R)$. When we included the functional form for the college's expectation of an applicant enrolling, $\exp(\boldsymbol{\beta}^{\boldsymbol{e}} \cdot \boldsymbol{Z}_i^{\boldsymbol{e}})/[1 + \exp(\boldsymbol{\beta}^{\boldsymbol{e}} \cdot \boldsymbol{Z}_i^{\boldsymbol{e}})]$, the likelihood was unable to converge.

First, we consider whether including the college's internal personal/non-academic ranking as a covariate alters our results. We assume this non-academic ranking captures unobservable applicant characteristics such as interview and recommendation letters quality that may be correlated with SATI scores and admission decisions. As a practical matter, College Y reports a non-academic rating on a 2 point scale directly and we generate a non-academic rating for College X's by subtracting the reported academic ranking (on a scale of 7) from the overall ranking (on a scale of 7).

Table 4 reports estimates of the key coefficients when we include the non-academic rating as a covariate in the colleges' objective function and the colleges' predicted SATI score. First note that while the non-academic rating significantly increases the probability of acceptance, it does not increase the predicted SATI score. More importantly, including the non-academic rating as a covariate does not appreciably change the estimates of χ. These results suggest that our estimates of χ in Table 2 are not biased due to non-academic unobservables being correlated with the difference between predicted and actual SATI scores.

Table 5 contains the most relevant estimates from a likelihood function using college performance measures as the outcome variable instead of the acceptance decision. The performance measures for College X include grade point average (gpa) at college graduation, whether the applicant graduated, recipient of honor or distinction in major, Phi Beta Kappa recipient, Summa or Magna Cum Laude recipient and number of collegiate sports. The college performance measure are more limited for College Y – just freshman year gpa, whether the student completed the freshman year and number of collegiate sports. We assume a linear model when the performance measure is gpa or number of collegiate sports with the same functional form as the academic rating and an error term which has a normal distribution. We estimate a likelihood function similar to that for acceptance when the performance measure is a dichotomous variable except we do not model the expected probability of enrollment but do control for early decision.[28] As Table 5 indicates, the estimates corresponding to χ vary in sign depending on the performance measure and are never statistically significant. With the caveat of smaller sample sizes among enrolled students and the non-random selection of students who decide to enroll, the lack of a consistently positive estimate of χ suggests that our estimates of χ from the acceptance decision specifications are not biased due to the difference in predicted and actual SATI scores being correlated with favorable academic characteristics only observable to the college.

While the results in Tables 4 and 5 provide evidence in support of the identification assumption of greatest concern, we also test the robustness of our structural estimation results along two other dimensions. First, we consider whether our results change if the college's belief of the withheld SATI score is based on the correlations between the SATI scores and other characteristics for all applicants; not just applicants who choose to submit. Specifically, we assume $SAT_{i,un} = \boldsymbol{\alpha} \cdot \boldsymbol{Z_i} + \varepsilon_{i,un}$ where the estimates of $\boldsymbol{\alpha}$ are based on information from all applicants. Table 6 presents the estimates of χ, $\beta_{SAT}^{P'} + \beta_{SAT}^{R'}$ and $\beta_{SAT}^{P'}$ from this specification. Constructing the predicted values in this way does not change any of the parameter estimates appreciably, primarily because the correlation between SATI scores and other characteristics of the applicants varies little based on whether the applicant submits SATI scores.

The final robustness check endogenizes the applicants' decision of whether to apply early decision, whether to submit her SATI scores and, conditional on acceptance, whether to enroll. While the applicant makes choices that maximize her expected utility, the college bases its acceptance decision on the objective function parameterization, as well as functional form and distributional assumptions presented in Sections II and III. The model permits us to derive

[28] We only observe college performance outcomes for applicants who enroll. Therefore, when constructing the likelihood function, we multiply the probability of achieving a specific college performance measure with an indicator variable that equals one if the applicant enrolls.

parametric expressions, as a function of observed data and structural parameters, for the probabilities an applicant applies early decision and/or submits her SATI score, the probability the college accepts an applicant, and the probability an applicant enrolls. A detailed description of the model, the derived likelihood function and the estimation is available on the authors' websites.[29] The estimates of the most relevant parameters from this model are presented in Table 6. Note that the estimates are similar to those in Table 2; most notably, the estimate of χ is 0.262 for College X and 0.349 for College Y. This is expected because the variation that identifies these parameters in our base specification is the same variation that identifies these parameters when we incorporate the applicant's decisions in the estimation.

