Optimal Dynamic Selection Under Costly Evaluation: Evidence from Graduate Admissions

Jeremy Rosen, *Georgetown University**

April 10, 2026

Abstract

This paper uses graduate admissions to study dynamic selection under costly evaluation. With a synthetic data protocol to protect applicants' privacy, I relate acceptance to measures of success, such as job placements. I then estimate structural models of an admissions committee's decision problem that capture first-round filtering and waitlists, and I develop statistical tests of whether observed decision rules are optimal given time and effort constraints. To determine which if any applicant characteristics are incorrectly valued, I quantify deviation gains from adopting counterfactual decision rules. This framework is applicable to multiple settings where decision-makers face uncertain outcomes and evaluation costs.

Keywords: Costly Evaluation, Dynamic Optimization, Education Economics, Graduate Admissions, Optimal Selection, Structural Estimation, Synthetic Data

JEL Codes: C61, C81, D81, I23

*Department of Economics, Georgetown University, Email: `jar361@georgetown.edu`

I thank James Albrecht, Satyam Anand, Axel Anderson, Susan Athey, Jessica Bai, Cauê Dobbin, Xavier Gabaix, Ian Gale, W. Bentley MacLeod, Nathan Miller, Alexandre Olbrecht, John Rust, and Susan Vroman for their helpful feedback. This research was approved under Georgetown University IRB protocol STUDY00006697.

1 Introduction

How do decision-makers select among a set of options in which each option's true quality is unobserved and costly to evaluate in terms of time and effort? In many such problems, decision-makers rely on a coarse, low-cost screening of all the options, followed by a precise, high-cost evaluation of the subset of options that survive the screening. Moreover, if the selected options have agency over whether to match with the decision-maker, final selections are often staggered over time through a deferred process to ensure the number of matches meets a target number.

Admissions and hiring are two prominent examples of such selection problems. There, committees first screen all the applications to filter out unqualified applicants, then read a subset of promising applications in full detail, then potentially interview the most promising applicants, and finally make early and late offers, depending on how many early offers are accepted. Other examples of selection with costly evaluation, and in some cases deferred offers, include conducting multi-stage clinical trials for medicines, messaging virtually on dating apps before meeting in real life, and simulating product prototypes on a computer before physically building them.

A fundamental question in such settings is how decision-makers should interpret the available information at each stage of evaluation to select either the best subset of options or all options above a quality threshold. A related question is whether observed decision rules reflect optimal dynamic selection given evaluation costs and other constraints, or whether those rules systematically deviate from optimal behavior. One could imagine a hiring committee undervaluing a skill or discriminating against a demographic. Moreover, if there are deviations, which decision rules should decision-makers implement instead? Answering these questions requires moving beyond identification of the determinants of success and explicitly modeling the decision problem.

Graduate admissions is a meaningful setting to study these questions, as evidenced by economics papers from Ehrenberg and Mavros (1995) to Bai, Esche, MacLeod, and Shi (2022). Graduate programs, especially PhD programs, invest at least a few years of time, along with tens of thousands of dollars of funding, into students' development. Moreover, high-ranking academic placements boost a program's reputation, and so increase its ranking from sources such as U.S. News. Better students also lead to better coauthoring opportunities for faculty, and since people with graduate degrees often find themselves in influential positions, such as professorships, studying how best to admit them is important for programs and society. In particular, if schools are undervaluing applicants with certain characteristics (e.g. from a particular demographic), future applicants with such characteristics may receive opportunities they would not have otherwise. For example, while Chetty, Deming, and Friedman (2026) focus on undergraduate admissions, they find that applicants from high-income families have higher admissions rates at Ivy-Plus colleges even if their standardized test scores are no higher than those of less fortunate applicants.

Moreover, graduate admissions is a setting with costly evaluation. Admissions committees observe for all applicants a limited set of objective characteristics, such as standardized test scores, but typically lack the time to read every file in detail. As such, applicants are initially filtered based on coarse signals, with more intensive evaluation reserved for the subset of applicants with promising objective characteristics. Final decisions may also involve waitlists or other forms of deferred acceptance, as uncertainty about applicant quality and matriculation status resolves over the period between the application and admissions deadlines. These features make graduate admissions well-suited for modeling dynamic selection under costly evaluation.

This paper addresses three specific research questions. The first question asks which *ex-ante* factors in application files determine admission and success in graduate school. Following the literature, success may be defined in multiple ways, including graduate course grades, passing comprehensive examinations, program completion, and job placement outcomes. I study this question as comprehensively as possible by extracting all available data from application files, including recommendation letters, application essays, and CVs. My data consist of all applicants to a research university's economics PhD program from 2015 to 2023, as well as all applicants to that university's government, history, linguistics, and sociology PhD programs from 2020 to 2023. Because these data are confidential, I implement a privacy-protecting research protocol that uses synthetic data for model development and validation (see Section 3). This protocol also serves as a template for researchers who wish to work with sensitive data in general.

The second question asks whether admissions committees optimize or make mistakes when admitting applicants. Specifically, if a factor predicts admission but not success, the committee may be overvaluing it, whereas if a factor predicts success only, the committee may be undervaluing it. For instance, the committee may be putting too much or too little weight on an elite undergraduate institution. To determine if the committee is optimizing, it is necessary to model the committee's optimization problem, estimate the parameters, and formally test the hypothesis of optimization, which to date, no other paper has done. To make such models tractable, per Bai et al. (2022), I reduce the committee's portfolio choice problem (i.e. from N applicants, choose the optimal subset N_A) to a cutoff rule problem (rank the N applicants by their expected success probabilities and select the N_A above the cutoff). I estimate multiple structural models in increasing order of complexity, modeling such features as first-round filtering and waitlists. Predictors that are strongly associated with admission but weakly associated with success may be overvalued, while predictors of success that play little role in admissions decisions may be undervalued. For example, committees may place too much or too little weight on elite undergraduate institutions, standardized test scores, or technical skills. Determining whether such patterns reflect optimal trade-offs or mistakes requires modeling the committee's optimization problem directly.

Perhaps as Simon (1956) suggests, if committees are not optimizing, they may be satisficing (i.e. aiming for a satisfactory result), namely because reading every application file would require too much effort.[1] This idea lends itself to the third question of whether committees can make better decisions without having to expend additional effort. To answer this question, I use

[1]Satisficing is distinct from rational inattention (see Section 2.2), which assumes that decision-makers are optimizing and so embeds the choice of how much effort to allocate into an optimization framework.

the structural models to calculate counterfactual deviation gains from adopting the optimal admissions strategy, while respecting the committee's time constraint of not being able to read every application file in detail. For robustness, I perturb the success parameter estimates about their asymptotic distribution when comparing the expected success probability for the class admitted by the model's decision rule vs. the committee's. Perhaps committees have other priorities than maximizing success probabilities, such as admitting a diverse class, but I show how the models can be extended to factor in those priorities. Using Monte Carlo simulated data whose data-generating parameters I calibrate with reduced-form results from actual data, I show that in a dataset where the committee behaves optimally, a Kolmogorov-Smirnov test fails to reject the null hypothesis of no deviation gains. But in a dataset where the committee makes mistakes, such as discriminating against a demographic, the test rejects the null with $p < .0001$.

Given that the university in this study is not necessarily representative of all graduate programs, some of this paper's results may not generalize to other settings. However, the methods are generalizable to admissions, hiring, and other settings, essentially to any setting where decision-makers must decide which options to select based on *ex-ante* information. In addition, one may find evidence of discrimination against applicants with certain features (e.g. a low-income background, as shown in Chetty et al. 2026), or of academic networks, or "bubbles." For instance, certain undergraduate or economic institutions, such as liberal arts colleges, feed into economics graduate programs (see Siegfried and Stock 2007 and Stock and Siegfried 2015 for related findings). In addition, I find that having a recommender who also wrote letters for other applicants is predictive of admission, which implies that it helps to have a well-connected recommender. Evidence of discrimination or bubbles contributes to the literatures on education economics (e.g. improving educational outcomes, see Ferreyra, Garriga, Martín-Ocampo, and Sánchez-Díaz 2023), matching (applicants to programs), and optimization under uncertainty.

The rest of this paper proceeds as follows. In Section 2, I review the economics literature on both graduate admissions and this paper's methodological contributions, and in Section 3, I describe my data and the synthetic data protocol. Then, in Sections 4 and 5, I present reduced-form and structural models and results. Specifically, in Section 4, I regress admission on its determinants for all five of the university's social science PhD programs, including economics, from 2020 to 2023. I find that economics admissions is more data-driven than the other programs (i.e. the observable factors explain admission better for economics), though more data-driven might not mean closer to optimal. I also implement a Wald test of whether the determinants of admission and success are the same or different across programs. Given that the synthetic data protocol is in progress, I run this test on reduced-form-calibrated simulated data, specifically on one dataset where the determinants are the same by construction, and one where they are different. As should happen, the test rejects the null on the first dataset but not the second.

After that, in Section 5, I develop four structural models of admissions: year-static, year-dynamic, waitlist-hybrid, and waitlist-dynamic. In all four models, there is a filtering first stage, in which the committee only observes the objective factors, such as GRE scores but not recommendation letters. I estimate the models on simulated datasets where the committee is optimizing and where it is not, and the estimation results are all well-behaved. That is, the structural hypothesis tests reject optimization, and I find counterfactual deviation gains on only the latter datasets. I also find that if the committee reads too few files in Round 1, then the marginal applicant rejected in Round 1 has a higher expected success probability than the marginal applicant accepted into the program. When I progress from simulated to actual data, I will calculate how many files "too few" is. Finally, in Section 7, I summarize the paper's results and methodological contributions, as well as how a data-driven approach to optimal selection can help decision-makers in both admissions and a variety of other settings who face costly evaluation.

2 Literature Review

I divide this literature review into two groups. The first group consists of papers on the determinants of admission and success in graduate education. The first subgroup within this group, the original literature, begins with Ehrenberg and Mavros (1995) and ends with Athey, Katz, Krueger, Levitt, and Poterba (2007). In these papers, the authors regress measures of admission and success on *ex-ante* predictors found in applications to top economics PhD programs. The second subgroup is a series of papers by Stock, Finegan, and Siegfried (SFS) from 2006 to 2015. These papers regress measures of success on both *ex-ante* predictors and graduate program characteristics for 586 students who matriculated at 27 economics PhD programs in Fall 2002. Finally, the third subgroup, the recent literature, begins with Grove, Dutkowsky, and Grodner (2007) and ends with Bai et al. (2022). These papers distinguish themselves from the original literature by studying more diverse samples and by incorporating new predictors and methods.

The second group of papers relate to this paper's methodological contributions. The first subgroup of this group focuses on optimal dynamic selection with costly evaluation, and it contains papers on such topics as exploration vs. exploitation and rational inattention. Meanwhile, the second subgroup features papers on synthetic data and privacy protection, from the foundations of synthetic data to its economic applications. Because there have been few published economics studies that use synthetic data, this paper may guide future researchers who wish to implement a synthetic data protocol to allow them to work with confidential data they cannot directly access.

2.1 Graduate Admissions Literature

While there have been earlier papers on graduate education, Ehrenberg and Mavros (1995) were the first to include predictors of success found in application files. They analyze Cornell PhD students, including economics, and find that verbal but not quantitative GRE scores are predictive of program completion. Subsequent research implies the reason is likely that their sample is restricted to only the applicants who attended Cornell, nearly all of whose quantitative GRE scores were concentrated in the top percentiles. Two years later, Attiyeh and Attiyeh (1997) use a sample across 48 universities and multiple programs from 1990 to 1991 to study determinants of admission. They find that all else equal, committees prioritized underrepresented minorities.

Krueger and Wu (2000) and Grove and Wu (2007) then study both admission and success at a single top economics department in 1989, taking care to include all applicants in their success regressions and track placement outcomes for those who did not matriculate. They find that admissions committee scores, GRE scores, and the prominence of recommendation letter writers predict good placements, but there is a lot of uncertainty. Finally, Athey et al. (2007) analyze the determinants of success for Chicago, Harvard, MIT, Princeton, and Stanford economics PhD students from 1990 to 1999, adding graduate course grades into their models. While GRE scores and other *ex-ante* factors predict good placements, they do not when controlling for course grades.

Stock et al. (2006) begin the series of SFS papers by compiling the aforementioned 586 student dataset of Fall 2002 economics PhD matriculants. They start by looking at attrition rates in the first two years and find that low GRE scores predict attrition, but RAships and shared department office space help programs retain students. Siegfried and Stock (2007) compile another, larger dataset of everyone who earned a U.S. economics PhD from 1966 to 2003 and observe that top liberal arts colleges feed economics PhD programs at disproportionally high rates.

Stock, Finegan, and Siegfried (2009a) and Stock, Finegan, and Siegfried (2009b) follow up on the progress of the students in Stock et al. (2006) and find similar results but emphasize that there is a lot of uncertainty in predicting completion after five years. Stock, Siegfried, and Finegan (2011) and Stock and Siegfried (2014) track the students after eight years and summarize their multi-paper study, concluding that while some factors like high GRE scores, a liberal arts undergraduate institution, and financial support predict completion, uncertainty remains very high. Finally, Stock and Siegfried (2015) add more recent years to Siegfried and Stock (2007)'s larger dataset, noting that foreign undergraduate institutions have become more common over time.

Although Grove et al. (2007) is from the same year as Athey et al. (2007), it employs newer methods. First, the data are from Syracuse economics PhD students, who have different characteristics than students attending top-ranked programs. The authors also incorporate factor analysis into their regressions, finding that different attributes predict success at different stages. For example, GRE scores predict passing comprehensive exams, whereas research motivation (a factor from their factor analysis) predicts program completion. Schlauch and Startz (2018) study economics PhD candidates from top 50 programs posting their CVs for the 2016–2017 job market and find that while RA experience predicts attending a highly-ranked PhD program, a masters degree does not. This result tracks with the increasing prevalence of predocs in economics. Meanwhile, Jones, Schuhmann, Soques, and Witman (2020) do not run regressions but rather survey individuals involved in admissions at 132 economics PhD programs. They find that higher-ranked programs care about different factors than lower-ranked programs. For example, higher programs care more about undergraduate rankings. Also, while recommendation letters matter for admission, programming skills usually do not, though they may matter for success.

Bai et al. (2022) represent the frontier of economics graduate admissions research. They study economics PhD applicants at a single top program from 2013 to 2019 and like Krueger and Wu (2000) collect data on admissions and placement outcomes for all applicants. Unlike all previous papers, they have a theory section, in which they prove that a committee's portfolio choice problem simplifies to a cutoff rule problem, which is more tractable for prospective modelers. I build on this result by directly modeling the cutoff rule problem and estimating the parameters, which is necessary to make claims as to whether committees are optimizing. In addition, they perform textual analysis of recommendation letters and find that subjective factors, such as letters, matter more for both admission and success more than objective factors, such as GREs. As such, the models in this paper also incorporate subjective as well as objective factors.

2.2 Methodological Literature

This paper also relates to a broader methodological literature on optimal dynamic selection with costly evaluation, as well as a literature on synthetic data and privacy protection in empirical economics. Within the former literature, earlier papers develop classic models of sequential search where agents face uncertainty about option quality and must decide when to stop searching given search costs. McCall (1970) and Mortensen (1970) model job search as a dynamic stopping problem in which workers optimally accept offers that exceed a reservation wage. Weitzman (1979) extends this logic to settings in which decision-makers can acquire information about multiple alternatives at a cost, which he denotes as "Pandora's Problem." Together, these papers establish that when information is costly, optimal decision rules are often threshold-based.

A more recent group of papers studies costly evaluation outside of search problems. In models of rational inattention, agents are fully optimizing but face constraints on their ability to process information. Sims (2003) introduces rational inattention as a framework in which information acquisition is costly, leading agents to optimally trade off decision accuracy with information costs. Matějka and McKay (2015) show that rational inattention can generate probabilistic discrete choice rules in line with multinomial logit models, while Caplin and Dean (2015) provide a revealed-preference characterization of optimal information acquisition that separates rational inattention from actual mistaken decisions. These models differ from Simon (1956)'s concept of satisficing, which proposes heuristic, "good enough" decision rules. In contrast, rational inattention models retain the mathematical framework of rationality but endogenize information acquisition. Notably, unlike this paper, none of these papers estimate their models on data.

This paper differs from both rational inattention and satisficing. I take the information environment faced by decision-makers as given, since admissions committees, at least in this framework, do not choose how much information is available at each stage of evaluation. Instead, they interpret the information they observe at each stage and decide how to act on it. The contribution of this paper is to model this dynamic selection problem explicitly, estimate parameters governing the resulting decision rules, and determine whether observed behavior is optimal conditional on respecting the committee's constraints. By doing so, this paper makes optimality testable and allows for the computation of counterfactual decision rules and deviation gains.

The second methodological literature addresses synthetic data and privacy protection. Rubin (1993) introduces synthetic data as a tool for protecting confidentiality by replacing sensitive microdata with simulated observations drawn from an estimated data-generating process. Before synthetic data, privacy protection often involved masking the actual data by, for example, adding random noise. Masking, however, induces a trade-off between privacy and information loss (Little 1993), and the rows in a masked dataset still correspond to real subjects. More recent work in economics emphasizes the challenges posed by restricted access to data. Abowd and Schmutte (2015) characterize ignorability, or when researchers can estimate their models without having to account for the effect of statistical disclosure limitations on their estimates. Meanwhile, Koenecke and Varian (2020) propose that researchers publicly release synthetic data, such as those generated by the Synthetic Data Vault (Patki, Wedge, & Veeramachaneni 2016), when they cannot release the actual data. Their proposed protocol differs from this paper's protocol, as this protocol is designed for a setting where the *researcher* cannot access the actual data.

Finally, Bertrand, Kamenica, and Pan (2015); Stanley and Totty (2024); and Carr, Wiemers, and Moffitt (2025) assess the efficacy of synthetic data by comparing results from the U.S. Census Bureau's Survey of Income and Program Participation dataset to those from its synthetic counterpart (Benedetto, Stanley, & Totty 2018). These papers develop protocols where researchers validate analyses on synthetic data, with final estimation conducted on the actual data by authorized intermediaries. This paper's protocol goes further because the synthetic datasets used here were not already available, and I use them for structural estimation as well as reduced-form. As such, the protocol spans the entire process from creating synthetic data to obtaining final results on actual data. Rather than drawing inference from synthetic data, synthetic datasets are used solely for model development and code validation, while all reported estimates are obtained from the actual administrative data. In this way, this paper provides a template for conducting both reduced-form and structural analyses on sensitive data when direct access is infeasible.

3 Data

This paper uses two datasets: one consisting of all applicants to a research university's economics PhD program from 2015 to 2023, and another of all applicants to that university's economics, government, history, linguistics, and sociology programs from 2020 to 2023. Unlike the latter dataset, the former dataset contains subjective variables (i.e. those derived from textual documents such as CVs and recommendation letters) and program outcomes data (e.g. post-program job placements). Therefore, I estimate this paper's structural models on the former dataset only. However, the latter dataset, which still contains objective variables and each applicant's admission status, is necessary for making reduced-form comparisons across programs.

Also, the implementation of this paper's privacy-protecting synthetic data protocol is in progress.

Therefore, while Section 4 contains reduced-form results from the latter dataset, the rest of the paper's results, particularly the structural results in Section 5, are from Monte Carlo simulated data whose data-generating parameters were calibrated with the available reduced-form results. The simulated data estimates verify that the structural models' parameters are identifiable and that the hypothesis tests based on those models reject optimality only when the simulated admissions committee misvalues certain characteristics, such as an arbitrary demographic. I will replace the simulated data results with actual results and move the simulated results to an appendix in a future version of this paper. In this section, I describe the variables and synthetic data protocol, the latter of which serves as a template for conducting research on confidential data.

3.1 Variables

There are three types of variables: dependent variables, objective independent variables, and subjective independent variables. Table 8 of Appendix A contains a full list of the variables. The dependent variables consist of measures of both admission and success. For admission, such measures include binary variables of whether the applicant makes it past the first-round filtering stage, whether they are accepted into the program, whether they are waitlisted, whether they attend the university's open house, whether they matriculate at the university, and whether they attend (and were therefore accepted at) a different university's program. I also have the U.S. News American (if applicable) and global rankings of the program they attended.

For success in economics, I have a binary variable of whether they complete the program, along with categorical and continuous variables of the type and ranking of their eventual job placement. Placement types include academic, government, consulting, and technology. For academic placements, I use the department's IDEAS/RePEc American and global rankings with a plus five penalty for business schools (Krueger & Wu 2000). For postdocs, I use a 2/3 to 1/3 weighted average of the university's ranking and the maximum ranking, which is 138 for American and 340 for global. For non-tenure-track positions, I use a 1/3 to 2/3 weighted average. For unranked tenure-track positions, I assign the maximum ranking plus one; for unranked postdocs and non-tenure-track positions, I assign 150 for American and 350 for global. Ultimately, for American placements, I use the American ranking, whereas for international placements, I use the global ranking times 3/7. Doing so ensures that all placements are ranked on the same scale.

Meanwhile, for non-academic placements, I construct academia-equivalent rankings shown in Appendix A. Following Krueger and Wu (2000), the most prestigious non-academic placements (Federal Reserve Board of Governors, IMF, World Bank, DeepMind, Google Research, and Microsoft Research) are given an academic-equivalent ranking of 40, with lower rankings down to 135 for less prestigious placements. As in Krueger and Wu (2000) and Bai et al. (2022), to avoid selection bias, I use publicly available LinkedIn profiles, graduate program websites, personal websites, PandaInUniv (2026), and García Guzmán (2026)'s *The Econ PhD placements dataset* to obtain outcomes data for all applicants, not just those who matriculate. The minority of applicants whose outcomes data are not available receive the lowest possible ranking of 150.

Regarding independent variables from the application files, I distinguish objective variables from subjective variables. Objective variables are those variables that the application software compiles into a spreadsheet for the committee's convenience, whereas subjective variables are those that the applicant uploads as pdf files that cannot be easily converted into a spreadsheet.[2] Because the economics committee does not have time to read everyone's file, they use the objective variables to filter out a predetermined percentage of the applicants before reading any files. For example, they filter out applicants with low quantitative GRE scores. For this reason, the structural models in this paper treat objective variables separately from subjective ones.

Examples of objective variables include demographics, such as application year, age, gender, and an ethnicity/region of the world categorical variable, in which the categories are U.S. White, U.S. African/Hispanic/Native American, U.S. Asian, International Asian, and International Other. There are also performance variables; such as verbal, quantitative, and analytical GRE scores; undergraduate GPA; TOEFL/IELTS scores for non-native English speakers; two binary variables for having a graduate degree and work experience respectively; and a categorical variable for the type of undergraduate institution (elite U.S. university, elite U.S. liberal arts college, other U.S. college/university, elite foreign university, and other foreign university), where elite is always defined as top 50. Moreover, recommender variables, namely the number of professor (Krueger & Wu 2000) and prolific recommenders the applicant has, measure how well-connected the applicant is. A prolific recommender is one who writes a letter for at least one other applicant in the sample, which Bai et al. (2022) find predictive of admission and success.

For non-economics applicants, there is a categorical variable for concentration, such as U.S. history vs. world history. For economics applicants who attended a masters program, I have their masters GPA and a categorical variable for the type of masters program (elite U.S., other U.S., elite international, and other international). I also have a binary variable for whether they have research experience and a regression-based score that the committee generated to summarize their objective characteristics. Finally, for the applicants whose files the committee read, I have the score that their assigned reader gave them relative to the scores that reader gave to other applicants. This score is a proxy for information that the committee can observe but I cannot.

Meanwhile, subjective variables in economics come from recommendation letters, application essays, CVs, and supplemental forms. I use textual processing algorithms on the letters to measure overall word count and counts of positive, negative, standout, ability, research, grindstone, teaching, communal, and agentic words (Bai et al. 2022). From application essays, the algorithms extract the word count and if the applicant mentioned a specific topic or professor's name (Grove et al. 2007). By conveying the applicant's interest in the program, these features are

[2] Bai et al. (2022) define objective variables as "verifiable performance measures" of applicant quality and subjective variables as committee member assessments. For most variables, our definitions coincide. While they do not categorize demographics, such as gender and nationality, I categorize demographics as objective variables.

intended to be predictive of whether the applicant will matriculate if accepted.

From CVs, there are binary variables for whether the applicant has economics-specific work experience, a working paper, a publication, and coding skills. Jones et al. (2020) find that economics committee members, especially at high-ranking programs, do not put much weight on coding skills. But given the increasing prevalence of empirical research that uses techniques like webscraping and machine learning, along with the increasing use of AI tools in economics, committees may be undervaluing such skills. Economics also has a supplemental form, in which the applicant lists the number of math courses they took, as well as the names of advanced math courses such as real analysis. I find that more math is predictive of getting accepted into the economics program, and I will determine whether it is also predictive of success.

The economics dataset contains all of the above variables, except when indicated otherwise and except for program completion and placements for the most recent years, considering those students are still in the program. In contrast, the dataset of all programs' applicants is more limited because it does not contain success data or subjective variables. But the latter dataset is still necessary for reduced-form analyses and statistical tests comparing the determinants of admission across programs. I present summary statistics from this dataset in Table 9 of Appendix B, as well as reduced-form results in Section 4. I find that for better or worse, economics admissions is more data-driven based on objective variables than the other programs.

3.2 Synthetic Data Protocol

In this section, I describe the synthetic data protocol used to protect the applicants' privacy, which I employ in conjunction with a university administrative office that is responsible for stewarding the data. While a handful of recent papers, such as Stanley and Totty (2024) and Carr et al. (2025) use synthetic data, the synthetic data in those papers had already been made available to researchers by the U.S. Census Bureau. In contrast, this paper's synthetic data protocol covers all the steps from creation of the synthetic data to obtaining the final results on the actual data.

The steps of the protocol are as follows. First, the administrative office receives the data from the graduate admissions office and individual departments, such as economics. They remove all direct identifiers, such as names and email addresses, and run algorithms I provide to extract variables from textual documents, such as recommendation letters. Second, they use Patki et al. (2016)'s publicly-available Synthetic Data Vault (SDV) to generate the synthetic data. The SDV estimates a Gaussian copula to capture correlations between variables and then simulates a new, synthetic dataset from the copula using a random number generator. Crucially, the rows in the synthetic dataset do not correspond to real subjects, which prevents identity disclosure (Chapelle & Falissard 2023). The SDV works best when the final versions of the variables are already created (e.g. turning a list of country names into a categorical variable), so without seeing any actual data, I work with the office to create the variables before implementing the SDV.

However, there remains the possibility of attribute disclosure. For example, if an applicant from an unusual country has an unusually low GRE score, information about that applicant may leak into the synthetic dataset. As such, if any cell in the synthetic data has fewer than five other cells from that variable with the same value, the variable's observations are binned and in each bin, replaced with the bin's sample mean. Only at this point do I receive the synthetic data. I use the data to finalize my estimation code, which I then give to the administrative department. The department takes that code, runs it on the actual data, and returns the parameter estimates.

As a demonstration of synthetic data, I use a Gaussian copula to simulate 1,000 observations of the following independent variables: quantitative GRE percentile (continuous), gender (binary), demographic (categorical with three demographics), advanced math (binary), and committee score (continuous). Then I use the inverse logit function to simulate binary dependent variables for transitioning past the first-round filter, getting accepted into the program, and obtaining a successful placement above a threshold ranking. The top frame of Figure 1 contains the first ten observations of this simulated dataset. Advanced math and committee score are missing for all applicants who did not survive the filter, a feature present in the actual data.

Transition	Accept	Success	GREQuantPct	Gender	Demographic1	Demographic2	AdvMath	ComScore
0	0	0	89	0	0	0	·	·
1	0	1	96	0	1	0	1	5
0	0	1	88	0	1	0	·	·
0	0	0	80	1	1	0	·	·
1	1	1	99	0	1	0	1	9
1	0	1	81	0	0	1	1	4
1	0	1	92	0	1	0	1	4
1	1	1	93	0	0	1	0	6
1	0	1	83	1	0	1	0	7
0	0	0	87	0	1	0	·	·

$$\Downarrow$$

Transition	Accept	Success	GREQuantPct	Gender	Demographic1	Demographic2	AdvMath	ComScore
1	0	0	90	1	1	0	0	5
1	0	0	83	0	0	1	1	6
1	1	1	91	0	0	1	0	5
1	1	1	92	0	0	1	0	5
0	0	1	83	0	0	0	·	·
1	0	0	93	0	0	1	0	9
0	0	0	84	0	0	1	·	·
1	1	1	88	0	1	0	1	4
1	0	1	86	0	1	0	0	6
0	0	1	86	1	0	1	·	·

Figure 1: Synthetic Data Example

I then pass this simulated dataset through a Gaussian copula and use the copula to generate a synthetic dataset from it, the first ten observations of which are in the bottom frame of the figure. The means, variances, and correlations of the variables are similar in both datasets, with only the correlations in the synthetic dataset slightly attenuated. However, none of the rows of the simulated dataset are present in the synthetic dataset. In addition, important statistical properties, such as *Accept* equaling 1 only when *Transition* equals 1, *Demographic1* and *2* never both equaling 1, and *AdvMath* and *ComScore* missing when *Transition* equals 0, are preserved.

Later in this paper, I simulate six datasets to evaluate the performance of the models and estimators while the synthetic data protocol is in progress. The first two datasets are for a Wald test comparing the determinants of admission and success across different programs. In one dataset, the determinants are the same, whereas in the other, they are different. The next two datasets are for the year-static and year-dynamic structural models. In one dataset, the committee admits applicants optimally, whereas in the other, it does not; for example, it discriminates against *Demographic1*. The last two datasets are for the waitlist-hybrid and waitlist-dynamic models; here, I add binary dependent variables for getting waitlisted and for matriculating if accepted. I also add a continuous matriculation score variable, and as before, the committee admits optimally in only one of the two datasets. In a future version of this paper, I will replace results from these simulated datasets with results from the actual data, though I will keep the simulated results in an appendix to show that the statistical tests are not prone to false positives or negatives.

4 Reduced-Form Analysis

In this section, I perform two reduced-form analyses. First, using the dataset of all 6,222 applicants to the university's economics, government, history, linguistics, and psychology PhD programs from 2020 to 2023, I regress a binary variable for acceptance on a set of objective variables, such as GRE scores. Summary statistics of the data are in Table 9 of Appendix B. According to this table, the economics program has the highest acceptance rate at 11.21%, and the government program has the lowest acceptance rate at 4.30%. I also run lasso and elastic net regressions to determine which variables are most predictive of acceptance. In a future version of this paper, I will regress success measures on objective and subjective variables for the economics program. The success measures will include binary variables for program completion and for obtaining an academic or non-academic job placement above a threshold ranking.

Second, on a simulated dataset of economics and non-economics programs, I run Wald tests of whether the determinants of acceptance, as well as the determinants of success, are the same across programs. As expected, for a simulated dataset where I set the data-generating parameters such that the determinants of acceptance and success are different, the test rejects the null. In contrast, for a dataset where these determinants are the same, the test does not reject. In general, I keep the data-generating parameters as close as possible to the actual reduced-form results.

4.1 Regressions and Lasso

In Table 1, I present logistic regressions of acceptance on objective independent variables for all five programs. The left subtable contains the regression coefficients, whereas the right subtable contains the average marginal effects (AMEs). There are several notable results. First, across programs, acceptance rates were lowest in Spring 2021, which was the year that Covid reduced department funding. Second, all else equal, economics favors younger applicants, and most programs favor women. Third, government and history favor underrepresented minorities, whereas economics disfavors applicants from East and South Asia. Fourth, GRE scores and undergraduate GPAs matter, but economics puts an especially large weight on the quantitative GRE score; a one point increase increases the probability of acceptance by over 1% point. Fifth, relative to a non-elite U.S. undergraduate institution, an elite institution matters more than a graduate degree or work experience. Sixth, as Bai et al. (2022) find, having a letter writer who recommended other applicants in the sample matters; this result provides evidence of network effects within the social sciences. Seventh, a strong math background matters for economics.

In addition, both in-sample and 5-fold-cross-validated (CV) Pseudo-R^2 values never exceed 0.3 for any program. However, they are highest for economics, meaning that economics admissions is more data-driven than the other programs, at least with respect to the data I can observe. For economics, the CV Pseudo-R^2 of 0.2105 is within a few percentage points of the in-sample Pseudo-R^2 of 0.2662, but that is not the case for the other programs. In fact, for linguistics and psychology, the CV Pseudo-R^2s are negative. Some of the other programs have additional admissions criteria, such as psychology, which interviews applicants. Although I cannot observe interview results, the interviews may explain some of the variation in acceptances.

Because of the low to negative CV Pseudo-R^2s, I run lasso (least absolute shrinkage and selection operator) and elastic net, with k-fold cross-validation for variable selection. In Table 2, I present the lasso estimates. All the CV Pseudo-R^2s are positive, and lasso selects more variables for economics than for any other program. The last three variables in this table are missing variable dummies. Any time an applicant's GRE scores, undergraduate GPA, or TOEFL/IELTS scores are missing, I impute the variable's sample mean for that observation and encode the missing variable dummy equal to 1 (Krueger & Wu 2000). I could implement group lasso (Yuan & Lin 2006) to force lasso to treat a variable and its corresponding missing dummy as a package, but allowing lasso to select one but not the other (see Bai et al. 2022) has a meaningful interpretation. For example, when lasso selects UgradGPA but not MissingUgradGPA for economics, that result means undergraduate GPAs, as opposed to TOEFL/IELTS scores, are missing at random.

Moreover, per Table 10 in Appendix B, several of the variables are correlated with one another. When variables are correlated, and lasso selects one but not the other, its decision may not be robust to sampling error. Therefore, in Table 11 of this appendix, I present elastic net results. Elastic net is a generalization of lasso, in which the estimator can select any combination of the L1 (lasso) or L2 (ridge) penalty. The closer α^* is to 0, the closer the penalty is to lasso, and the closer it is to 1, the closer the penalty is to ridge. Unlike lasso which removes some variables entirely, ridge only shrinks coefficients closer to 0. For all but economics, α^* is equal or close to 0. But for economics it equals 0.8, meaning that elastic net wants to keep most of the variables in the model. Finally, in Table 12 of the appendix, I present post-lasso regression results, in which I regress admission only on the predictors that lasso selects; these estimates are not shrunken.