An advantage of estimating a model that incorporates the applicant's enrollment decision is that we can better address counterfactuals involving student body composition. The two counterfactuals considered are: (i) if the colleges were not partially cursed and, therefore, did not underestimate the relationship between an applicant's decision not to submit SATI scores and the applicant's actual score (i.e., χ=0); and (ii) if all applicants were required to submit their SATI scores. Our simulations for both counterfactuals indicate that both the fraction of applicants who are accepted and the fraction who enroll decrease slightly for both colleges, compared to those predicted by the model. However, both counterfactuals would only minimally affect the student body characteristics and academic performance. While our estimates suggest that the colleges do underestimate the relationship between an applicant's decision to submit and her actual SATI score, the student body characteristics and academic performances do not change significantly because the colleges have a great deal of other information upon which to base their acceptance decision and infer SATI scores for those who do not submit.

[29] The model assumes that the applicant maximizes her utility when deciding whether to submit an SATI score and whether to apply early decision. Both of these choices affect the probability of being accepted. While the objective function of the admission's department is similar to equation (1), the model allows the probability an accepted applicant enrolls to depend on whether the applicant applied early decision.

A robustness test we are not able to conduct is distinguishing between whether the colleges misconstrue how an applicant's actual SATI score affects her decision to submit (applicants' strategies) and/or whether the colleges misconstrue the underlying distribution of SATI scores for applicants who choose not to submit.[30] Like Eyster and Rabin's cursed equilibrium framework, we assume that if the college conditions on applicant *i* not submitting, the colleges have unbiased expectations of the underlying distribution of SATI scores of applicants who choose not to submit. Empirically distinguishing between this assumption and an assumption that colleges fully understand the link between the applicants' actual score and their decision to reveal their SATI score would requires a restriction on exactly how the colleges misconstrue the distribution of SATI scores.

V. Conclusion

Our empirical estimates in this paper find that colleges underestimate the relationship between an applicant's decision not to submit her SATI score and her actual SATI score, consistent with results in the experimental literature on players underestimating the relationship between other players' actions and types. Specifically, our estimates suggest that the colleges partially update their beliefs about the applicants SATI scores based on their decision to withhold them, yet underestimate this relationship when making inference on the applicant's SATI score by approximately 15 points. This inaccurate inference would not appreciably change the composition of the student body nor would requiring all applicants to submit their SATI scores, primarily because the colleges have a great deal of other information to base their acceptance decision and to infer SATI scores. Given the

[30] Dekel, Fudenberg and Levine (2004) model an environment where uninformed players learn about the distribution of other players' types as well as their strategies and consider how this learning is influenced by how informative are the signals received by the uniformed players. This model is relevant for our empirical analysis because the applicant data are from five and six years after College X instituted an optional SATI policy and from the first year after College Y instituted an optional SATI policy. While the admission departments are likely to learn more over time about their pool of applicants and what factors influence the applicants' decision not to submit, the admission departments only observe college performance for those applicants who enroll and many of these observations occur several years after the admissions decisions are made.

arguments made by colleges for adopting these policies, which are often to increase the number of underrepresented students, these estimates suggest that the ability to infer SAT scores is so high, the admissions' margin is not where the colleges will accomplish their stated goal of increased diversity. The pool of applicants may be another margin over which the optional SAT policy could affect the student body, but our data from only the period when the colleges have the optional SAT policy in place do not allow us to address this.

If players do systematically underestimate the relationship between other players' actions and types, it has implications for all games with private information including the common value auctions, bilateral trade, voting and signaling games discussed in Eyster and Rabin (2005). Assuming a similar level of underestimation as in the optional SATI environment, we would expect "cursedness" to have a much more significant economic impact in these other situations because players in other games are unlikely to have access to the enormous amount of information colleges have on applicants. With less public information, misconstruing other players' private information will likely result in larger welfare affects. This is what Brown, Camerer and Lovalla (2012, 2013) found in the movie industry, where much more private information exists, where film studios can choose not to release the movie to critics prior to it opening. These papers also find that Bayesian updating does not fit the movie data as well as an alternative model, in their case, a cognitive hierarchy model.[31] By estimating "cursedness" in an environment where college applicants choose whether to submit their SATI scores, we provide direction for further empirical analyses focusing on whether and to what extent players underestimate the relationship between other players' actions and types.

[31] The literature on saliency and shrouding is related to how players make inferences. Gabiax and Laibson (2006) develop an interesting model where firms in competitive markets have incentive to hide information if a subset of consumers is "myopic". In the context of groceries and tolls, Chetty, Looney, & Kroft (2009) and Finkelstein (2009) find that demand for groceries and tolls are significantly influenced by the saliency of information. Using data from both field and natural experiments, Brown, Hossain and Morgan (2010) find that sellers in on-line auctions can increase profits by shrouding their shipping charges.

References:

Avery, Christopher, Andrew Fairbanks, and Richard Zeckhauser. 2003. *The Early Admissions Game*. Cambridge: Harvard University Press.

Banks, Jeffery S. and Joel Sobel. 1987. “Equilibrium Selection in Signaling Games.” *Econometrica*, 55(3): 647-661.

Brownstein, Andrew. 2001. “Colleges Debate Whether Dropping the SAT Makes Them More Competitive.” The Chronicle of Higher Education, 48(9): A14.

Brown, Jennifer, Tanjim Hossain and John Morgan. 2010. “Shrouded Attributes and Information Suppression: Evidence from the Field.” *The Quarterly Journal of Economics*, 125(2): 859-876.

Brown, Alexander L., Colin F. Camerer and Dan Lovallo. 2012. “To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office.” *American Economic Journal: Microeconomics*, 4(2): 1-26.

Brown, Alexander L., Colin F. Camerer and Dan Lovallo. 2013. “Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews.” *Management Science*, 59: 733-747.

Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong. 2004. “A Cognitive Hierarchy Model of Games.” *The Quarterly Journal of Economics*, 119(3): 861-898.

Cho, In-Koo and David Kreps. 1987. “Signaling Games and Stable Equilibria.” *Quarterly Journal of Economics*, 102: 179-221.

Chetty, Raj, Adam Looney and Kory Kroft. 2009. “Salience and Taxation: Theory and Evidence.” *American Economic Review*, 99(4): 1145-1177.

College Board. 2009. “Score Choice: New SAT Score-Reporting Policy.” http://www.collegeboard.com/prod_downloads/sat/sat-score-reporting-policy-flyer.pdf (accessed 4/3/17).

College Board. Unknown date. “SAT Program.” http://www.collegeboard.com/prod_downloads/highered/ra/SAT_Highlights.pdf (accessed 4/3/17).

Conlin, Michael, Stacy Dickert-Conlin and Gabrielle Chapman. 2013. “Voluntary Disclosure and the Strategic Behavior of Colleges.” *Journal of Economic Behavior and Organizations*, 96: 48-64.

Dekel, Eddie, Drew Fudenberg, and David K. Levine. 2004. “Learning to play Bayesian games.” *Games and Economic Behavior*, 46: 282–303 .

Ehrenberg, Ronald G. 2005. “Method or Madness? Inside the UNSWR College Rankings.” *Journal of College Admissions.* 189 (Fall): 29-35.

Eyster, Erik and Matthew Rabin. 2005. “Cursed Equilibrium.” *Econometrica*, 73(5): 1623-72.

Farrell, Elizabeth F. and Martin Van Der Werf. “Playing the Ranking Game.” Chronicle of Higher Education, May 25, 2007. http://chronicle.com/free/v53/i38/38a01101.htm.

Finkelstein, Amy. 2009. “E-ZTax: Tax Salience and Tax Rates,” *Quarterly Journal of Economics*, 124(3): 969-1010.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. The MIT Press.

Gabaix, Xavier and David Laibson. 2006. “Shrouding Attributes, Consumer Myopia, and Infromation Suppression in Competitive Markets.” *Quarterly Journal of Economics*, 121: 505-540.

Grossman, Sanford J. and Hart, Oliver D.. 1980. “Disclosure Laws and Takeover Bids.” *The Journal of Finance* 35: 323-334.

Grossman, Sanford J. 1981. “The Informational Role of Warranties and Private Disclosure about Product Quality.” *Journal of Law and Economics* 24: 461-89.

Harsanyi, John 1967. “Games with Incomplete Information Played By ‘Bayesian’ Players.” Management Science, 14(3): 159-182.