Table 1: Logistic Regression Results

(a) Coefficients

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.3949**	-0.4337*	-0.5735*	-0.5501	-0.1766
	(0.2002)	(0.2612)	(0.3479)	(0.4129)	(0.5545)
WinsorAge	-0.1010***	0.0453*	0.0380	0.0222	-0.0520
	(0.0357)	(0.0271)	(0.0327)	(0.0532)	(0.0899)
Female	0.4100**	0.6717***	0.6529**	0.1311	-0.5060
	(0.1680)	(0.2299)	(0.2972)	(0.3452)	(0.5606)
USAfrHispNat	0.4147	1.0088***	1.1601***	0.0120	-0.8263
	(0.5212)	(0.3544)	(0.4425)	(0.8411)	(1.1611)
USAsian	-0.5283	-0.5657	-0.3263	0.8638	1.1711
	(0.4184)	(0.6504)	(1.2511)	(0.7201)	(0.7276)
IntAsian	-0.8279***	-0.1835	0.9396	0.4890	-0.7827
	(0.2940)	(0.4147)	(0.7230)	(0.6729)	(0.9323)
IntOther	0.5049	0.3812	1.3526***	-0.1061	0.6239
	(0.3159)	(0.4140)	(0.5085)	(0.7483)	(1.1070)
GREVerbalPct	0.0161***	0.0284*	0.0753**	-0.0092	0.0607***
	(0.0058)	(0.0159)	(0.0344)	(0.0154)	(0.0218)
GREQuantPct	0.1498***	0.0160*	0.0210*	0.0211	-0.0138
	(0.0177)	(0.0086)	(0.0116)	(0.0150)	(0.0159)
GREAnalyticPct	0.0176***	0.0198**	0.0030	0.0175	-0.0202
	(0.0041)	(0.0094)	(0.0134)	(0.0127)	(0.0218)
UgradGPA	2.2151***	0.6974	1.3647**	1.2645	1.7772*
	(0.7192)	(0.5477)	(0.5512)	(0.9703)	(0.9487)
TOEFL/IELTS	0.3851	0.6450	1.0949**	0.9655*	-0.4656
	(0.2888)	(0.4705)	(0.5519)	(0.5456)	(0.6112)
EliteUSUni	0.8855***	0.3934	0.6542*	-0.6057	1.4303**
	(0.2949)	(0.3136)	(0.3802)	(0.4985)	(0.6192)
EliteUSLA	1.1033***	0.6271*	1.0695**	-0.6088	2.3151**
	(0.3797)	(0.3754)	(0.5399)	(1.4186)	(1.0642)
EliteForeign	1.0154*	0.4747	1.7711**	1.6195*	0.3889
	(0.6051)	(0.8226)	(0.7940)	(0.8587)	(1.9889)
OtherForeign	0.5430	0.5767	1.3744**	0.4188	-1.0737
	(0.5766)	(0.7687)	(0.6994)	(0.7157)	(1.3016)
GradDegree	0.0494	0.0748	0.3131	0.4049	0.4892
	(0.2235)	(0.2586)	(0.3279)	(0.4896)	(0.6399)
WorkExper	0.3926**	0.4013	0.4959	0.1495	1.8201**
	(0.1888)	(0.2992)	(0.3297)	(0.3778)	(0.8564)
#ProfRecom	0.0370	-0.0180	0.5503**	0.4781*	-0.0981
	(0.1318)	(0.1339)	(0.2552)	(0.2662)	(0.2660)
#ProlificRecom	0.2663***	0.3025**	0.1876	0.3284**	0.5901*
	(0.0856)	(0.1261)	(0.1577)	(0.1544)	(0.3145)
#MathCourses	0.1644***				
	(0.0594)				
AdvMath	0.2424				
	(0.1791)				
MissingGRE	1.3967	-0.1169	0.6830	-0.3579	-1.0955*
	(0.9047)	(0.4078)	(0.6751)	(0.3763)	(0.6582)
MissingUgradGPA	-0.3762	-0.5233	-1.4922**	-1.0516	0.3958
	(0.4947)	(0.5416)	(0.6905)	(0.6524)	(1.1854)
MissingTOEFL/IELTS	-0.3703	0.6333	0.5708	-0.0184	-1.6283*
	(0.2444)	(0.5716)	(0.5128)	(0.5111)	(0.9429)
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486
Pseudo-R^2	0.2662	0.1526	0.1983	0.1871	0.2475
CV Pseudo-R^2	0.2105	0.0721	0.0605	-0.0265	-0.2787

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) AMEs

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.0303**	-0.0164	-0.0430	-0.0298	-0.0075
	(0.0153)	(0.0100)	(0.0261)	(0.0224)	(0.0238)
WinsorAge	-0.0077***	0.0017*	0.0028	0.0012	-0.0022
	(0.0027)	(0.0010)	(0.0025)	(0.0029)	(0.0038)
Female	0.0314**	0.0254***	0.0489**	0.0071	-0.0216
	(0.0129)	(0.0088)	(0.0220)	(0.0187)	(0.0239)
USAfrHispNat	0.0318	0.0382***	0.0869***	0.0007	-0.0352
	(0.0400)	(0.0137)	(0.0332)	(0.0455)	(0.0487)
USAsian	-0.0405	-0.0214	-0.0244	0.0468	0.0499
	(0.0321)	(0.0247)	(0.0936)	(0.0391)	(0.0313)
IntAsian	-0.0635***	-0.0069	0.0704	0.0265	-0.0334
	(0.0224)	(0.0157)	(0.0542)	(0.0363)	(0.0393)
IntOther	0.0387	0.0144	0.1013***	-0.0057	0.0266
	(0.0243)	(0.0157)	(0.0385)	(0.0406)	(0.0471)
GREVerbalPct	0.0012***	0.0011*	0.0056**	-0.0005	0.0026***
	(0.0004)	(0.0006)	(0.0026)	(0.0008)	(0.0010)
GREQuantPct	0.0115***	0.0006*	0.0016*	0.0011	-0.0006
	(0.0013)	(0.0003)	(0.0009)	(0.0008)	(0.0007)
GREAnalyticPct	0.0013***	0.0008**	0.0002	0.0009	-0.0009
	(0.0003)	(0.0004)	(0.0010)	(0.0007)	(0.0009)
UgradGPA	0.1699***	0.0264	0.1022**	0.0684	0.0757*
	(0.0550)	(0.0208)	(0.0423)	(0.0522)	(0.0418)
TOEFL/IELTS	0.0295	0.0244	0.0820**	0.0523*	-0.0198
	(0.0220)	(0.0179)	(0.0416)	(0.0301)	(0.0259)
EliteUSUni	0.0679***	0.0149	0.0490*	-0.0328	0.0610**
	(0.0224)	(0.0119)	(0.0281)	(0.0264)	(0.0271)
EliteUSLA	0.0846***	0.0237*	0.0801**	-0.0330	0.0987**
	(0.0287)	(0.0143)	(0.0402)	(0.0761)	(0.0480)
EliteForeign	0.0779*	0.0180	0.1326**	0.0877*	0.0166
	(0.0463)	(0.0311)	(0.0591)	(0.0481)	(0.0850)
OtherForeign	0.0416	0.0218	0.1029*	0.0227	-0.0458
	(0.0442)	(0.0291)	(0.0527)	(0.0391)	(0.0556)
GradDegree	0.0038	0.0028	0.0234	0.0219	0.0208
	(0.0171)	(0.0098)	(0.0247)	(0.0262)	(0.0273)
WorkExper	0.0301**	0.0152	0.0371	0.0081	0.0776**
	(0.0144)	(0.0113)	(0.0248)	(0.0205)	(0.0368)
#ProfRecom	0.0028	-0.0007	0.0412**	0.0259*	-0.0042
	(0.0101)	(0.0051)	(0.0189)	(0.0142)	(0.0113)
#ProlificRecom	0.0204***	0.0114**	0.0140	0.0178**	0.0251*
	(0.0065)	(0.0048)	(0.0118)	(0.0082)	(0.0130)
#MathCourses	0.0126***				
	(0.0045)				
AdvMath	0.0186				
	(0.0137)				
MissingGRE	0.1071	-0.0044	0.0512	-0.0194	-0.0467
	(0.0696)	(0.0154)	(0.0506)	(0.0202)	(0.0286)
MissingUgradGPA	-0.0289	-0.0198	-0.1118**	-0.0569	0.0169
	(0.0380)	(0.0205)	(0.0524)	(0.0357)	(0.0506)
MissingTOEFL/IELTS	-0.0284	0.0240	0.0428	-0.0010	-0.0694
	(0.0189)	(0.0217)	(0.0383)	(0.0277)	(0.0423)
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Before proceeding to the Wald tests, I will discuss inference vs. prediction, as well as sample selection. Both will be important when running regressions with success measures as dependent variables. Regarding inference vs. prediction, omitted independent variables will make it difficult to obtain consistent estimates in any success regression. For example, I do not observe the effort that students put in after matriculating. Even effort, let alone all the variables in the application files that I do not observe, can bias success parameter estimates upward if their corresponding variables are correlated with effort and with success irrespective of effort (Ferreyra et al. 2023).

Table 2: Lasso Logistic Regression Results

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.3334	-0.3120	-0.3871	-0.2924	
WinsorAge	-0.0773	0.0255	0.0202		
Female	0.3347	0.5050	0.5004		-0.0352
USAfrHispNat	0.2628	0.8088	0.6977		
USAsian	-0.2655	-0.2768		0.1859	0.0379
IntAsian	-0.6301	-0.0534		0.1624	
IntOther	0.5106	0.1989	0.7227	-0.1071	
GREVerbalPct	0.0142	0.0264	0.0335		0.0330
GREQuantPct	0.1296	0.0104	0.0184	0.0094	
GREAnalyticPct	0.0163	0.0210		0.0107	
UgradGPA	1.9458	0.4405	0.7810	0.3589	
TOEFL/IELTS	0.3969	0.3010	0.6236	0.5785	
EliteUSUni	0.6505	0.2684	0.3456		0.4364
EliteUSLA	0.8434	0.5163	0.6424		0.3557
EliteForeign	0.4020		0.1087	0.5337	
OtherForeign					
GradDegree			0.1312		
WorkExper	0.2807	0.3413	0.3321		0.4843
#ProfRecom			0.3858	0.2626	
#ProlificRecom	0.2280	0.2335	0.1710	0.2436	0.0275
#MathCourses	0.1448				
AdvMath	0.2339				
MissingGRE	0.5886			-0.2281	-0.5146
MissingUgradGPA					
MissingTOEFL/IELTS	-0.2503	0.2865	0.0504		
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486
Pseudo-R^2	0.2623	0.1453	0.1647	0.1552	0.1269
CV Pseudo-R^2	0.2279	0.0923	0.0598	0.0804	0.0535

However, there are reasons for reassurance. First, in a later version of this paper, I will incorporate committee scores as an independent variable, which will proxy variables that the committee observes and I do not. I will also follow Bai et al. (2022) and include as a proxy of effort the number of "grindstone terms," such as depend* and diligen*, in recommendation letters. In addition, in this paper, prediction matters as much as inference. Mullainathan and Spiess (2017) distinguish $\hat{\beta}$ (e.g. the effect of hiring an additional teacher) from $\hat{y}$ (e.g. which teacher to hire) problems. For graduate admissions, if GRE scores predict success, the committee should use them regardless of whether it knows the mechanism through which they affect success.

In addition, the determinants of success conditional on applying (see Krueger and Wu 2000, and Grove and Wu 2007) are different from those conditional on matriculating (see Athey et al. 2007). Bai et al. (2022) prove that if $P(\text{Success}|\text{Skill})$ is an S-curve, then $\hat{\beta}$ will be smaller for more selected samples, a theoretical finding consistent with the significance of the quantitative GRE being weaker in studies on matriculants only. This problem is akin to height not predicting success in a sample of professional basketball players. Like Bai et al. (2022), I collect placement outcomes for all applicants, not just those who attended the university. But there are other measures of success, such as course grades and passing comprehensive examinations, that are only available for matriculants, and that have been used in papers like Athey et al. (2007).

As such, for these measures, one can implement a Heckman (1979) correction. Notationally, let $y_n^* = x_n\beta + u_n$, $d_n = \mathbb{1}(z_n\gamma + v_n > 0)$, and $y_n = d_n y_n^*$. y is a measure of success, x are variables in applications, and d is matriculation. The first step is to run a probit regression of d on z and compute the inverse Mills ratio $\hat{\lambda}_n = \frac{\phi(z_n\hat{\gamma})}{\Phi(z_n\hat{\gamma})}$. The second step is to run OLS, probit, or ordered probit of y on $(x, \hat{\lambda})$ to obtain $\hat{\beta}$. For Heckman instruments, one needs variables that affect d but not y. Such variables include the applicant's enthusiasm about the program (e.g. whether they included a specific professor's name in their application essay), along with the average acceptance rate and average quality of accepted applicants in the applicant's year. If the correction works well for success measures for which I have data for all applicants and can therefore verify its effectiveness, then it may also work well for measures for which I do not. In Appendix C, I incorporate a Heckman correction into a structural model estimated on simulated data.

4.2 Wald Tests Across Programs

In addition to regressions and lasso, I test on simulated data if the determinants of acceptance and success are the same across programs. For simplicity, I simulate data for only the economics program and two non-economics programs. However, I simulate two different datasets. In the first dataset, the means, standard deviations, and correlations (see Tables 13 and 14 in Appendix B) are similar to the actual data (at least for acceptance), meaning that the determinants of acceptance and success are different across programs.[3] In the second dataset, I modify the correlations, so the determinants for the non-economics programs are the same as those for economics.

The Wald test's setup is as follows. Assume $\left(Y^E \perp\!\!\!\perp Y^{NE}\right)\big|X$; that is, conditional on the independent variables X, the dependent variable (either admission or success) for economics Y^E is independent of its non-economics counterpart Y^{NE}. Also, suppose the dependent variables are binary. For example, accepted vs. rejected or successful vs. unsuccessful. As such, I parameterize the economics and non-economics outcome probabilities with the inverse logit functional form:

[3]For some programs, the initial job placement matters less career-wise than it does for economics, which is worth considering when interpreting the Wald test results.

$P\left(Y_n^E=1\middle|x_n^E;\theta^E\right)=\frac{\exp(x_{n,0}^E\theta_0^E+x_{n,1}^E\theta_1^E)}{1+\exp(x_{n,0}^E\theta_0^E+x_{n,1}^E\theta_1^E)}$ and $P\left(Y_n^{NE}=1\middle|x_n^{NE};\theta^{NE}\right)=\frac{\exp(x_{n,0}^{NE}\theta_0^{NE}+x_{n,1}^{NE}\theta_1^{NE})}{1+\exp(x_{n,0}^{NE}\theta_0^{NE}+x_{n,1}^{NE}\theta_1^{NE})}$.

Notably, I separate the intercept θ_0 from the slope θ_1, since the goal is to test for equal slopes, not equal intercepts. That said, I can add slopes I do not test to θ_0. Then the likelihood is:

$$\mathcal{L}_N^U\left(y\middle|x;\theta\right)=\prod_{n=1}^{N_E}P\left(Y_n^E=1\middle|x_n^E;\theta^E\right)^{Y_n^E}\left[1-P\left(Y_n^E=1\middle|x_n^E;\theta^E\right)\right]^{1-Y_n^E}$$
$$\prod_{n=1}^{N_{NE}}P\left(Y_n^{NE}=1\middle|x_n^{NE};\theta^{NE}\right)^{Y_n^{NE}}\left[1-P\left(Y_n^{NE}=1\middle|x_n^{NE};\theta^{NE}\right)\right]^{1-Y_n^{NE}}. \quad (1)$$

Let $g(\hat\theta)=\frac{1}{N_E}\sum_{n=1}^{N_E}\hat P_n^E(1-\hat P_n^E)\hat\theta_1^E-\frac{1}{N_{NE}}\sum_{n=1}^{N_{NE}}\hat P_n^{NE}(1-\hat P_n^{NE})\hat\theta_1^{NE}$. The Wald test statistic is therefore $\mathrm{W}=g(\hat\theta)'[\nabla_\theta g(\hat\theta)\hat V(\hat\theta)\nabla_\theta g(\hat\theta)']^{-1}g(\hat\theta)\sim\chi_Q^2$, where Q is the number of restrictions; that is, the number of determinants I test at a time. Note that the null hypothesis is of equal AMEs, not equal logit coefficients, since only AMEs can be compared across samples (Mood 2010).

If the dependent variable is instead continuous (e.g. placement rank), I parameterize its economics and non-economics distributions as: $f\left(y_n^E\middle|x_n^E;\beta^E,\sigma_E^2\right)=\frac{1}{\sqrt{2\pi\sigma_E^2}}\exp\left[\frac{-1}{2\sigma_E^2}\left(y_n^E-x_{n,0}^E\beta_0^E-x_{n,1}^E\beta_1^E\right)^2\right]$ and $f\left(y_n^{NE}\middle|x_n^{NE};\beta^{NE},\sigma_S^2\right)=\frac{1}{\sqrt{2\pi\sigma_{NE}^2}}\exp\left[\frac{-1}{2\sigma_{NE}^2}\left(y_n^{NE}-x_{n,0}^{NE}\beta_0^{NE}-x_{n,1}^{NE}\beta_1^{NE}\right)^2\right]$. Then:

$$\mathcal{L}_N^U\left(y\middle|x;\beta,\sigma^2\right)=\prod_{n=1}^{N_E}f\left(y_n^E\middle|x_n^E;\beta^E,\sigma_E^2\right)\prod_{n=1}^{N_{NE}}f\left(y_n^{NE}\middle|x_n^{NE};\beta^{NE},\sigma_{NE}^2\right). \quad (2)$$

Since linear coefficients can be compared across samples, the test is simpler (and can also be done with a likelihood-ratio test). Let $g(\hat\beta_1)=\hat\beta_1^E-\hat\beta_1^{NE}$. Under homoskedasticity (i.e. $\forall n,\sigma_n^2=\sigma^2$), the Wald statistic is $\mathrm{W}=g(\hat\beta_1)'[\nabla_{\beta,\sigma^2}g(\hat\beta_1)\hat V(\hat\beta,\hat\sigma^2)\nabla_{\beta,\sigma^2}g(\hat\beta_1)']^{-1}g(\hat\beta_1)\sim\chi_Q^2$.

In Table 3, I present binary Wald test results on the simulated data. In the dataset where the data-generating parameters are such that the true AMEs of the determinants of acceptance and success are different, the test rejects the null with $p<0.0001$. But in the dataset where the determinants are the same, the test does not reject; I obtain $p=0.6843$ for the acceptance dependent variable and $p=0.6993$ for the success dependent variable. Therefore, on the simulated data, the Wald test works properly, rejecting resoundingly when it should and not rejecting when it should not. Because the simulated dataset has 5,000 observations, which is of the same order of magnitude as the number of observations in the actual data, the test is sufficiently powerful.

In addition to testing if the determinants of acceptance or success are the same across programs, could I test if the determinants of acceptance are the same as those of success within a program? For example, does the quantitative GRE have the same AME for admission as for success? Intuitively, provided the committee wants to select applicants based on a “success” model, predictors of admission only may be overvalued, and vice versa. Such a test would be fatally flawed, however. First, there may be unwanted rejections of the null. For example, the committee may value diversity if it does not hurt $P(\text{Success})$ too much, in which case it may favor demographics not strictly more likely to succeed. There may also be unwanted non-rejections.

Table 3: Wald Test Results

(a) Different Determinants

	DV = Accept				DV = Success			
	Economics		Non-Economics		Economics		Non-Economics	
	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-16.3291***	-1.5784***	-8.4753***	-0.3642***	-10.4581***	-1.9104***	-3.2236***	-0.5997***
	(1.1274)	(0.0955)	(0.8673)	(0.0412)	(0.7237)	(0.1070)	(0.4018)	(0.0728)
GREQuantPct	0.1502***	0.0145***	0.0402***	0.0017***	0.1058***	0.0193***	0.0034	0.0006
	(0.0131)	(0.0012)	(0.0128)	(0.0006)	(0.0088)	(0.0014)	(0.0062)	(0.0012)
Gender	0.4702***	0.0454***	0.5942***	0.0255***	-0.1372	-0.0251	0.0375	0.0070
	(0.1473)	(0.0140)	(0.1820)	(0.0079)	(0.1095)	(0.0200)	(0.0868)	(0.0161)
Demographic1	-0.6246***	-0.0604***	-0.2541	-0.0109	-0.3544**	-0.0647**	-0.1152	-0.0214
	(0.2085)	(0.0200)	(0.2393)	(0.0103)	(0.1467)	(0.0267)	(0.1103)	(0.0205)
Demographic2	0.1555	0.0150	0.0081	0.0003	0.1713	0.0313	-0.0562	-0.0105
	(0.2070)	(0.0200)	(0.2109)	(0.0091)	(0.1557)	(0.0284)	(0.1030)	(0.0192)
AdvMath	0.2886*	0.0279*			0.3284***	0.0600***		
	(0.1611)	(0.0156)			(0.1112)	(0.0202)		
ComScore	0.2241***	0.0217***	0.3817***	0.0164***	0.1026***	0.0187***	0.1808***	0.0336***
	(0.0413)	(0.0039)	(0.0528)	(0.0025)	(0.0287)	(0.0052)	(0.0228)	(0.0041)
NonEconomics1			-0.0452	-0.0019			1.1393***	0.2120***
			(0.1995)	(0.0086)			(0.1219)	(0.0219)
Obs | Pseudo-R^2	5000	0.1488			5000	0.0758		
LL | Param | AIC	-1155.7	14	2339.5		-2747.2	14	5522.4	
Wald Stat | DF | P-Val	112.5897	5	0.0000		124.7379	5	0.0000	

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Same Determinants

	DV = Accept				DV = Success			
	Economics		Non-Economics		Economics		Non-Economics	
	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-16.3848***	-1.5822***	-28.5524***	-1.3677***	-10.4586***	-1.9105***	-10.2610***	-1.8514***
	(1.1315)	(0.0956)	(1.4963)	(0.0500)	(0.7236)	(0.1070)	(0.4815)	(0.0628)
GREQuantPct	0.1509***	0.0146***	0.3245***	0.0155***	0.1058***	0.0193***	0.1154***	0.0208***
	(0.0131)	(0.0012)	(0.0179)	(0.0007)	(0.0088)	(0.0014)	(0.0070)	(0.0011)
Gender	0.4714***	0.0455***	0.7880***	0.0377***	-0.1384	-0.0253	-0.0497	-0.0090
	(0.1475)	(0.0140)	(0.1696)	(0.0081)	(0.1095)	(0.0200)	(0.0895)	(0.0161)
Demographic1	-0.6287***	-0.0607***	-1.8488***	-0.0886***	-0.3486**	-0.0637**	-0.3079***	-0.0556***
	(0.2088)	(0.0200)	(0.2445)	(0.0115)	(0.1467)	(0.0267)	(0.1123)	(0.0202)
Demographic2	0.1531	0.0148	0.3703*	0.0177*	0.1756	0.0321	0.0282	0.0051
	(0.2072)	(0.0200)	(0.1929)	(0.0091)	(0.1558)	(0.0284)	(0.1069)	(0.0193)
AdvMath	0.2889*	0.0279*			0.3267***	0.0597***		
	(0.1612)	(0.0156)			(0.1112)	(0.0202)		
ComScore	0.2241***	0.0216***	0.3994***	0.0191***	0.1026***	0.0187***	0.0757***	0.0137***
	(0.0413)	(0.0039)	(0.0506)	(0.0023)	(0.0287)	(0.0052)	(0.0228)	(0.0041)
NonEconomics1			0.8376***	0.0401***			1.9483***	0.3515***
			(0.2297)	(0.0107)			(0.1236)	(0.0205)
Obs | Pseudo-R^2	5000	0.3251			5000	0.1431		
LL | Param | AIC	-1130.4	14	2288.8		-2691.4	14	5410.8	
Wald Stat | DF | P-Val	3.1018	5	0.6843		3.0042	5	0.6993	

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

If there is discrimination in the job market, the committee may rationalize discriminating in the same way, which would make the determinants of admission equal to those of success.

But more importantly, this test would not account for the mechanics of the admissions process. As a simple example, suppose the following conditions hold: (1) Advanced math has a large positive effect on success, and coding skills have a smaller positive effect. (2) 25% of applicants have both skills and therefore the highest $P(\text{Success})$, 25% have math only (2nd-highest), 25% have coding only (3rd-highest), and 25% have neither (lowest). (3) This university accepts 50% of the applicants, and the other 50% will attend a lower-ranked school, ensuring everyone's outcomes are observable. (4) Unlike applicant quality, school quality does not affect outcomes.

Under these suppositions, an optimizing committee will accept the 50% of applicants with math and ignore coding. As a result, an acceptance regression will have a degenerate math AME of 1 and a degenerate coding AME of 0, while a success regression will have the following AMEs: $1 > \text{math} > \text{coding} > 0$. Also, for the acceptance AMEs to equal the success AMEs, the committee would suboptimally have to accept applicants without math. Therefore, unlike with, for instance, wages equaling marginal revenue product in a competitive labor market which allows for cross-regression comparisons, the AMEs for admission and success need not be equal for the committee to be optimizing. In turn, to draw conclusions about optimality, structural models of the committee's selection problem are necessary; these models are the basis for Section 5.

That said, within the Wald test framework, I can answer additional questions, as long as those questions are not about optimality. For example, I can test if there are different determinants of different types of success. Specifically, applicants may need different skills at different stages of the program; Athey et al. (2007) find that the quantitative GRE matters for course grades, whereas research experience matters more for placements. Statistically testing this hypothesis may be useful for committees who must weigh the early vs. late stages when deciding which applicants to accept. I can also test if the determinants of admission or success are different for U.S. vs. international applicants (see Krueger and Wu 2000). International applicants may have different levels of preparation or research interests, and be more likely to complete the program for lack of outside options and return to their native country after graduation. Finally, I can test if the determinants have changed over time. Reasons for such changes may include Covid and other macro/geopolitical factors, and for economics in particular, a shift in the profession toward empirics, coding/AI-intensive research, and elite predocs and Federal Reserve RAships.

5 Structural Analysis

There are two main reasons to estimate structural models of the admissions committee's decision problem. First, to make any claims about whether committees are optimizing, it is necessary to define what optimization is by explicitly modeling their selection problem. Second, structural models can uncover feasible counterfactual deviation decision rules to help committees admit stronger applicants, and the models can quantify any deviation gains from such rules.

The question then is how to set up the models. Bai et al. (2022) prove that if all the variables in an application file can be reduced to a single number (i.e. an expected success probability),[4] then accepting the optimal subset of applicants is the same as using a cutoff rule. A cutoff rule problem is far more tractable than a portfolio choice problem because I do not have to consider all the subsets of the hundreds of the applicants in each year who can be accepted; I only have to rank the applicants and choose the ones above the cutoff. Because the committee may be more selective in some years than others, I will allow the cutoff probability to vary across years.

After recasting into this paper's notation Bai et al. (2022)'s proof that the committee's portfolio choice problem reduces to a cutoff rule, I start by modeling the committee's problem as occurring over two rounds. In Round 1, the committee observes only each applicant's objective variables, such as GRE scores. Based on those variables, it can either transition the applicant to Round 2, or it can issue an immediate rejection. If the applicant survives Round 1, the committee then takes the time to read the application file and access all variables, objective and subjective (e.g. recommendation letters). Finally, it accepts or rejects the applicant. Rather than setting up a two-round model, why do I not use regressions, lasso, neural networks, or any other econometric or machine learning method to rank the applicants based on $P(\text{Success})$ conditional on all data and compare this simple model's ranking to the committee's?

The reason is that while I can still use lasso for variable selection to reduce the state space of the two-round models, a one-round model is cheating because the committee does not have time to read everyone's files (though in the future, perhaps committees can use textual processing algorithms to "read" files quickly and generate additional objective variables without too much effort). Moreover, two-round models are useful not just for graduate admissions but for any other setting with first-stage filtering, such as hiring, clinical trials for medicines, etc.

I start with two different two-round models: (single) year-static and year-dynamic. In the year-static model, the committee ranks the applicants based on their objective variables, reads the files of those whose expected success probabilities are above the first cutoff probability, re-ranks the surviving applicants based on all their data, and accepts those above the second cutoff probability. I refer to this model as static because each round is a separate static problem. In contrast, in the year-dynamic model, the committee only receives a payoff from accepted applicants and forms a conditional expectation in Round 1 about those who survive to Round 2. As such, the first cutoff is not the expected success probability of the marginal transitioned applicant, but the level of that applicant's expected Round 2 value. As such, it can take a value greater than 1. In Figure 5's visualization, T stands for transition, A for accept, and R for reject.

This model appears more sophisticated because it favors applicants with objective characteristics that are positively correlated with good subjective characteristics. However, the year-static

[4]The assumption that committees select based on expected success may have become more realistic in light of the 2023 Supreme Court ruling against race-based affirmative action in *Students for Fair Admissions v. Harvard*.

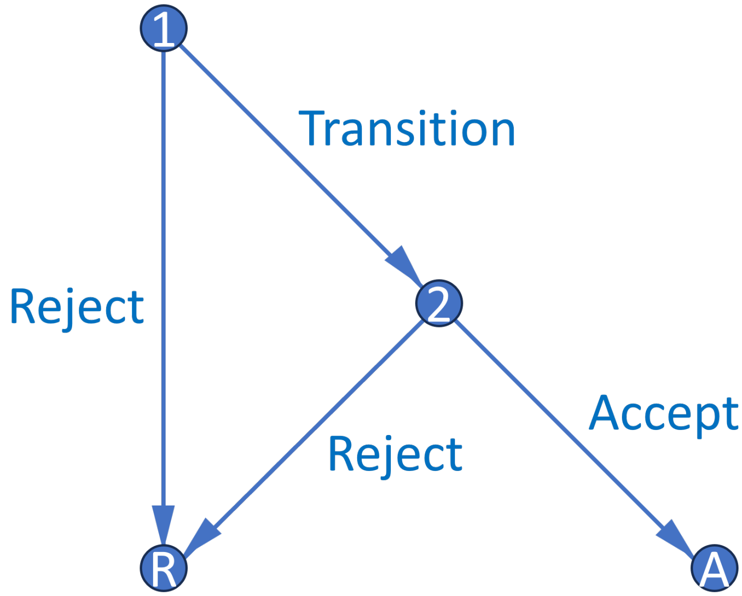


Figure 2: Decision Graph

model also correctly favors those applicants because as a result of omitted variable bias, the parameter estimates for such objective characteristics are biased upward. Only a year-myopic model, in which the setup is identical to the year-dynamic model, except for the fact that the parameters governing the expectation of the subjective characteristics conditional on the objective characteristics are restricted to zero, will not favor such applicants. Later in this section, when I present maximum likelihood estimation results of these models on simulated data, I include the Akaike Information Criterion $\text{AIC} = 2[K - \log(\hat{\mathcal{L}}^R)]$ to compare different models and determine whether the extra complexity meaningfully improves the model's fit. K is the number of parameters, and $\hat{\mathcal{L}}^R$ is the restricted likelihood's value; lower AIC values are better.

Still later in this section, I present three-round models that include waitlists. In Round 2 of these models, the committee does not accept or reject the surviving applicants. Rather, it early sends them an early acceptance, or it waitlists them. It then sees which of the early accepted applicants matriculate and which do not. Finally, in Round 3, it accepts or rejects the waitlisted applicants. The key dynamic is that the committee favors likely matriculants in Round 2, in addition to applicants with high expected success probabilities. It does so to avert a Round 3 situation in which it has barely filled up its class, applicants it could have accepted in Round 2 chose other schools, and now it has to accept weaker applicants. Finally, at the end of this section, I show how the models can be used to uncover counterfactual deviation decision rules. I also show how much more effective these rules are at admitting the best applicants.

5.1 Portfolio Choice to Cutoff Rule

Bai et al. (2022) prove that selecting the best N_A applicants out of N is equivalent to selecting the N_A whose success probability exceeds a cutoff. Let x be a vector of objective variables (e.g. GRE scores), and z be a vector of subjective variables (e.g. variables from recommendation letters). Also, let d_n denote the accept or reject decision for the nth applicant, and y_n denote that applicant's outcome, where $S =$ success and $F =$ failure. The committee solves a constrained optimization problem, in which it maximizes the sum of the success probabilities of the accepted applicants under the constraint that it cannot accept more than N_A applicants. For simplicity, suppose there is one round in one year, and all accepted applicants matriculate:

$$\max_{(d_1,\ldots,d_N)\in\{A,R\}^N} \sum_{n=1}^{N} P(y_n = S|x_n,z_n)\mathbb{1}(d_n = A), \text{ such that } \sum_{n=1}^{N} \mathbb{1}(d_n = A) \leq N_A. \tag{3}$$

Because success probabilities are never negative, the unconstrained objective increases every time $d_n = A$, which means that the constraint binds. The Lagrangian with multiplier λ is:

$$\begin{aligned} \mathcal{L}(d_1,\ldots,d_N;\lambda) &= \sum_{n=1}^{N} P(y_n = S|x_n,z_n)\mathbb{1}(d_n = A) + \lambda\left[N_A - \sum_{n=1}^{N} \mathbb{1}(d_n = A)\right] \\ &= \sum_{n=1}^{N} \left[P(y_n = S|x_n,z_n) - \lambda\right]\mathbb{1}(d_n = A) + \lambda N_A. \end{aligned} \tag{4}$$

Suppose without loss of generality that the applicants are sorted in decreasing order such that $P(y_1 = S|x_1,z_1) \geq \cdots \geq P(y_N = S|x_N,z_N)$. To ensure that the committee accepts exactly N_A applicants, it must be the case that $\lambda^* \in \left[P(y_{N_A+1} = S|x_{N_A+1},z_{N_A+1}), P(y_{N_A} = S|x_{N_A},z_{N_A})\right]$. That is, in this setup, λ^* is set identified.[5] Then the solution is to accept all applicants whose success probability exceeds λ^* and reject the rest, which means that λ^* is a cutoff probability:

$$\forall n \in \{1,\ldots,N\},\, d_n(x_n,z_n|\lambda^*) = \begin{cases} A & \text{if } P(y_n = S|x_n,z_n) > \lambda^* \\ \{A,R\} & \text{if } P(y_n = S|x_n,z_n) = \lambda^* \\ R & \text{if } P(y_n = S|x_n,z_n) < \lambda^*. \end{cases} \tag{5}$$

If $P(y_n = S|x_n,z_n) = \lambda^*$, then the committee is indifferent between acceptance and rejection, which means that in the event of a tie at the margin, the committee will randomly accept enough applicants to reach N_A acceptances. Also, while λ^* is set identified in this setup, I show in the proof of Theorem 1 in the next subsection how to point identify it with admissions data by modeling a committee that compares each applicant's success probability to a single-valued cutoff, and then by exploiting the observed committee's exact selectivity. Finally, given a large number of applicants, λ^* can be interpreted as the marginal accepted applicant's success probability. Therefore, I denote it going forward as $P(y^* = S)$, or the success probability of that applicant.

[5] If there were a continuum of applicants as in Bai et al. (2022)'s proof, the set's length would be 0, and so λ^* would be point identified. But all the models in this paper have N applicants, as there are in the data.

5.2 Year-Static Model

The notation for the year-static model is as follows. $r \in \{1,2\}$ is the round, x are objective variables, z are subjective variables, t is the year, and ε_r is a type 1 extreme value shock with scale parameter σ. One can imagine the shocks randomly affecting the committee's perception of each applicant's success probability. The shocks allow the model to fit the data, given that a real-world committee will not always accept the applicants with the highest success probabilities.

Starting in Round 2, the committee chooses $d_2 \in \{\text{Accept}, \text{Reject}\}$. If it accepts an applicant, it gets the applicant's success probability $P(y = S|x,z)$, whereas if it rejects the applicant, it gets the Round 2 cutoff probability $P(y_2^* = S|t)$, which is the success probability of t's marginal accepted applicant, whose success probability is conditional on x and z. There are various measures of success, from completing the program to obtaining a job placement above a threshold ranking.

Meanwhile in Round 1, the committee chooses $d_1 \in \{\text{Transition}, \text{Reject}\}$. If it transitions an applicant, it gets $P(y = S|x)$, or the applicant's success probability, which may differ from $P(y = S|x,z)$. However, if it rejects the applicant, it gets $P(y_1^* = S|t)$, or the success probability of t's marginal transitioned applicant, whose success probability is conditional on x only. $P(y_1^* = S|t)$ can be less or greater than $P(y_2^* = S|t)$ depending on how much the committee values its time vs. admitting the best possible applicants. As such, the Round 2 and Round 1 value functions are:

$$v_2(x,z,t,\varepsilon_2) = \max_{d_2 \in \{A,R\}} u_2(x,z,t,d_2) + \varepsilon_2(d_2), \text{ such that:} \tag{6}$$

$$u_2(x,z,t,d_2) = P(y = S|x,z)\mathbb{1}(d_2 = A) + P(y_2^* = S|t)\mathbb{1}(d_2 = R)$$

$$v_1(x,t,\varepsilon_1) = \max_{d_1 \in \{T,R\}} u_1(x,t,d_1) + \varepsilon_1(d_1), \text{ such that:} \tag{7}$$

$$u_1(x,t,d_1) = P(y = S|x)\mathbb{1}(d_1 = T) + P(y_1^* = S|t)\mathbb{1}(d_1 = R).$$

Each round's conditional choice probabilities (CCPs) then have the closed-form expression:

$$P(d_r|x,(z),t) = \frac{\exp(u_r(x,(z),t,d_r)/\sigma)}{\sum_{\widetilde{d}_r \in D_r} \exp(u_r(x,(z),t,\widetilde{d}_r)/\sigma)}. \tag{8}$$

To estimate this model, the first step is to set up the unrestricted likelihood. I parameterize the unrestricted CCPs (P_{d_1},P_{d_2}), success probabilities (P_{y_1},P_{y_2}), and cutoff probabilities ($P_{y_1^*}$,$P_{y_2^*}$) with the following flexible logit specifications with parameters θ:

$$P_{d_1} = P\big(d_1 = T\big|x,t;\theta_T\big) = \frac{\exp(f_T(x,t)\theta_T)}{1+\exp(f_T(x,t)\theta_T)}, \quad P_{d_2} = P\big(d_2 = A\big|x,z,t;\theta_A\big) = \frac{\exp(f_A(x,z,t)\theta_A)}{1+\exp(f_A(x,z,t)\theta_A)} \tag{9}$$

$$P_{y_1} = P\big(y = S\big|x;\theta_{S_1}\big) = \frac{\exp(f_{S_1}(x)\theta_{S_1})}{1+\exp(f_{S_1}(x)\theta_{S_1})}, \quad P_{y_2} = P\big(y = S\big|x,z;\theta_{S_2}\big) = \frac{\exp(f_{S_2}(x,z)\theta_{S_2})}{1+\exp(f_{S_2}(x,z)\theta_{S_2})} \tag{10}$$

$$P_{y_1^*} = P\big(y_1^* = S\big|t;\theta_{S_1^*}\big) = \frac{\exp(f_{S_1^*}(t)\theta_{S_1^*})}{1+\exp(f_{S_1^*}(t)\theta_{S_1^*})}, \quad P_{y_2^*} = P\big(y_2^* = S\big|t;\theta_{S_2^*}\big) = \frac{\exp(f_{S_2^*}(t)\theta_{S_2^*})}{1+\exp(f_{S_2^*}(t)\theta_{S_2^*})}. \tag{11}$$

Let N_T be the number of N total applicants who survive to Round 2. Also, T = transition, A = accept, R = reject, S = success, and F = failure. The unrestricted likelihood is then:

$$\mathcal{L}_N^U(\theta) = \prod_{n=1}^{N_T} P_{d_{1,n}}(\theta_T) P_{d_{2,n}}(\theta_A)^{\mathbb{1}(d_{2,n}=A)} (1-P_{d_{2,n}}(\theta_A))^{\mathbb{1}(d_{2,n}=R)} \left[P_{y_{1,n}}(\theta_{S_1}) P_{y_{2,n}}(\theta_{S_2})\right]^{\mathbb{1}(y_n=S)}$$

$$\left[(1-P_{y_{1,n}}(\theta_{S_1}))(1-P_{y_{2,n}}(\theta_{S_2}))\right]^{\mathbb{1}(y_n=F)} \prod_{n=N_T+1}^{N} (1-P_{d_{1,n}}(\theta_T)) P_{y_{1,n}}(\theta_{S_1})^{\mathbb{1}(y_n=S)} (1-P_{y_{1,n}}(\theta_{S_1}))^{\mathbb{1}(y_n=F)}. \quad (12)$$

While I cannot observe success for the most recent years, I can modify the likelihood to contain a separate block for recent years that does not contain the success subblocks. I only have to assume that the determinants of success in older years are representative of those in recent years.