Jehiel, Philippe. 2005. “Analogy-based expectation equilibrium.” *Journal of Economic Theory*, 123: 81–104.

Jehiel, Philippe and Frédéric Koessler. 2008. “Revisiting games of incomplete information with analogy-based expectations.” *Games and Economic Behavior*, 62: 533-557.

Jin, Ginger. 2005. “Competition and Disclosure Incentives: Empirical Incentive from HMOs.” *Rand Journal of Economics*, 36: 93-112.

Jin, G. and P. Leslie. 2003. “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *Quarterly Journal of Economics*, 118(2): 409-51.

Kreps, David and Robert Wilson. 1982. “Sequential Equilibria.” *Econometrica*, 50(4): 863-894.

Levin, Daniel E. 2002. “The Uses and Abuses of the *U.S. News* Rankings.” *Priorities*. Association of Governing Boards of Universities and Colleges. 20: 1-19.

Major, B., Spencer, S., Schmader, T., Wolfe, C., and Crocker, J.. 1998. Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, 24: 34-50.

Mathios, Alan. 2000. “The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market.” *Journal of Law and Economics*. XLIII. 651-677.

Milgrom, Paul R. 1981. “Good News and Bad News: Representation Theorems and Applications.” *Bell Journal of Economics* 12: 380-91.

Monks, James and Ronald G. Ehrenberg. 1999. “The Impact of the U.S. News & World Report College Rankings on Admissions Outcomes and Pricing Policies at Selective Private Institutions.” NBER Working Paper 7277, July.

Robinson, Michael and James Monks. 2005. “Making SAT Scores Optional in Selective College Admissions: A Case Study.” *Economics of Education Review*, 24(4): 393-405.

Rooney, C., & Schaeffer, B.. 1998. “Test scores do not equal merit: Enhancing equity & excellence in college admissions by de-emphasizing SAT and ACT scores.” Cambridge, MA: National Center for Fair & Open Testing (FairTest).

Ross, Lee. 1977. “The intuitive psychologist and his shortcomings: Distortions in the attribution process.” In L. Berkowitz (Ed.), *Advances in experimental social psychology* (10, 173–220). New York: Academic Press.

Rothstein, Jesse. 2004. “College performance predictions and the SAT.” *Journal of Econometrics*, 121(1-2): 297-317.

Shavell, Steven. 1994. “Acquisition and Disclosure of Information Prior to Sale.” *RAND Journal of Economics* 25: 20-36.

Vigdor J.L. and C.T. Clotfelter. 2003. "Retaking the SAT." *Journal of Human Resources* 38(1): 1-33.

Table 1
Summary Statistics: Means and Standard Deviations

	COLLEGE X			COLLEGE Y		
	All Applicants	Submit	Don't Submit	All Applicants	Submit	Don't Submit
N	6,557	5,540	1,017	3,602	2,734	868
Attended Private HS	0.481 (0.500)	0.477 (0.500)	0.503 (0.500)	0.383 (0.486)	0.364 (0.481)	0.440 (0.497)
Income Missing (sr)	0.457 (0.498)	0.456 (0.498)	0.464 (0.499)	0.562 (0.496)	0.563 (0.496)	0.560 (0.497)
Income >100K (sr)	0.266 (0.442)	0.271 (0.445)	0.238 (0.426)	0.206 (0.404)	0.204 (0.403)	0.212 (0.409)
White	0.886 (0.318)	0.888 (0.315)	0.874 (0.332)	0.869 (0.337)	0.874 (0.332)	0.854 (0.354)
Not from Northeast	0.247 (0.432)	0.240 (0.427)	0.289 (0.454)	0.166 (0.372)	0.164 (0.371)	0.173 (0.378)
Female Student	0.676 (0.468)	0.657 (0.475)	0.778 (0.416)	0.503 (0.500)	0.487 (0.500)	0.553 (0.497)
SATI Score (math+verbal) (/1600)	1251 (131)	1272 (124)	1139 (112)	1257 (139)	1266 (144)	1228 (120)
No High School GPA reported (sr)	0.255 (0.436)	0.258 (0.438)	0.237 (0.425)	0.342 (0.474)	0.345 (0.475)	0.332 (0.471)
HS GPA A or A+ (sr)	0.199 (0.400)	0.197 (0.398)	0.209 (0.407)	0.226 (0.418)	0.239 (0.426)	0.184 (0.388)
HS GPA A- (sr)	0.227 (0.419)	0.227 (0.420)	0.228 (0.420)	0.172 (0.377)	0.166 (0.372)	0.190 (0.393)
Class rank missing	0.308 (0.462)	0.304 (0.460)	0.332 (0.471)	0.354 (0.478)	0.346 (0.476)	0.379 (0.485)
Class rank 1st 10th	0.220 (0.414)	0.225 (0.418)	0.195 (0.396)	0.205 (0.403)	0.212 (0.409)	0.182 (0.386)
SATII Score(s) Submitted (1=yes)	0.731 (0.444)	0.751 (0.432)	0.815 (0.388)	0.382 (0.486)	0.338 (0.473)	0.522 (0.500)
Average SATII Score (when submitted) (/800)	633 (68)	638 (68)	595 (67)	637 (69)	640 (72)	632 (60)
ACT Score Submitted (1=yes)	0.17 (0.68)	0.13 (0.34)	0.34 (0.47)	0.03 (0.18)	0 (0)	0.14 (0.34)
Average ACT Score (when submitted) (/36)	26.30 (3.47)	26.59 (3.53)	25.70 (3.26)	26.14 (3.28)		26.14 (3.28)
Applied Early Decision	0.068 (0.252)	0.058 (0.235)	0.120 (0.325)	0.105 (0.306)	0.108 (0.310)	0.096 (0.294)

Table 1 (continued)
Summary Statistics: Means and Standard Deviations

	COLLEGE X			COLLEGE Y		
	All Applicants	Submit	Don't Submit	All Applicants	Submit	Don't Submit
N	6,557	5,540	1,017	3,602	2,734	868
Overall Rating (/7)	3.290 (0.984)	3.310 (0.996)	3.178 (0.909)			
Personal/Non-Academic Rating (/2)				1.031 (0.298)	1.027 (0.328)	1.073 (0.379)
Academic Rating (/7)	3.253 (1.085)	3.289 (1.085)	3.058 (0.964)	4.570 (1.543)	4.545 (1.611)	4.649 (1.295)
Accepted (1=yes)	0.415 (0.493)	0.419 (0.493)	0.396 (0.489)	0.457 (0.498)	0.446 (0.497)	0.492 (0.500)
Enroll (1=yes)	0.135 (0.341)	0.125 (0.341)	0.187 (0.390)	0.139 (0.346)	0.132 (0.339)	0.159 (0.366)
Enroll Conditional on Acceptance	0.324 (0.468)	0.298 (0.458)	0.471 (0.500)	0.303 (0.460)	0.296 (0.457)	0.323 (0.468)
SATI Score Conditional on Acceptance (/1600)	1301 (118)	1323 (107)	1172 (099)	1321 (117)	1343 (114)	1260 (103)
SATI Score Conditional on Enrollment (/1600)	1254 (117)	1281 (107)	1156 (101)	1280 (114)	1301 (113)	1225 (98)