After estimating the unrestricted likelihood, I can estimate the restricted likelihood $\mathcal{L}^R$ by substituting in the model's CCPs, which are functions of the success probability $(\theta_{S_1}, \theta_{S_2})$ and structural $(\sigma, \theta_{S_1^*}, \theta_{S_2^*})$ parameters. Because I can observe whether the committee read each applicant's file and whether the committee accepted each applicant, I identify the cutoff probability parameters $(\theta_{S_1^*}, \theta_{S_2^*})$ for each year by how selective the committee is that year.[6] Without loss of generality, every time $d_2 = A$, $\frac{\exp(P(y=S|x,z)/\sigma)}{\exp(P(y=S|x,z)/\sigma)+\exp(P(y_2^*=S|t)/\sigma)}$ enters $\mathcal{L}^R$. This term decreases in $P(y_2^* = S|t)$, so every time an applicant is accepted, it puts downward pressure on $P(y_2^* = S|t)$, considering the objective is to maximize the likelihood. Intuitively, the committee's selectivity identifies the cutoff probability. Also, if the committee engages in sorting, or selecting the applicants with the highest success probabilities $P(y = S|x,z)$, that improves the likelihood. Formally:

Theorem 1. *For simplicity, suppose there is one round. The maximum likelihood estimate $\hat{P}_{y^*} = \hat{P}(y^* = S|t)$ is such that the number of acceptances N_A equals the expected N_A. Also, for any $\hat{P}_{y^*}$ and N_A, $\mathcal{L}^R$ is highest when the committee accepts the N_A highest $P_{y_n} = P(y_n = S|x_n, z_n)$.*

Proof. See Appendix F.

When estimating this model, it is necessary to use public outcomes data to track success for all applicants, including those who did not matriculate at the university. Bai et al. (2022) prove that success parameter estimates will be biased toward zero if the sample is restricted to matriculants only. Intuitively, the reason is that all matriculants must have been accepted, and accepted applicants tend to have high values of those determinants. For example, conditional on matriculating, the relationship between the quantitative GRE score and success is weakened because nearly all matriculants have very high quantitative GRE scores. In addition to tracking success for all applicants, it is also necessary to track which, if any, graduate program each applicant attended. Then controlling for the applicant's expected U.S. News program ranking conditional on x and z relaxes the assumption that success depends only on applicant characteristics. For instance, a rejected applicant who succeeded at some other university might not have succeeded at this one, and so the committee may not have made a mistake by rejecting that applicant.

[6] If selectivity does not vary much across years, $(\theta_{S_1^*}, \theta_{S_2^*})$ can be made intercept only.

I must also handle the problem of characteristics that the committee observes that I do not. For example, the committee may personally know a recommendation letter writer, or the letter may contain information that algorithms cannot detect, such as exactly how rigorous the applicant's masters program was. Failing to address this problem may bias the success parameter estimates and make the model's predictions worse than the committee's. Therefore, including the committee score of each applicant in z to proxy such characteristics is an additional robustness check. Optimization may ultimately require merging human and data-driven approaches. While humans have access to more information and can use their intuition to evaluate less quantifiable characteristics, they may also fall prey to human biases that models do not.

There is also the matter of the likelihood function, in which I assume $\big[(d_1,d_2) \perp\!\!\!\perp y\big]\big|(x,z,t)$. That is, conditional on the independent variables, success is independent of admission. If an applicant is rejected at this university, they can still succeed on the job market by doing well at another university. But if they are rejected at this university, they are also more likely get rejected at all programs ranked at least as high as this university, which may hurt their ability to succeed if job market success is contingent on graduate program ranking. If so, $\hat{\theta}_S$ may be biased upward because without loss of generality, a low GRE score is correlated with not getting into a highly-ranked program, and not getting into such a program can prevent job market success.

Given that, would it help to replace $P\big(y = S\big|x;\theta_S\big)$ with $P\big(y = S\big|d,x;\theta_S\big)$, considering that in general, the joint density that constitutes the likelihood $f(d,y|x)$ equals $f(d|x) \times f(y|d,x)$? There are three reasons why I do not do so. First, if I condition y on x and $d = \mathbb{1}(\text{Accept})$ in the restricted likelihood, I would effectively be regressing d on y on d, which will fail due to perfect collinearity. Second, because d is a function of x, even conditioning y on x and $d = \mathbb{1}(\text{AttendProgram})$ (i.e. whether the applicant attended any university's program) will collapse $\hat{\theta}_S$ toward 0, except for the one parameter $\hat{\theta}_S^d$. For example, in Athey et al. (2007), quantitative GRE scores predict course grades on their own but not when the committee score, which is highly correlated with the acceptance decision, is included. Finally, the committee does not know *ex-ante* whether d will be 1 or 0, so it cannot factor this information into its decision problem.

Once I have estimated the unrestricted and restricted likelihoods, I perform the likelihood-ratio test $\text{LR} = 2\big[\log(\hat{\mathcal{L}}^U) - \log(\hat{\mathcal{L}}^R)\big] \sim \chi^2_Q$, where $Q = \dim(\theta^U) - \dim(\theta^R)$. The purpose of this test is to determine whether the model fits the data well. If it does, then $\hat{\mathcal{L}}^R$ will be close to $\hat{\mathcal{L}}^U$, and the test will not reject.[7] With this test, I can determine whether the committee is

[7]Due to the σ parameter, the restricted likelihood parameters are not nested within the unrestricted ones. While I could use Vuong (1989)'s non-nested specification test, I stick with the likelihood-ratio test, as Anderson, Rosen, Rust, and Wong (2025) do. The likelihood-ratio test is simpler, and there is only one non-nested parameter.

optimizing in the same way the model says it should.[8] If the test rejects, then the committee may be behaving suboptimally. However, the test can still reject even if the committee's deviations from the model do not translate into a meaningfully weaker set of admitted applicants. Therefore, the more practical optimality test is the deviation gains test, which I present in Section 6.

I now present estimation results on simulated data for the year-static model. First, in Tables 15 and 16 of Appendix B, I provide summary statistics for two simulated datasets: one in which the committee is not optimizing and one in which it is. The non-optimizing committee discriminates aginst *Demographic1*, undervalues *AdvMath* (i.e. whether the applicant took advanced math, such as real analysis), and overvalues *ComScore* (i.e. the committee reader's score of the applicant). Then, in Table 4, I present the unrestricted parameter estimates; the simulated data are such that the quantitative GRE score and *AdvMath* both predict success significantly at the 1% level. Finally, Table 5 contains the restricted parameter estimates. As expected, the likelihood ratio test rejects optimality at the 5% level for the suboptimal dataset, but it does not reject optimality for the optimal one. Also as expected, the unrestricted AIC is better than the restricted AIC for the suboptimal dataset, but the restricted AIC is better for the optimal dataset. To make the simulated datasets have the above properties, the data-generating parameters could not exactly equal those implied by the reduced-form results, but I kept them as close as possible.

Although the data are simulated, Table 5 contains an interesting result. There are two years in each of the simulated datasets: one in which the committee transitions about 45% of the applicants to Round 2, and one in which it only transitions about 25%. Essentially, in *YearTransMore*, the committee is more concerned with admitting a strong class than saving time and effort. And in that year, as expected, the Round 1 cutoff probability is lower than the Round 2 cutoff probability That is, it is easier to move on to Round 2 than it ultimately is to get into the program. However, in the other year, which is represented by the intercept term, the Round 1 cutoff is greater than the Round 2 cutoff, meaning the $P(\text{Success}|x)$ of the marginal applicant rejected in Round 1 is greater than the $P(\text{Success}|x,z)$ of the marginal applicant accepted into the program.

How can this be? Sometimes when a committee transitions a seemingly strong applicant to Round 2, the applicant's z reveals they are weaker than previously thought, e.g. a bad recommendation letter. If the committee transitions too few applicants to Round 2, it can end up in an unfortunate situation where it threw out so many files that it now has to accept those weaker applicants. As a simple example, suppose there are three applicants: A, B, and C. The committee has time to read two applicants' files and space to accept one applicant. Suppose that overall, A is the best applicant and C is the worst, but A is the worst in objective variables. Then the committee will reject A in Round 1 and accept B, who is actually worse than A. In the simulated

[8] A one-round version of the year-static model with $d = \mathbb{1}(\text{AttendProgram})$ can test if the model fits the profession's behavior, as opposed to a single program's behavior. Although a one-round model as the main specification is cheating, it is necessary for this test because the transition decision is unobservable for other programs. Also, with a one-round model, I can observe how much better at predicting success a one-round process, in which the committee reads every file, would be. Perhaps in some settings it could be worth the additional time cost.

data, the minimum percentage of applications the committee must read in Round 2 to avoid this situation is between 25%, where the situation occurs, and 45%, where it does not. In a later version of the paper, I will determine whether this situation occurs in any of the actual years, and if so, the minimum percentage of files the committee must read to avoid this situation.

5.3 Year-Dynamic Model

I now present the year-dynamic model. Relative to the year-static model, this model is more computationally intensive, especially if the dimensions of x and z are large. This model also has the downside that the Round 1 cutoff cannot be interpreted as a probability because its value can exceed 1, which is why I denote the cutoff with L for level instead of P for probability. However, the year-dynamic model has two advantages. First, while its parameter estimates may not be entirely causal, there is no omitted variable bias of the kind that occurs when estimating $P(y=S|x)$. Second, for any multiyear model, such as those in Appendix E, it is necessary for the Round 1 value function to be dynamically linked to the Round 2 value function.

The year-dynamic model's Round 2 value function is the same as the year-static's:

$$v_2(x,z,t,\varepsilon_2) = \max_{d_2\in\{A,R\}} u_2(x,z,t,d_2)+\varepsilon_2(d_2), \text{ such that:} \tag{13}$$

$$u_2(x,z,t,d_2) = P(y=S|x,z)\mathbb{1}(d_2=A)+P(y_2^*=S|t)\mathbb{1}(d_2=R).$$

However, in Round 1, if the committee transitions an applicant, it gets $\mathbb{E}[v_2(x,z,t,\varepsilon_2)|x]$, which is the expected value of Round 2, where z is integrated out conditional on x:

$$v_1(x,t,\varepsilon_1) = \max_{d_1\in\{T,R\}} u_1(x,t,d_1)+\varepsilon_1(d_1), \text{ such that:} \tag{14}$$

$$u_1(x,t,d_1) = \mathbb{E}[v_2(x,z,t,\varepsilon_2)|x,t]\mathbb{1}(d_1=T)+L(y_1^*=S|t)\mathbb{1}(d_1=R)^*, \text{ such that:}$$

$$\mathbb{E}[v_2(x,z,t,\varepsilon_2)|x,t] = \int_{\widetilde{z}}\sigma\log\left(\sum_{\widetilde{d}_2\in\{A,R\}} e^{u_2(x,\widetilde{z},t,\widetilde{d}_2)/\sigma}\right)dF(\widetilde{z}|x).$$

The restricted CCPs are the same as in the year-static model:

$$P(d_r|x,(z),t) = \frac{\exp(u_r(x,(z),t,d_r)/\sigma)}{\sum_{\widetilde{d}_r\in D_r}\exp(u_r(x,(z),t,\widetilde{d}_r)/\sigma)} \tag{15}$$

Rather than write out the logit specifications for the unrestricted CCPs, success probabilities, and cutoff probabilities, I reference Equations (9)–(11). That said, there are two differences from the year-static framework. First, P_{y_1} is not part of the year-dynamic likelihood, and second, $L_{y_1^*} = L(y_1^*=S|t;\theta_{S_1^*}) = \exp(f_{S_1^*}(t)\theta_{S_1^*})$, since the cutoff level can be greater than 1. In addition, letting F_j be the conditional CDF for each z_j, I parameterize the integral in Equation (14) as:

$$\mathbb{E}[v_2(x,z,t,\varepsilon_2)|x,t] = \int_{\widetilde{z}_1}\cdots\int_{\widetilde{z}_J}\left[\cdot\right]dF_1(\widetilde{z}_1|\widetilde{z}_2,\ldots\widetilde{z}_J,x)\cdots dF_J(\widetilde{z}_J|x). \tag{16}$$

Table 4: Year-Static Unrestricted Estimates

(a) Suboptimal Committee

	Transition		Accept		Success1		Success2	
	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-8.5777***	-1.8014***	-4.7849***	-0.9157***	-10.5425***	-1.9666***	-8.3580***	-1.7285***
	(0.9369)	(0.1610)	(1.6496)	(0.2990)	(1.0566)	(0.1529)	(1.7683)	(0.3127)
GREQuantPct	0.0919***	0.0193***	0.0401**	0.0077**	0.1140***	0.0213***	0.0839***	0.0173***
	(0.0110)	(0.0020)	(0.0193)	(0.0036)	(0.0125)	(0.0019)	(0.0210)	(0.0039)
Gender	0.1789	0.0376	0.3374	0.0646	0.0044	0.0008	-0.0594	-0.0123
	(0.1417)	(0.0297)	(0.2315)	(0.0442)	(0.1509)	(0.0281)	(0.2300)	(0.0476)
Demographic1	-0.4843**	-0.1017**	-0.5416	-0.1036	-0.2944	-0.0549	-0.2355	-0.0487
	(0.1988)	(0.0412)	(0.3380)	(0.0639)	(0.2115)	(0.0392)	(0.3420)	(0.0701)
Demographic2	0.0338	0.0071	-0.0696	-0.0133	0.2505	0.0467	0.5612*	0.1161*
	(0.2101)	(0.0441)	(0.3531)	(0.0676)	(0.2214)	(0.0413)	(0.3209)	(0.0661)
AdvMath			0.4165	0.0797*			0.6465***	0.1337***
			(0.2555)	(0.0482)			(0.2502)	(0.0504)
ComScore			0.2219***	0.0425***			0.0059	0.0012
			(0.0668)	(0.0120)			(0.0583)	(0.0121)
YearTransMore	0.8980***	0.1886***	-1.2106***	-0.2317***				
	(0.1416)	(0.0273)	(0.2388)	(0.0397)				
Obs | Pseudo-R^2	1000	0.0988						
LL | Param | AIC	-1616.1	26	3284.3					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	Transition		Accept		Success1		Success2	
	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-9.1577***	-1.9164***	-9.2245***	-1.6855***	-10.5425***	-1.9666***	-8.4374***	-1.7657***
	(0.9469)	(0.1576)	(1.8167)	(0.2808)	(1.0566)	(0.1529)	(1.7772)	(0.3171)
GREQuantPct	0.0992***	0.0208***	0.0971***	0.0177***	0.1140***	0.0213***	0.0847***	0.0177***
	(0.0111)	(0.0019)	(0.0207)	(0.0033)	(0.0125)	(0.0019)	(0.0210)	(0.0039)
Gender	0.0002	0.0000	0.1897	0.0347	0.0044	0.0008	-0.0412	-0.0086
	(0.1422)	(0.0298)	(0.2448)	(0.0448)	(0.1509)	(0.0281)	(0.2299)	(0.0481)
Demographic1	-0.3518*	-0.0736*	-0.4248	-0.0776	-0.2944	-0.0549	-0.2785	-0.0583
	(0.2004)	(0.0416)	(0.3637)	(0.0662)	(0.2115)	(0.0392)	(0.3423)	(0.0709)
Demographic2	0.0435	0.0091	-0.1174	-0.0215	0.2505	0.0467	0.5504*	0.1152*
	(0.2137)	(0.0447)	(0.3788)	(0.0693)	(0.2214)	(0.0413)	(0.3256)	(0.0678)
AdvMath			0.7137***	0.1304***			0.5665**	0.1185**
			(0.2626)	(0.0463)			(0.2501)	(0.0513)
ComScore			0.0952	0.0174			0.0237	0.0050
			(0.0628)	(0.0113)			(0.0572)	(0.0119)
YearTransMore	0.8248***	0.1726***	-1.2833***	-0.2345***				
	(0.1419)	(0.0276)	(0.2467)	(0.0378)				
Obs | Pseudo-R^2	1000	0.1058						
LL | Param | AIC	-1611.7	26	3275.4					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 5: Year-Static Restricted Estimates

(a) Suboptimal Committee

	Cutoff1	$P(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success1	Success2	Sigma
Intercept	0.1539	53.84%	-0.5306**	37.04%	-10.3327***	-7.2584***	
	(0.2749)	(40.49%, 66.66%)	(0.2154)	(27.83%, 47.29%)	(1.9333)	(2.2514)	
GREQuantPct					0.1123***	0.0663**	
					(0.0227)	(0.0260)	
Gender					0.1010	0.0839	
					(0.1144)	(0.1650)	
Demographic1					-0.4233**	-0.3725	
					(0.1783)	(0.2547)	
Demographic2					0.1588	0.2935	
					(0.1773)	(0.2525)	
AdvMath						0.5525***	
						(0.1957)	
ComScore						0.1161***	
						(0.0447)	
YearTransMore	-0.9253***	31.62%	1.2103***	66.37%			
	(0.2622)	(26.62%, 37.07%)	(0.4178)	(47.08%, 81.40%)			
Sigma							0.2473***
							(0.0680)
Observations	1000						
LL | Param | AIC	-1625.8	17	3285.6				
LR Stat | DF | P-Val	19.2965	9	0.0228				

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	Cutoff1	$P(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success1	Success2	Sigma
Intercept	-0.0094	49.77%	-0.5492***	36.60%	-10.0587***	-8.9338***	
	(0.2357)	(38.43%, 61.13%)	(0.1882)	(28.53%, 45.50%)	(2.2082)	(2.4363)	
GREQuantPct					0.1090***	0.0882***	
					(0.0259)	(0.0275)	
Gender					0.0070	0.0220	
					(0.1080)	(0.1620)	
Demographic1					-0.3346*	-0.2945	
					(0.1803)	(0.2626)	
Demographic2					0.1444	0.2881	
					(0.1702)	(0.2608)	
AdvMath						0.6055***	
						(0.1854)	
ComScore						0.0657	
						(0.0419)	
YearTransMore	-0.7509***	31.86%	1.1488***	64.56%			
	(0.2200)	(27.27%, 36.83%)	(0.3847)	(47.28%, 78.72%)			
Sigma							0.2149***
							(0.0597)
Observations	1000						
LL | Param | AIC	-1615.9	17	3265.9				
LR Stat | DF | P-Val	8.4729	9	0.4873				

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Depending on if each z_j is continuous, binary, or multinomial, I use one of the following PDFs: $f_{z|x}(z_j|x;\beta_{F_{z|x,j}},\sigma^2_{F_{z|x,j}}) = \frac{1}{\sqrt{2\pi\sigma^2_{F_{z|x,j}}}}\exp\left[\frac{-1}{2\sigma^2_{F_{z|x,j}}}(z_j - g(x)\beta_{F_{z|x,j}})^2\right]$, $P(z_j = 1|x;\theta_{F_{z|x,j}}) = \frac{\exp\left(g(x)\theta_{F_{z|x,j}}\right)}{1+\exp\left(g(x)\theta_{F_{z|x,j}}\right)}$,

or $P(z_j = (k \neq K)|x;\theta_{F^k_{z|x,j}}) = \frac{\exp\left(g(x)\theta_{F^k_{z|x,j}}\right)}{1+\sum_{\ell=1}^{K-1}\exp\left(g(x)\theta_{F^\ell_{z|x,j}}\right)}$, respectively.