Table 2
Estimates of Structural Parameters for College X and College Y

Parameter Estimates	College X			College Y		
χ , weight on expected SATI score not conditioning on submission choice	0.191* (0.095)			0.339** (0.120)		
$\beta^{P'}_{SAT}+\beta^{R'}_{SAT}$, coefficient associated with actual SATI score/100 for applicants who submit	0.619** (0.054)			0.592** (0.049)		
$\beta^{P'}_{SAT}$, coefficient associated with inferred SATI score/100 for applicants who do not submit	0.652** (0.057)			0.629** (0.053)		
$B^{YR'}$, coefficient associated with inverse of expected probability of enrollment	-0.684** (0.052)			-0.282** (0.033)		
σ_u , standard deviation associated with inference of unconditional SATI score	0.084** (0.001)			0.117** (0.002)		
σ_c , standard deviation associated with conditional SATI score	0.084** (0.038)			0.145** (0.051)		
	Acceptance Decision β's	Predicted SATI α's	Expected Enrollment β^e's	Acceptance Decision β's	Predicted SATI α's	Expected Enrollment β^e's
SATII Score(s) Submitted (1=yes)	-3.323** (0.675)	-0.737** (0.013)	0.210 (0.291)	-2.227* (0.923)	-0.741** (0.035)	0.337 (0.753)
Average SATII Score/100 (when submitted)	5.967** (1.072)	1.225** (0.020)	-0.526 (0.439)	3.304* (1.429)	1.234** (0.054)	-0.404 (1.146)
ACT Score Submitted (1=yes)	-1.201 (1.033)	-0.657** (0.024)	-0.139 (0.447)	2.041 (1.327)		
Average ACT Score (when submitted)	0.064 (0.038)	0.024** (0.001)	-0.001 (0.016)	-0.066 (0.050)		
Attended Private HS	0.436** (0.115)	0.001 (0.002)	-0.134** (0.047)	0.411** (0.112)	0.013** (0.005)	-0.189* (0.090)
Female Student	-0.598** (0.107)	-0.021** (0.002)	0.030 (0.044)	0.360** (0.106)	-0.031** (0.005)	-0.012 (0.080)
No High School GPA reported (sr)	0.771** (0.176)	0.018** (0.004)	-0.133 (0.080)	0.655** (0.220)	0.063** (0.009)	-0.109 (0.216)
HS GPA A or A+ (sr)	1.310** (0.218)	0.026** (0.005)	-0.176 (0.091)	1.432** (0.285)	0.078** (0.010)	-0.422 (0.244)
HS GPA A- (sr)	1.036** (0.185)	0.026** (0.004)	-0.159* (0.080)	0.891** (0.245)	0.058** (0.009)	-0.184 (0.228)
HS GPA B+ (sr)	0.702** (0.175)	0.016** (0.004)	-0.109 (0.078)	0.340 (0.229)	0.042** (0.009)	0.190 (0.234)

Table 2 (cont.)
Estimates of Structural Parameters for College X and College Y

	College X			College Y		
	Acceptance Decision β's	Predicted SATI α's	Expected Enrollment β^{e}'s	Acceptance Decision β's	Predicted SATI α's	Expected Enrollment β^{e}'s
Class rank missing	0.191	0.010**	-0.029	-0.119	0.026**	0.066
	(0.134)	(0.003)	(0.059)	(0.143)	(0.006)	(0.123)
Class rank 1st 10th	0.922**	0.023**	-0.206*	0.380	0.069**	0.062
	(0.207)	(0.004)	(0.082)	(0.248)	(0.009)	(0.176)
Class rank 2nd 10th	0.349**	0.009*	-0.052	-0.016	0.024**	0.006
	(0.161)	(0.004)	(0.069)	(0.201)	(0.009)	(0.154)
Income Missing (sr)	-0.394**	-0.006	0.215**	0.087	-0.007	-0.012
	(0.147)	(0.003)	(0.057)	(0.157)	(0.007)	(0.130)
Income <100K (sr)	-0.527**	-0.013**	0.276**	0.179	-0.043**	-0.328**
	(0.160)	(0.003)	(0.066)	(0.180)	(0.007)	(0.125)
Legacy (1=yes)	0.745*	0.006	0.016	0.053	-0.005	0.274
	(0.310)	(0.007)	(0.130)	(0.192)	(0.009)	(0.183)
White	-2.130**	0.050**	0.548**	-0.832**	0.081**	-0.656**
	(0.167)	(0.004)	(0.052)	(0.161)	(0.007)	(0.171)
Not from Northeast	0.197	0.001	-0.082	0.012	0.006	0.120
	(0.134)	(0.003)	(0.052)	(0.138)	(0.006)	(0.113)
Apply Early Decision		-0.013**	4.848**		-0.042**	4.067**
		(0.005)	(0.341)		(0.007)	(0.262)
SATI score/100			-0.157**			0.083*
			(0.022)			(0.040)
SATI score Submitted			0.053			0.099
			(0.079)			(0.107)
Constant	-4.540**	1.189**	0.494	-6.758**	1.124**	-1.855**
	(0.663)	(0.006)	(0.293)	(0.588)	(0.011)	(0.478)
Log Likelihood		1,280			249	
Observations		6,557			3,602	

Standard errors in parentheses: * significant at 5% and ** significant at 1%.