The unrestricted likelihood is then:

$$\mathcal{L}_N^U(\theta) = \prod_{n=1}^{N_T} f(z_n|x_n;\theta_F)P_{d_{1,n}}(\theta_T)P_{d_{2,n}}(\theta_A)^{\mathbb{1}(d_{2,n}=A)}(1-P_{d_{2,n}}(\theta_A))^{\mathbb{1}(d_{2,n}=R)}$$
$$P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)}\prod_{n=N_T+1}^{N} f(z_n|x_n;\theta_F)(1-P_{d_{1,n}}(\theta_T)). \quad (17)$$

As before, I estimate the restricted likelihood by substituting the restricted CCPs for the unrestricted ones. To save computation time, I use (θ_F^U,θ_S^U) as an initial guess of (θ_F^R,θ_S^R).

There is an added complication with my dataset, namely that it does not contain subjective variables for the $N-N_T$ of N applicants who were rejected in Round 1. This omission is a problem when estimating $f(z_n|x_n;\theta_F)$ in the likelihood. Because I am unable to obtain z for everyone, I must instead estimate the following Heckman-corrected full-information likelihood:

$$\mathcal{L}_N^U(\theta) = \prod_{n=1}^{N_T} \Phi(x_n,t_n;\theta_H) f\left(z_n|x_n,\frac{\phi(x_n,t_n;\theta_H)}{\Phi(x_n,t_n;\theta_H)};\theta_F\right)P_{d_{1,n}}(\theta_T)P_{d_{2,n}}(\theta_A)^{\mathbb{1}(d_{2,n}=A)}(1-P_{d_{2,n}}(\theta_A))^{\mathbb{1}(d_{2,n}=R)}$$
$$P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)}\prod_{n=N_T+1}^{N}(1-\Phi(x_n,t_n;\theta_H))(1-P_{d_{1,n}}(\theta_T)). \quad (18)$$

The binary probit "first stage" is $\Phi(x_n,t_n;\theta_H)$ or $1-\Phi(x_n,t_n;\theta_H)$, and the inverse Mills ratio is $\frac{\phi(x_n,t_n;\theta_H)}{\Phi(x_n,t_n;\theta_H)}$. In addition, I parameterize as probit the densities of any $z_j \in z$ that were originally logit. The Heckman instrument here is t, since in any given year, the committee may transition more or fewer applications for reasons that have nothing to do with the applicant's z.

I estimate the year-dynamic model on the same simulated data as the year-static model. The unrestricted and restricted estimates are in Tables 17 and 18 of Appendix C, respectively. Since I have two z variables, namely the advanced math dummy and the applicant's committee score, I have two PDFs to estimate. Committee score is a continuous z, and since I parameterize continuous z's with the normal distribution, I have an additional σ_F parameter that equals the square root of the distribution's mean squared error. I also let committee score depend on advanced math, so I can estimate the joint distriubiton of the two z's without assuming the z's are independent.

Unlike before, the likelihood-ratio test rejects the year-dynamic model even for the optimizing committee. The reason is that when I generated the simulated data, I had the committee optimize in a year-static manner. That is, I did not explicitly form an expectation of Round 2 by considering how x and z are related to each other. However, even though the data reject the year-dynamic model, I show in Section 6 that by following the year-dynamic model, the committee would admit a class nearly indistinguishable from the year-static model. Tables 19 and 20 of Appendix C contain Heckman-corrected estimates for the case where I deleted all z's in the simulated data for applicants who do not survive Round 1. The inverse Mills ratio is never statistically significant, meaning that in these simulated datasets, even though the missing z's are not missing at random, there is no significant selection bias in estimating the distribution of $z|x$.

Finally, in Tables 21–24 of Appendix C, I present year-myopic estimates. The year-myopic model is a special case of the year-dynamic model in which all parameters of $f_{z|x}$ except intercept and variance parameters are restricted to equal 0, effectively making the distribution of z unconditional. In this model, if there is an objective variable that is positively associated with z in addition to being associated with success in its own right, the committee will not put any extra weight on that variable in Round 1. Therefore, unlike the year-static model, the year-myopic model will necessarily be less effective at admitting applicants than the year-dynamic model.

5.4 Waitlist Models

An admissions committee does not send out all its acceptances at once, but rather in batches. It starts by accepting its most preferred applicants, and then it puts the others who survived Round 1 on a waitlist. Then provided not all its early acceptances matriculate, it accepts applicants off the waitlist and ultimately moves down the waitlist until it has filled its cohort. I model this process as concisely as possible by increasing the number of rounds to three. As shown in Figure 6, the committee observes only the objective variables in Round 1 and either transitions the applicant to Round 2 or rejects them out of courtesy. But in Round 2, instead of accepting or rejecting the survivors based on the objective and subjective variables, it accepts or waitlists them.

Then in Round 3, it accepts or rejects the waitlisted applicants. Between Rounds 2 and 3, it observes how many of the early acceptances matriculate and makes its Round 3 decisions accordingly. If more early acceptances matriculate, it accepts fewer applicants off the waitlist because there are fewer openings left to fill. Intuitively, this situation is better than the other way around. More early acceptances matriculating means the committee gets more of its top choices. In contrast, if the committee has to accept many applicants off the waitlist, it may have to lower its standards because applicants it could have gotten with early offers may have already committed to other schools. Therefore, it should err on the side of giving Round 2 offers to applicants who are relatively likely to matriculate. I could simply have the committee value applicants by a combination of their success and matriculation probabilities, but such a model could only direct the committee to give early offers to likely matriculants if the committee in the data were already doing so. Otherwise, the estimated weight on the matriculation probability would be zero.

Thus, I model this dynamic by letting the Round 3 cutoff probability be $P(y_3^* = S|\overline{m_{2A}}, n_3, \overline{q_3}, \overline{m_3})$

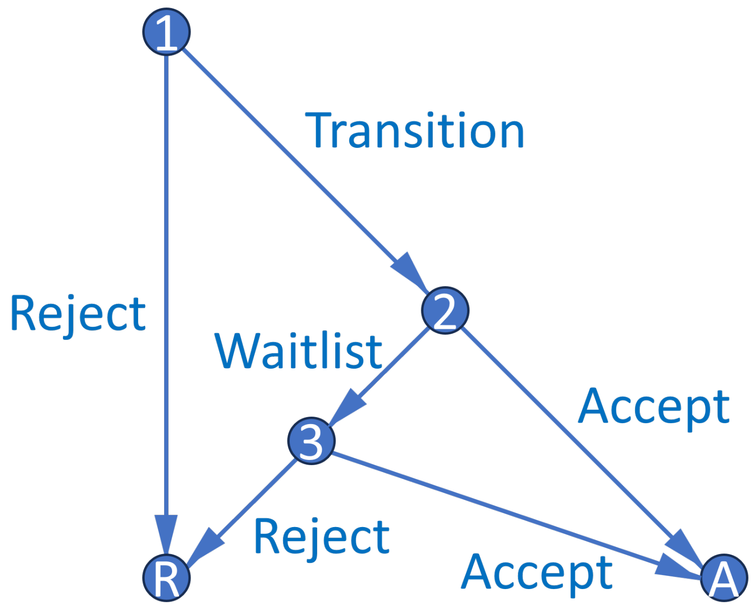


Figure 3: Waitlist Decision Graph

instead of $P(y_3^* = S|t)$. Here, $\overline{m_{2A}}$ is the average early accepted applicant's matriculation score[9] in a given year. A matriculation score represents how likely the committee thinks the applicant is to matriculate if accepted, and it can be estimated if necessary by regressing matriculation on x and z for the sample of accepted applicants. n_3 is that year's number of waitlisted applicants, $\overline{q_3}$ is that year's average waitlisted applicant quality, and $\overline{m_3}$ is that year's average waitlisted applicant matriculation score. $P(\text{Cutoff})_3$ increases in all four; in particular, it increases in $\overline{m_{2A}}$ because a larger $\overline{m_{2A}}$ means there will be fewer openings for waitlisted applicants. To identify parameters for all those variables, I would need at least five years of data, but the only variable necessary to make the dynamics work is $\overline{m_{2A}}$. Therefore, when I estimate the waitlist-dynamic model on the usual two years worth of simulated data later in this section, I include only $\overline{m_{2A}}$.

Given that, what information available at the start of Round 2 determines $\overline{m_{2A}}$? For all applicants in each year, I define $\ddot{m}_2$ as $\overline{m_2}$ but with m_2 excluded, where m_2 is the applicant's own matriculation score. Within each year, a low m_2 necessarily implies a high $\ddot{m}_2$, since $\overline{m_2}$ is a within-year average. I then estimate $f_{\overline{m_{2A}}|\ddot{m}_2}$ on all transitioned applicants; the parameter estimate should be positive provided there exists variation in $\overline{m_2}$ across years. For example, by decreasing acceptance probabilities, Covid may have made applicants more likely to matriculate if accepted. For simplicity, I suppose the simulated committee provides its own matriculation scores, but with the actual data, I will estimate the parameters of m_2, and therefore $\ddot{m}_2$, in a first-stage likelihood. I can also estimate the parameters of $\overline{m_1}$ in a first-stage likelihood, where m_1 is a matriculation

[9] A matriculation score can but does not have to be a probability. For example, it can be a number from 1 to 10.

score based on x and not z, and use it for Round 1's distribution of $\overline{m_2}$ conditional on $\overline{m_1}$.

The model implies that the optimal dynamic decision in Round 2 may not be the same as the optimal myopic one. Specifically, a committee may waitlist, as opposed to accept, an applicant with a low m_2 even if their $P(\text{Success})$ is higher than an applicant with a higher m_2. The reason is that compared to a high m_2 applicant, a low m_2 applicant has a relatively high $\ddot{m}_2$. That implies a high $\overline{F}_{\overline{m_{2A}}|\ddot{m}_2}$, which in turn implies a high Round 3 cutoff probability, Round 3 value function, and payoff for waitlisting the low m_2 applicant in Round 2. Perhaps I will find that in the actual data, the committee does not think dynamically; that is, removing these "memory" parameters increases the AIC. However, if $\ddot{m}_2$ affects $\overline{F}_{\overline{m_{2A}}|\ddot{m}_2}$, which in turn affects $\mathbb{E}[P(\text{Cutoff})_3]$, perhaps the committee should. Note that success data for recent years, which does not yet exist, is unnecessary to identify these two effects because admission data has all the necessary information.

The upshot of this model is clear: accepting applicants who want to come benefits the program. I also discount the expected Round 3 value with $\beta \in (0,1)$ so that the payoff from accepting an applicant in Round 3 is less than that from accepting them in Round 2. The discount factor effectively captures the chance of losing them to another school. In theory, a committee could just give early offers to both a top applicant with a low m_2 and a weaker applicant with a higher m_2. But it has to be careful because it does not want to overshoot its target cohort size,[10] and it may also face a capacity constraint if it invites accepted applicants to an open house. I now present the value functions for the waitlist-dynamic model. The full equations are in Appendix D, but for simplicity here, I include only the necessary state variables to make the dynamics work. That is, I write $P(y_3^* = S|\overline{m_{2A}})$ instead of $P(y_3^* = S|\overline{m_{2A}}, n_3, \overline{q_3}, \overline{m_3})$. Starting with Round 3:

$$v_3(x,z,\overline{m_{2A}},\varepsilon_3) = \max_{d_3 \in \{A,R\}} u_3(x,z,\overline{m_{2A}},d_3) + \varepsilon_3(d_3), \text{ such that:} \tag{19}$$

$$u_3(x,z,\overline{m_{2A}},d_3) = P(y = S|x,z)\mathbb{1}(d_3 = A) + P(y_3^* = S|\overline{m_{2A}})\mathbb{1}(d_3 = R).$$

The Round 2 value function is:

$$v_2(x,z,\ddot{m}_2,\varepsilon_2) = \max_{d_2 \in \{A,W\}} u_2(x,z,\ddot{m}_2,d_2) + \varepsilon_2(d_2), \text{ such that:} \tag{20}$$

$$u_2(x,z,\ddot{m}_2,d_2) = P(y = S|x,z)\mathbb{1}(d_2 = A) + \beta\mathbb{E}[v_3(x,z,\overline{m_{2A}},\varepsilon_3)|x,z,\ddot{m}_2]\mathbb{1}(d_2 = W), \text{ such that:}$$

$$\mathbb{E}[v_3(x,z,\overline{m_{2A}},\varepsilon_3)|x,z,\ddot{m}_2] = \int_{\widetilde{\overline{m_{2A}}}} \sigma \log\left(\sum_{\tilde{d}_3 \in \{A,R\}} e^{u_3(x,z,\widetilde{\overline{m_{2A}}},\tilde{d}_3)/\sigma}\right) dF_{\overline{m_{2A}}|\ddot{m}_2}(\widetilde{\overline{m_{2A}}}|\ddot{m}_2).$$

The Round 1 value function depends on whether I include a conditional expectation, as in the year-dynamic model, or whether I do not, as in the year-static model. I refer to the three-round

[10]To model a committee that explicitly avoids over and undershooting, I could add the committee's target cohort size minus the number of early matriculants as a covariate to Round 3's cutoff probability. The parameter estimate should be positive because fewer spots open means a more selective committee. Therefore, the committee will accept more applicants when the number of early matriculants is small relative to its target cohort size. However, the parameter may be unidentifiable, especially if estimated alongside $\overline{m_{2A}}$'s parameter, and so have to be calibrated.

model without a Round 1 conditional expectation as the waitlist-hybrid model, since it has a static Round 1 and dynamic Round 2. Its Round 1 value function is:

$$v_1(x,t,\varepsilon_1) = \max_{d_1 \in \{T,R\}} u_1(x,t,d_1) + \varepsilon_1(d_1), \text{ such that:} \tag{21}$$

$$u_1(x,t,d_1) = P(y=S|x)\mathbb{1}(d_1=T) + P(y_1^*=S|t)\mathbb{1}(d_1=R).$$

In contrast, the utility function in the waitlist-dynamic's Round 1 value function is:

$$u_1(x,t,d_1) = \mathbb{E}[v_2(x,z,\ddot{m}_2,\varepsilon_2)|x]\mathbb{1}(d_1=T) + L(y_1^*=S|t)\mathbb{1}(d_1=R), \text{ such that:} \tag{22}$$

$$\mathbb{E}[v_2(x,z,\ddot{m}_2,\varepsilon_2)|x] = \int_{\widetilde{z}}\int_{\widetilde{\ddot{m}_2}} \sigma \log\left(\sum_{\widetilde{d}_2 \in \{A,R\}} e^{u_2(x,\widetilde{z},\widetilde{\ddot{m}_2},\widetilde{d}_2)/\sigma}\right) dF_{\ddot{m}_2|x,z}(\widetilde{\ddot{m}_2}|x,\widetilde{z})dF_{z|x}(\widetilde{z}|x).$$

Expressed as simply as possible, the restricted CCPs are:

$$P(d_r|\text{states}) = \frac{\exp(u_r(\text{states},d_r)/\sigma)}{\sum_{\widetilde{d}_r \in D_r}\exp(u_r(\text{states},\widetilde{d}_r)/\sigma)}. \tag{23}$$

The waitlist-dynamic model's unrestricted CCPs, success probabilities,, and cutoff probabilities are $P_{d_1} = P\big(d_1=T\big|x,t;\theta_T\big)$, $P_{d_2} = P\big(d_2=A\big|x,z,\ddot{m}_2;\theta_{A_2}\big)$, $P_{d_3} = P\big(d_3=A\big|x,z,\overline{m_{2A}};\theta_{A_3}\big)$, $P_y = P\big(y=S\big|x,z;\theta_S\big)$, $L_{y_1^*} = L\big(y_1^*=S\big|t;\theta_{S_1^*}\big)$, and $P_{y_3^*} = P\big(y_3^*=S\big|\overline{m_{2A}};\theta_{S_3^*}\big)$. These probabilities have the same logit specifications as in the year-dynamic model (see Equations (9)–(11) for the functional forms).[11] Let applicants 1 to N_3 be waitlisted, N_3+1 to N_2 accepted in Round 2, and N_2+1 to $N_1=N$ rejected in Round 1. Then the unrestricted waitlist-dynamic likelihood is:

$$\begin{aligned}
\mathcal{L}_{N_1}^U(\theta) = &\prod_{n=1}^{N_3} f_{z|x}(z_n|x_n;\theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n,z_n;\theta_{F_{\ddot{m}_2|x,z}}) f_{\overline{m_{2A}}|\ddot{m}_2}(\overline{m_{2A,n}}|\ddot{m}_{2,n};\theta_{F_{\overline{m_{2A}}|\ddot{m}_2}}) \\
&P_{d_{1,n}}(\theta_T)(1-P_{d_{2,n}}(\theta_{A_2}))P_{d_{3,n}}(\theta_{A_3})^{\mathbb{1}(d_{3,n}=A)}(1-P_{d_{3,n}}(\theta_{A_3}))^{\mathbb{1}(d_{3,n}=R)}P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)} \\
&\prod_{n=N_3+1}^{N_2} f_{z|x}(z_n|x_n;\theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n,z_n;\theta_{F_{\ddot{m}_2|x,z}}) f_{\overline{m_{2A}}|\ddot{m}_2}(\overline{m_{2A,n}}|\ddot{m}_{2,n};\theta_{F_{\overline{m_{2A}}|\ddot{m}_2}}) \\
&P_{d_{1,n}}(\theta_T)P_{d_{2,n}}(\theta_{A_2})P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)} \\
&\prod_{n=N_2+1}^{N_1} f_{z|x}(z_n|x_n;\theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n,z_n;\theta_{F_{\ddot{m}_2|x,z}})(1-P_{d_{1,n}}(\theta_T)).
\end{aligned} \tag{24}$$

As with the year-dynamic likelihood, I use (θ_F^U,θ_S^U) as an initial guess of (θ_F^R,θ_S^R) for the restricted likelihood. Also, Heckman corrections are necessary for $f_{z|x}$ and $f_{\ddot{m}_2|x,z}$ due to the missing z's for the applicants rejected in Round 1. To show that the dynamics of the waitlist models work as intended before proceeding with estimation, I prove the following theorem:

[11] $L_{y_1^*}$ is parameterized as exponential because it can exceed 1. However, in the waitlist-hybrid model, it is denoted as $P_{y_1^*}$ and parameterized as logit. The waitlist-hybrid model also includes $P_{y_1}\big(y=S\big|x;\theta_{S_1}\big)$.

Theorem 2. *Success probability equal, an applicant's Round 2 acceptance probability is strictly increasing in their matriculation score.*

Proof. See Appendix F.

To estimate the waitlist-hybrid and dynamic models, I simulate different datasets from the ones used for the year-static and dynamic models. Again, for simplicity, I use $P(y_3^* = S|\overline{m_{2A}})$ instead of $P(y_3^* = S|\overline{m_{2A}}, n_3, \overline{q_3}, \overline{m_3})$, which simplifies the likelihood and reduces the number of parameters. But I still must generate a dummy variable that equals 1 if an applicant is waitlisted, a dummy that equals 1 if the applicant matriculates, and a matriculation score variable that is a function of each applicant's matriculation probability. Summary statistics and correlation matrices for these datasets are in Tables 25 and 26 of Appendix C. As before, the committee optimizes in one of the datasets and does not in the other. Also, in *YearReadMat*, the committee transitions more of the applicants, and the applicants have higher matriculation scores on average.

The unrestricted waitlist-hybrid estimates are in Table 6. As expected, the $\ddot{m}_2$ parameter in the F$\overline{m_{2A}}$ block is positive and significant at the 1% level in both datasets. In addition, the $\ddot{m}_2$ parameter in the *Accept2* block is negative and in the non-optimizing dataset significant at the 5% level, which is also as expected. The reason is that the committee is more likely to give an early offer to an applicant with a high m_2 and therefore a low $\ddot{m}_2$. Finally, the $\overline{m_{2A}}$ parameter is negative and significant at the 1% level because a high $\overline{m_{2A}}$ means fewer spots will be open in Round 3 due to more Round 2 matriculants, and so it will be harder to get accepted off the waitlist.

Table 7 contains restricted estimates for the waitlist-hybrid model. In the dataset where the committee is not optimizing, the likelihood-ratio test rejects the model with $p < 0.0001$, whereas in the other dataset, the test does not reject with $p = 0.6465$. As with the year-static model, in the year that is not *YearReadMat*, the Round 1 cutoff probability is higher than the Round 3 cutoff probability. As such, the marginal applicant rejected in Round 1's $P(y = S|x)$ is greater than than the marginal applicant accepted into the program's $P(y = S|x,z)$. Moreover, in *YearReadMat*, the Round 3 cutoff probability is higher not only because there are more applicants who survive Round 1, but also because those applicants have higher matriculation scores on average.

Meanwhile, in Tables 27 and 28 of Appendix C, I present estimates for the waitlist-dynamic model. As with the year-dynamic model, the likelihood-ratio test rejects this model for both datasets, though as expected, the test statistic is higher for the suboptimal dataset. Finally, Tables 29 and 30 of Appendix C contain Heckman-corrected estimates for the case where all the z's, as well as the matriculation scores, are missing for the applicants who do not survive Round 1. As with the year-dynamic model estimated on data with missing z's, the inverse Mills ratios are never statistically significant, so I cannot reject the null hypothesis of no selection bias.

Finally, in Appendix E, I present two different infinite-horizon multiyear models. A limitation of the waitlist models is that while they give early acceptances to applicants with higher matriuclation scores all else equal, matriculation scores have no bearing on the probability of a late acceptance. For this reason, the first multiyear model is constructed so that an applicant's overall acceptance probability is increasing in their matriculation score, the logic being that more

Table 6: Waitlist-Hybrid Unrestricted Estimates

(a) Suboptimal Committee

	$F\overline{m_{2A}}$	Transition		Accept2		Accept3		Success1		Success3	
	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	0.1489***	-10.7143***	-2.1065***	1.4111	0.1810	5.2706**	0.6876**	-7.5197***	-1.6198***	-7.8305***	-1.6644***
	(0.0018)	(1.0455)	(0.1549)	(2.3121)	(0.2948)	(2.3583)	(0.3035)	(0.9558)	(0.1909)	(1.7348)	(0.3277)
GREQuantPct		0.1111***	0.0218***	-0.0217	-0.0028	0.0304	0.0040	0.0810***	0.0175***	0.0802***	0.0171***
		(0.0119)	(0.0019)	(0.0253)	(0.0032)	(0.0266)	(0.0034)	(0.0113)	(0.0023)	(0.0206)	(0.0040)
Gender		0.3463**	0.0681**	0.3925	0.0503	-0.7569**	-0.0987**	-0.2071	-0.0446	-0.0143	-0.0030
		(0.1473)	(0.0286)	(0.2997)	(0.0386)	(0.3458)	(0.0441)	(0.1473)	(0.0314)	(0.2373)	(0.0504)
Demographic1		-0.2186	-0.0430	-0.1530	-0.0196	-1.1218**	-0.1463**	-0.2002	-0.0431	-0.1192	-0.0253
		(0.2038)	(0.0400)	(0.4589)	(0.0590)	(0.5553)	(0.0709)	(0.1986)	(0.0427)	(0.3732)	(0.0792)
Demographic2		0.4335**	0.0852**	0.2764	0.0355	-1.0364**	-0.1352**	-0.1484	-0.0320	-0.2127	-0.0452
		(0.2208)	(0.0431)	(0.4718)	(0.0604)	(0.5085)	(0.0661)	(0.2220)	(0.0477)	(0.3782)	(0.0803)
AdvMath				0.4911	0.0630	0.0415	0.0054			0.1456	0.0309
				(0.3246)	(0.0415)	(0.3585)	(0.0467)			(0.2555)	(0.0543)
ComScore				0.0047	0.0006	0.1279	0.0167			0.0319	0.0068
				(0.0677)	(0.0087)	(0.0788)	(0.0102)			(0.0577)	(0.0122)
YearTransMat		1.0063***	0.1979***								
		(0.1492)	(0.0265)								
$\ddot{m}_2$	1.0524***			-0.3346**	-0.0429**						
	(0.0003)			(0.1516)	(0.0192)						
$\overline{m_{2A}}$						-1.5313***	-0.1998***				
						(0.1683)	(0.0074)				
$\sqrt{\text{MeanSqError}}$	0.0041***										
	(0.0002)										
Obs | Pseudo-R^2	1000	0.9309									
LL | Param | AIC	-165.4	37	404.9								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	$F\overline{m_{2A}}$	Transition		Accept2		Accept3		Success1		Success3	
	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-1.2247***	-8.0729***	-1.6884***	-7.7449***	-0.7422***	-0.1816	-0.0241	-10.5573***	-2.1618***	-9.8892***	-1.9466***
	(0.0038)	(0.9326)	(0.1635)	(2.5864)	(0.2352)	(2.0599)	(0.2733)	(1.0083)	(0.1698)	(1.7557)	(0.2871)
GREQuantPct		0.0848***	0.0177***	0.0690**	0.0066**	0.0617**	0.0082**	0.1188***	0.0243***	0.0974***	0.0192***
		(0.0109)	(0.0020)	(0.0313)	(0.0029)	(0.0266)	(0.0034)	(0.0119)	(0.0020)	(0.0203)	(0.0035)
Gender		0.0593	0.0124	-0.5311	-0.0509	-0.1300	-0.0173	-0.0708	-0.0145	-0.0561	-0.0111
		(0.1415)	(0.0296)	(0.3773)	(0.0363)	(0.3031)	(0.0403)	(0.1494)	(0.0306)	(0.2465)	(0.0486)
Demographic1		-0.2873	-0.0601	0.2456	0.0235	0.4599	0.0610	-0.7173***	-0.1469***	-0.4344	-0.0855
		(0.1955)	(0.0407)	(0.6410)	(0.0617)	(0.4960)	(0.0653)	(0.2061)	(0.0414)	(0.3668)	(0.0718)
Demographic2		-0.0791	-0.0165	0.4982	0.0477	0.1980	0.0263	-0.0743	-0.0152	0.5745	0.1131
		(0.2150)	(0.0450)	(0.6312)	(0.0606)	(0.5096)	(0.0675)	(0.2199)	(0.0450)	(0.3821)	(0.0743)
AdvMath				0.2702	0.0259	-0.0617	-0.0082			-0.1088	-0.0214
				(0.3837)	(0.0367)	(0.3646)	(0.0484)			(0.2639)	(0.0520)
ComScore				-0.0175	-0.0017	0.0656	0.0087			0.1487**	0.0293**
				(0.0989)	(0.0095)	(0.0890)	(0.0118)			(0.0684)	(0.0133)
YearTransMat		0.8006***	0.1674***								
		(0.1410)	(0.0277)								
$\ddot{m}_2$	1.4277***			-0.1255	-0.0120						
	(0.0007)			(0.1679)	(0.0159)						
$\overline{m_{2A}}$						-1.1005***	-0.1461***				
						(0.1176)	(0.0034)				
$\sqrt{\text{MeanSqError}}$	0.0057***										
	(0.0005)										
Obs | Pseudo-R^2	1000	0.8919									
LL | Param | AIC	-268.3	37	610.6								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 7: Waitlist-Hybrid Restricted Estimates

(a) Suboptimal Committee

	F$\overline{m_{2A}}$	Cutoff1	$P(\text{Cutoff})_1$	Cutoff3	$P(\text{Cutoff})_3$	Success1	Success3	Beta	Sigma
Intercept	0.1489***	-0.7116***	32.92%	-5.4270***	23.52%	-7.6626***	-3.1538**		
	(0.0018)	(0.1242)	(27.79%, 38.50%)	(1.9647)	(13.98%, 36.77%)	(1.7500)	(1.5900)		
GREQuantPct						0.0803***	0.0283		
						(0.0206)	(0.0181)		
Gender						0.0907	-0.0455		
						(0.0878)	(0.1258)		
Demographic1						-0.1575	-0.3403*		
						(0.1325)	(0.2031)		
Demographic2						0.1673	-0.2274		
						(0.1341)	(0.2014)		
AdvMath							0.1762		
							(0.1350)		
ComScore							0.0376		
							(0.0304)		
YearTransMat		-0.3883***	24.98%		70.26%				
		(0.1056)	(22.60%, 27.52%)		(42.57%, 88.27%)				
$\dot{m}_2$	1.0524***								
	(0.0003)								
$\overline{m_{2A}}$				0.9855**					
				(0.3937)					
$\sqrt{\text{MeanSqError}}$	0.0041***								
	(0.0002)								
Beta								0.6194***	
								(0.0889)	
Sigma									0.1589***
									(0.0526)
Observations	1000								
LL | Param | AIC	-189.7	21	421.5						
LR Stat | DF | P-Val	48.5585	16	0.0000						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	F$\overline{m_{2A}}$	Cutoff1	$P(\text{Cutoff})_1$	Cutoff3	$P(\text{Cutoff})_3$	Success1	Success3	Beta	Sigma
Intercept	-1.2247***	-0.5082***	37.56%	-11.0710	7.70%	-10.6778***	-9.8095***		
	(0.0038)	(0.1163)	(32.38%, 43.04%)	(7.1634)	(0.88%, 44.03%)	(1.4717)	(2.1222)		
GREQuantPct						0.1193***	0.0978***		
						(0.0176)	(0.0239)		
Gender						-0.0260	-0.1737		
						(0.1189)	(0.1951)		
Demographic1						-0.5947***	-0.2075		
						(0.1843)	(0.3184)		
Demographic2						-0.0824	0.4580		
						(0.1758)	(0.2897)		
AdvMath							-0.0104		
							(0.2250)		
ComScore							0.1049*		
							(0.0570)		
YearTransMat		-0.4722***	27.28%		95.10%				
		(0.1070)	(23.97%, 30.86%)		(2.93%, 99.99%)				
$\dot{m}_2$	1.4277***								
	(0.0007)								
$\overline{m_{2A}}$				1.9422					
				(1.4272)					
$\sqrt{\text{MeanSqError}}$	0.0057***								
	(0.0005)								
Beta								0.9225***	
								(0.1209)	
Sigma									0.2832***
									(0.0617)
Observations	1000								
LL | Param | AIC	-275.0	21	592.0						
LR Stat | DF | P-Val	13.3576	16	0.6465						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

matriculants this year will give the committee the option to be more selective next year. Meanwhile, the second multiyear model rewards more diverse applicants, such as those interested in economic theory as opposed to applied research. In this example, by accepting more theory applicants this year, the committee will have to worry less next year about having too many applied students relative to the number of potential faculty advisors. But this is just one example; the model applies to any form of diversity that increases an applicant's success probability.

6 Optimality Tests

In the previous section, I used likelihood-ratio tests comparing the unrestricted and restricted likelihoods to assess the fit of the various models. For the year-static and waitlist-hybrid models, these tests rejected the null on only the datasets where the simulated committee was optimizing. However, for the year-dynamic, year-myopic, and waitlist-dynamic models, the tests always rejected the null. But a likelihood-ratio test only determines whether the model fits the data, not whether the committee can make any meaningful improvements by following the model's decision rule as opposed to its own rule. To make the latter determination, a different kind of statistical test is necessary, one that compares the model's decision rule to the committee's.

Suppose the committee adopts the year-static/dynamic/myopic or the waitlist-static/dynamic model's decision rules. Will the average success rate of the set of accepted (or matriculated) applicants increase, and if so, by how much? To answer these questions, I solve the model with $\sigma = 0$ in the conditional choice probabilities (CCPs), compute the yearly success rates of the model's acceptance or matriculation set, and compare those success rates to the success rates of the committee's acceptance or matriculation set. When I solve the model, I plug in the unrestricted success probability parameter estimates, not the restricted ones. The reason is that the restricted ones may be biased to rationalize suboptimal behavior (see Anderson et al. 2025). For example, if the committee puts too little weight on the quantitative GRE score, then its restricted parameter estimate will be biased down since the restricted CCPs assume optimality. I also adjust the cutoff probability parameter estimates to ensure that the transition, waitlist, and acceptance rates are unchanged. Doing so effectively makes the models follow a quota rule.

For the year-static model, I rank the applicants in each year by $P(y = S|x)$, set $\theta_{S_1^*}$ to match that year's transition rate, rank the survivors by $P(y = S|x,z)$, and set $\theta_{S_2^*}$ to match that year's acceptance rate. As such, the model transitions and accepts the same number of applicants as are transitioned and accepted in the data. In contrast, for the year-dynamic model, I rank the applicants in each year not by $P(y = S|x)$, but rather by the expected Round 2 social surplus function $\mathbb{E}[v_2(x,z,\varepsilon_2)|x] = \int_{\tilde{z}} \sigma \log \left(e^{P(y=S|x,\tilde{z})/\sigma} + e^{P(y_2^*=S)/\sigma} \right) dF(\tilde{z}|x)$.[12] σ and $\theta_{S_{2^*}}$ are structural parameters, so I use their restricted estimates in this step. After that, the procedure is the

[12]For the year-myopic model, expected social surplus is $\int_{\tilde{z}} \sigma \log \left(e^{P(y=S|x,\tilde{z})/\sigma} + e^{P(y_2^*=S)/\sigma} \right) dF(\tilde{z})$.

same as the year-static model's: I set $\theta_{S_1^*}$ to match that year's transition rate, rank the survivors by $P(y=S|x,z)$, and reset $\theta_{S_2^*}$ to match that year's acceptance rate. Once I have determined which applicants each model accepts, I compute their average success probability and compare it to the average success probability of the applicants that the committee's decision rule selects. Determining which applicants the committee's decision rule selects is straightforward. In Round 1, I rank the applicants by $P(d_1 = T|x)$, and in Round 2, I rank the survivors by $P(d_2 = A|x,z)$.

For the waitlist-hybrid and waitlist-dynamic models, the procedure is more complicated because the purpose of these models is to evaluate the committee's performance based not on the applicants the committee accepts but rather those who matriculate at the university. A committee that accepts strong applicants who do not matriculate due to getting offers from higher-ranked universities is not doing its job. For the waitlist-hybrid model, I proceed as follows for each year:

1. **Round 1:** Rank the applicants by $P(y=S|x)$, and set $\theta_{S_1^*}$ to match the transition rate.