Table 3
Estimates of Selective Structural Parameters for College X and College Y
With Academic Rating as the College's Choice

Parameter Estimates	College X	College Y
χ, weight on expected SATI score not conditioning on submission choice	0.358** (0.104)	0.286** (0.039)
$\beta^{P'}_{SAT}+\beta^{R'}_{SAT}$, coefficient associated with actual SATI score/100 for applicants who submit	0.266** (0.013)	0.901** (0.013)
$\beta^{P'}_{SAT}$, coefficient associated with inferred SATI score/100 for applicants who do not submit	0.254** (0.012)	0.924** (0.014)
Log Likelihood	-1,701	-2,199
Observations	6,510	3,481

Notes: Set of covariates is the same as in Table 2.
Standard errors in parentheses: * significant at 5% and ** significant at 1%.

Table 4
Estimates of Selective Structural Parameters for College X and College Y Including Non-Academic Ratings as Covariates

Parameter Estimates	College X	College Y
χ, weight on expected SATI score not conditioning on submission choice	0.186*	0.380**
	(0.098)	(0.130)
$\beta^{P'}{}_{SAT}+\beta^{R'}{}_{SAT}$, coefficient associated with actual SATI score/100 for applicants who submit	0.613**	0.613**
	(0.054)	(0.052)
$\beta^{P'}{}_{SAT}$, coefficient associated with inferred SATI score/100 for applicants who do not submit	0.646**	0.655**
	(0.057)	(0.056)
Acceptance Decision:		
Non-Academic Rating	0.207**	0.854**
	(0.030)	(0.098)
Predicted SATI:		
Non-Academic Ranking	-0.004*	0.006
	(0.002)	(0.007)
Log Likelihood	1,300	-356
Observations	6,507	3,478

Notes: Set of covariates is the same as in Table 2 except include non-academic rating. Non-Academic Rating is Overall Rating minus Academic Rating for College X and Personal Rating for College Y. Standard errors in parentheses: * significant at 5% and ** significant at 1%.

Table 5
College Performance Measures

College X:	GPA at graduation	Graduated	Honor or Distinction in Major	Phi Beta Kappa	Summa or Magna Cum Laude	Number of Sports
χ , weight on expected SATI score not conditioning on submission choice	36,911 (403,187)	7.98 (126)	-0.834 (1.201)	-0.074 (1.213)	0.570 (0.598)	-0.521 (1.449)
$\beta^{P'}_{SAT}+\beta^{R'}_{SAT}$, coefficient associated with actual SATI score for applicants who submit	-0.003 (0.002)	-0.018 (0.027)	0.124 (0.073)	0.182 (0.116)	0.262** (0.092)	0.078 (0.078)
$\beta^{P'}_{SAT}$, coefficient associated with inferred SATI score for applicants who do not submit	-0.00000 (0.00001)	-0.002 (0.027)	0.151 (0.075)	0.201 (0.120)	0.274** (0.095)	0.121 (0.079)
Log Likelihood	482	4,146	5,476	5,697	5,570	5,466
Mean (SD) of Y	3.450 (0.289)	0.823 (0.382)	0.617 (0.486)	0.106 (0.308)	0.245 (0.430)	0.406 (0.634)
Observations with Performance Measure	687	882	687	687	687	687

College Y:	Freshman Year GPA	Completed Freshman Year	Number of Sports
χ , weight on expected SATI score not conditioning on submission choice	-0.365 (1.177)	-3.15 (6.90)	0.878 (0.421)
$\beta^{P'}_{SAT}+\beta^{R'}_{SAT}$, coefficient associated with actual SATI score for applicants who submit	1.098 (0.961)	0.055 (0.142)	-0.123** (0.032)
$\beta^{P'}_{SAT}$, coefficient associated with inferred SATI score for applicants who do not submit	1.358 (1.013)	0.093 (0.149)	-0.126** (0.034)
Log Likelihood	-1,075	1,919	1,576
Mean (SD) of Y	3.450 (0.289)	0.960 (0.196)	0.343 (0.582)
Observations with Performance Measure	479	499	499

Notes: Set of covariates is the same as in Table 2. Standard errors in parentheses: * significant at 5% and ** significant at 1%.

Table 6
Robustness Results

	College X		College Y	
	Parameterization: Inference based on all applicants	Endogenize Enrollment	Parameterization: Inference based on all applicants	Endogenize Enrollment
χ , weight on expected SATI score not conditioning on submission choice	0.181 (0.096)	0.262 (0.182)	0.361** (0.122)	0.349** (0.088)
$\beta^{P'}{}_{SAT}+\beta^{R'}{}_{SAT}$, coefficient associated with actual SATI score/100 for applicants who submit	0.618** (0.054)	0.274** (0.025)	0.592** (0.049)	0.658** (0.033)
$\beta^{P'}{}_{SAT}$, coefficient associated with inferred SATI score/100 for applicants who do not submit	0.652** (0.058)	0.300** (0.025)	0.630** (0.053)	0.688** (0.033)
Log Likelihood	1,978	-3,544	229	-3,851
Observations	6,557	6,557	3,602	3,602

For alternative parameterization of college's inference, the set of covariates is the same as in Table 2. Standard errors in parentheses: * significant at 5% and ** significant at 1%.

Appendix A

Substituting the functional form assumptions in [2] into the college's acceptance decision in [1] results in the college accepting applicant *i* if the following inequality is satisfied.

$$P_i^e\left[\beta_{SAT}^P(SAT_{+i}^P - SAT_{-i}^P) + \boldsymbol{\beta^P}\cdot\left(\boldsymbol{Z_{+i}^P} - \boldsymbol{Z_{-i}^P}\right) + \varepsilon_{qi} + \beta_{SAT}^R(SAT_{+i}^R - SAT_{-i}^R) + \boldsymbol{\beta^R}\cdot\left(\boldsymbol{Z_{+i}^R} - \boldsymbol{Z_{-i}^R}\right) + \boldsymbol{\beta^D}\cdot\left(\boldsymbol{Z_{+i}^D} - \boldsymbol{Z_{-i}^D}\right)\right] + \beta^{YR}(YR_{ai} - YR_{ri}) > 0$$

We can express changes in the expected average or fraction of characteristic *j* of the student body as $Z_{+i}^j - Z_{-i}^j = \frac{\left(N^{sb}Z_{-i}^j + Z_i^j\right)}{(N^{sb}+1)} - Z_{-i}^j = \frac{Z_i^j}{(N^{sb}+1)} - \frac{Z_{-i}^j}{(N^{sb}+1)}$ where N^{sb} is expected student body size excluding applicant *i*. Similarly, $SAT_{+i}^P - SAT_{-i}^P = \frac{SAT_i^P}{(N^{sb}+1)} - \frac{SAT_{-i}^P}{(N^{sb}+1)}$ and $SAT_{+i}^R - SAT_{-i}^R = \frac{SAT_i^R}{(N^{sub}+1)} - \frac{SAT_{-i}^R}{(N^{sub}+1)}$ where N^{sub} is the expected number of enrolled students that submit excluding applicant *i*. Likewise, the change in expected yield rate from applicant *i* being accepted can be expressed as $YR_{ai} - YR_{ri} = \frac{P_i^e}{(N^a+1)} - \frac{YR_{ri}}{(N^a+1)}$ where N^a is the expected number of other applicants accepted. Assuming the college's expectations of $Z_{-i}^j (j \in P, R, D)$, SAT_{-i}^k $(k \in P, R)$, YR_{ri}, N^{sb}, N^{sub} and N^a do not vary across applicants, we can denote the probability applicant *i* is accepted as

$$P_s^a = Prob\left\{\varepsilon_{qi} > \beta^{int} - \beta_{SAT}^{P'}SAT_i - \boldsymbol{\beta^{P'}}\cdot\boldsymbol{Z_i^P} - \beta_{SAT}^{R'}SAT_i - \boldsymbol{\beta^{R'}}\cdot\boldsymbol{Z_i^R} - \boldsymbol{\beta^{D'}}\cdot\boldsymbol{Z_i^D} + \beta^{YR'}/P_i^e\right\} \quad [3]$$

where:

$$\beta_{SAT}^{P'} = \frac{\beta_{SAT}^P}{(N^{sb}+1)},\ \boldsymbol{\beta^{P'}} = \frac{\boldsymbol{\beta^P}}{(N^{sb}+1)},\ \beta_{SAT}^{R'} = \frac{\beta_{SAT}^R}{(N^{sb}+1)},\ \boldsymbol{\beta^{R'}} = \frac{\boldsymbol{\beta^R}}{(N^{sub}+1)},\ \boldsymbol{\beta^{D'}} = \frac{\boldsymbol{\beta^D}}{(N^{sb}+1)},\ \beta^{YR'} = \frac{\beta^{YR}}{(N^a+1)}$$

and β^{int} is a function of Z_{-i}^j (for all *j*), N^{sb}, N^{sub} and N^a.