2. **Round 2:** Rank the applicants by the payoff difference $P(y=S|x,z) - \beta\mathbb{E}[v_3(x,z,\overline{m_{2A}},\varepsilon_3)|x,z,\ddot{m}_2] = P(y=S|x,z) - \beta\int_{\widetilde{\overline{m_{2A}}}} \sigma \log\left(e^{P(y=S|x,z)/\sigma} + e^{P(y_3^*=S|\widetilde{\overline{m_{2A}}})/\sigma}\right) dF_{\overline{m_{2A}}|\ddot{m}_2}(\widetilde{\overline{m_{2A}}}|\ddot{m}_2)$. Like σ and $\theta_{S_3^*}$, β is a structural parameter, so I use its restricted estimate in this step. I observe the number of early acceptances N_{2A}, so the top N_{2A} applicants receive early acceptances.

3. **Round 2.5:** To determine how many and which early acceptances matriculate, first determine which applicants receive early acceptances from the committee's decision rule. Rank those applicants by their matriculation scores. Set the committee's number of early matriculants to match the actual number of early matriculants, and assume the applicants with the highest matriculation scores are the early matriculants. Now for the model, rank the model's early acceptances by their matriculation scores. However, the model's number of early matriculants is such that the model's marginal early matriculant's matriculation score is the lowest matriculation score at least as large as the committee's marginal early matriculant's matriculation score. As such, if the model is selecting applicants with lower matriculation scores than the committee, it will obtain fewer early matriculants.

4. **Round 3:** Assume that the early acceptances who do not matriculate in Round 2.5 will not matriculate at all. As such, rank the waitlisted applicants by their $P(y=S|x,z)$, and give late acceptances to the applicants atop this ranking. The committee's number of late acceptances (note that the committee ranks the applicants based on $P(d_3 = A|x,z)$) will match the data's number of late acceptances. To calculate the model's number of late acceptances, find the difference in the number of spots it must fill relative to the committee, where the final number of matriculants will match that in the data. Add this difference to the data's number of late acceptances. Then find the ratio of the number of spots it must fill relative to the committee, and multiply this ratio by the data's number. Finally, compute

the average of those two adjustments to the data's number of late acceptances.[13]

5. **Round 3.5:** For both the committee and the model, rank the late acceptances (who will be different for the committee vs. the model) by their matriculation scores. The committee will obtain the set of applicants atop its ranking whose size matches the number of late matriculants in the data. In contrast, the model will obtain the set of applicants atop its own ranking whose size matches the number of spots it needs to fill. Ultimately, both the committee and model will obtain the same number of matriculants as in the data.

The waitlist-dynamic's procedure is the same as the waitlist-hybrid's except for Round 1. In Round 1, the waitlist-dynamic model ranks the applicants based on the expected social surplus $\mathbb{E}[v_2(x,z,\ddot{m}_2,\varepsilon_2)|x] = \int_{\tilde{z}}\int_{\widetilde{\ddot{m}_2}} \sigma \log\left(e^{P(y=S|x,\tilde{z})/\sigma} + e^{\beta\mathbb{E}[v_3(x,\tilde{z},\widetilde{m_{2A}},\tilde{\varepsilon}_3)|x,\tilde{z},\widetilde{\ddot{m}_2}]/\sigma}\right) dF_{\ddot{m}_2|x,z}(\widetilde{\ddot{m}_2}|x,\tilde{z})\,dF_{z|x}(\tilde{z}|x)$. It then sets the cutoff level parameter $\theta_{S_1^*}$ to match the transition rate. After that, the procedure is identical to Steps 2–5, and I compute the average success probability of the model's matriculants.

Suppose I find deviation gains. Could committees have incorrect beliefs about which applicants are most likely to succeed? Also, if so, can they do a better job admitting applicants? Of course, it is also possible to find deviation losses if the committee knows something the model does not, such as which recommenders are most trustworthy. To prevent losses, algorithmically extracting as much information as possible from the application files is essential, and adding the committee score to z will also help prevent losses. In addition, if there are gains in the year-static, dynamic, or myopic models, the matriculation rate of the model's (stronger) acceptance set may be less than that of the committee's (weaker) acceptance set. In that case, the estimated gains would be an upper bound. However, the waitlist-hybrid and dynamic models accept those more likely to matriculate to prevent losing good, interested applicants by making them wait too long.

However, I must be careful of sampling error. The models have many parameters, so it is possible for the models to overfit the data. Ideally, I would test the performance of each model's deviation strategy on a holdout set. But given the limited number of observations, it is necessary to use all of them to estimate the parameters. k-fold cross-validation with large k would mitigate this problem, but it would introduce a new computational problem, as the models take far too long to estimate for it to be feasible, even with parallelization across multiple cores. Therefore, I follow Anderson et al. (2025) by using a randomization test to address sampling error.

If I assume for simplicity that the model is a two-round model that chooses the same transition set of applicants as the committee, then the following inequality holds:

[13] Suppose the committee gives 40 late acceptances to fill 5 spots, but the model must fill 10 spots because it obtained 5 fewer early matriculants. The model will have to give more than 40 late acceptances, but how many more? The additive adjustment implies the model will give 45 late acceptances, which is undoubtedly too low. However, the multiplicative adjustment of 80 late acceptances is likely too high because the additional late acceptances will be weaker applicants who are therefore more likely to matriculate if accepted. Therefore, I take the average of these adjustments, which rounds to 63 late acceptances. However, the results are robust to both the additive and multiplicative adjustments on their own. Analogous logic holds if the model must fill fewer spots than the committee.

Theorem 3. *Let d^R be the model's decision rule with $\sigma = 0$ in the CCPs, and let d^U be the committee's decision rule. By optimality:*

$$\frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S\middle|x_n, z_n; \hat{\theta}_S^U\right)\mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^R\left[P_{y_n}\left(\hat{\theta}_S^U\right)\right] = A\right\}$$
$$\geq \frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S\middle|x_n, z_n; \hat{\theta}_S^U\right)\mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^U\left[P_{y_n}\left(\hat{\theta}_A^U\right)\right] = A\right\}. \quad (25)$$

Proof. See Appendix F.

In this inequality, I am careful to assume that the committee's Round 1 decision rule is the same as the model's. Without this assumption, it is possible in rare cases for the sign to flip if, for example, the committee favors high-upside applicants in Round 1, and it turns out that favoring such applicants ultimately helps it attain a better set of accepted applicants (see Appendix F).

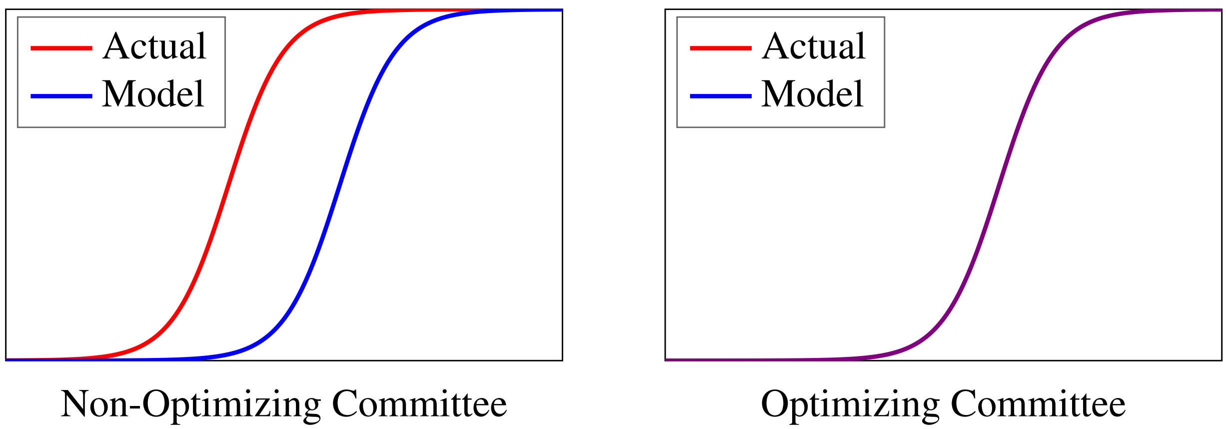


Figure 4: Deviation Gains CDFs

Regardless, for a perturbation $\widetilde{\theta_{S,t}^U}$ about $\hat{\theta}_S^U$'s asymptotic distribution, the sign can flip:

$$\frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S\middle|x_n, z_n; \widetilde{\theta_{S,t}^U}\right)\mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^R\left[P_{y_n}\left(\hat{\theta}_S^U\right)\right] = A\right\}$$
$$< \frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S\middle|x_n, z_n; \widetilde{\theta_{S,t}^U}\right)\mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^U\left[P_{y_n}\left(\hat{\theta}_A^U\right)\right] = A\right\}. \quad (26)$$

Given that the sign can flip if the committee optimizes or comes close to it, I can use such perturbations of the success parameters to test whether the deviation gains are robust to sampling error. Across $T = 500$ perturbations, I plot the CDFs of $\overline{P(\text{Success})}$ of the acceptance sets of the year-static, year-dynamic, year-myopic and reduced-form models. While the acceptance sets are computed with the actual parameter estimates, the $\overline{P(\text{Success})}$ values are computed with the perturbed estimates. The reduced-form model represents the committee's actutal decision rule.

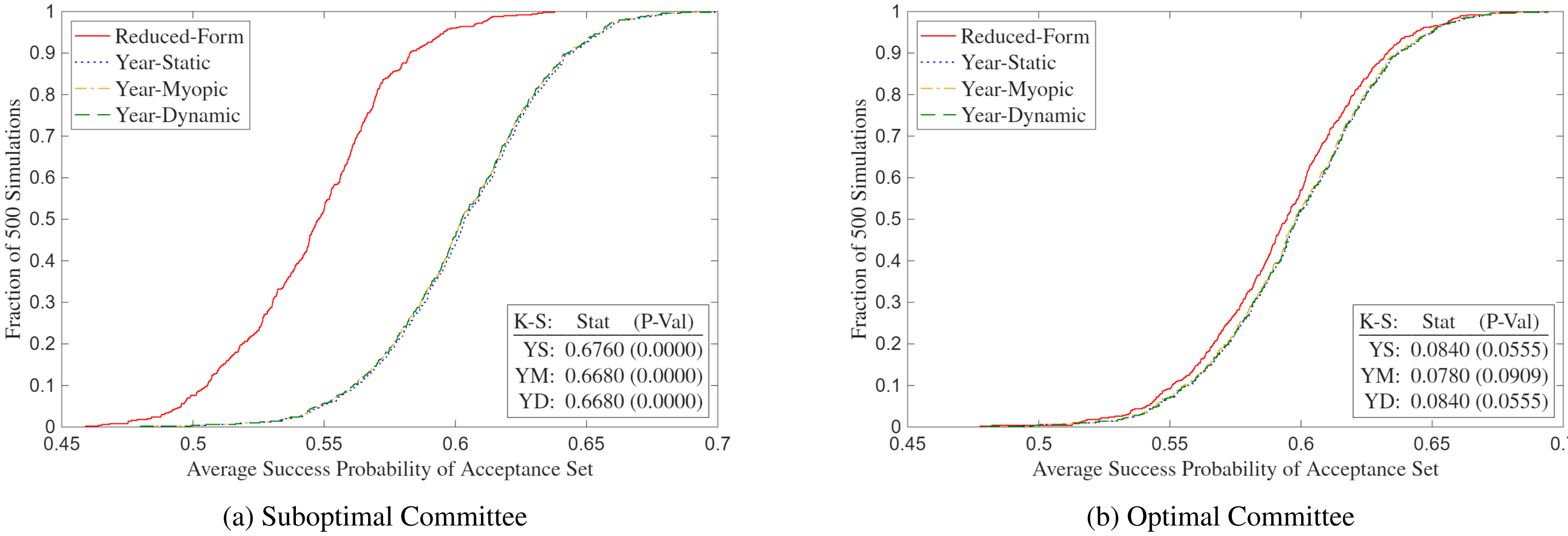
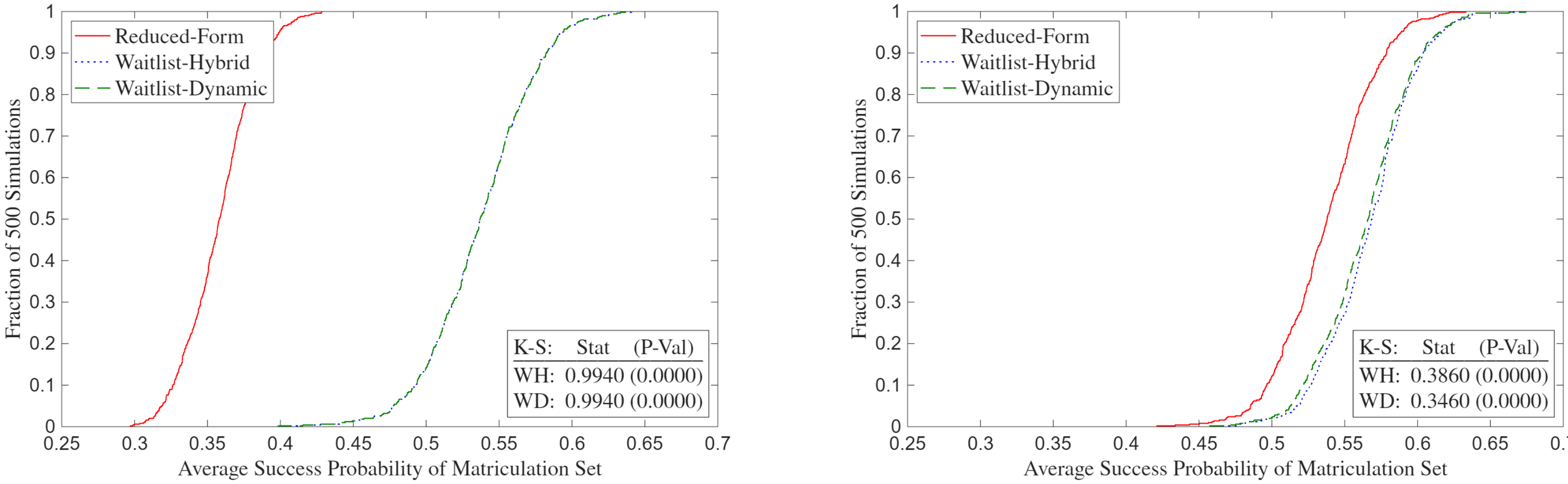


Figure 5: Deviation Gains

Figure 4 illustrates what the CDFs will look like in general. When a committee switches from a suboptimal decision rule to the optimal one, the actual decision rule's CDF shifts right to match the model's CDF. Per Figure 5, in the dataset where the committee is not optimizing, all of the models' CDFs stochastically dominate the reduced-form CDF. The reason is that the committee is discriminating against *Demographic1* and putting too much weight on the committee score relative to advanced math courses. But in the dataset where the committee is optimizing, the CDFs are almost identical. The models' CDFs still stochastically dominate the reduced-form CDF, but barely. Also, in both datasets, the year-dynamic CDF weakly stochastically dominates the year-myopic CDF because the year-dynamic model more effectively ranks the applicants in Round 1 by conditioning z's distribution on x, which the year-myopic model does not do.

To formally test deviation gains, I perform two-sample Kolmogorov-Smirnov (KS) tests of the equality of each of the models' CDFs and the reduced-form CDF. For this test, let $X = \overline{P_y\left[d\left(\hat{\theta}^U\right), \widetilde{\theta_S^U}\right]}$, and $F_T(x)$ be X's CDF over T perturbations. Then the test statistic $D_T = \sup_x \left|F_T^R(x) - F_T^U(x)\right|$, and I reject the null of equality at the α level if $D_T > \sqrt{\frac{-\log(\alpha/2)}{T}}$. As the test should, it rejects equality of the CDFs with $p < 0.0001$ all three times in the suboptimal dataset. But in the optimal dataset, it does not reject equality of any of the CDFs at the 5% level. In Figure 7 of Appendix C, I repeat the test for the datasets in which the subjective z variables are missing for the applicants rejected in Round 1, and the results are effectively the same.

Finally, I simulate deviation gains for the waitlist models. Analogous to the two-round models, I plot CDFs for the waitlist-hybrid, waitlist-dynamic, and reduced-form models, though this reduced-form model is the three-round version. However, as explained earlier in this section, the committee does not ultimately care about the $\overline{P(\text{Success})}$ of its acceptance set, but rather of its matriculation set. Therefore, the expected success probabilities are of the matriculation sets.



(a) Suboptimal Committee

(b) Optimal Committee

Figure 6: Waitlist Deviation Gains

The results are in Figure 6, and as expected, the deviation gains are much larger for the suboptimal dataset. The KS tests reject the null for both the suboptimal and optimal datasets, considering the challenge of constructing a three-round dataset that is perfectly optimal. But for the optimal dataset, the deviation gains for the 50th percentile across perturbations are 3.67% points for the waitlist-hybrid model and 3.65% points for the waitlist-dynamic model. In contrast, for the suboptimal dataset, the deviation gains are 15.08% points for both models. In Figure 8 of Appendix C, I repeat this test for the datasets in which the subjective z variables, as well as the matriculation scores, are missing for the applicants rejected in Round 1. The results are effectively the same, though for both the suboptimal and optimal datasets, the waitlist-hybrid model outperforms the waitlist-dynamic model by a greater margin than in Figure 6. Ultimately however, the main question is this: would deviation gains be borne out in practice? The only way to know for sure would be for a real-world admissions committee to try implementing the various models' decision rules and see if the result of doing so is stronger cohorts over the years.

Finally, it is possible to use the model's acceptance or matriculation set in year t to put a dollar value on success in t. Assume that it takes h hours for a committee member to read an application file, and also that each committee member makes an hourly wage of w. Therefore, reading N_T files has a monetary cost of $c = N_T \times h \times w$. If I counterfactually increase N_T, then $\overline{P(\text{Success})}$ of the acceptance or matriculation set increases because the committee discards fewer applicants in Round 1, who could have a high $P(y = S|x,z)$. As such, I can compute $\overline{P(\text{Success})}$ for all possible values of N_T, where the lower bound is the number of acceptances N_A, and the upper bound is the number of applicants N. At that point, I can plot c vs. $\overline{P(\text{Success})}$ and calculate the derivative of the curve numerically. This derivative is the dollar value of a marginal increase in $\overline{P(\text{Success})}$, provided the committee is choosing the optimal value of N_T. While the derivative will be positive, it may not be constant, e.g. under diminishing returns to reading more files.

7 Conclusion

This paper studies how decision-makers select among options whose true quality is uncertain and costly to evaluate. Graduate admissions provides a natural setting for such a study, as it features a low-cost filtering process followed by a high-cost review of the application files that survive the filter. Because the committee has a target cohort size, and not all accepted applicants will matriculate, there is also staggered timing of offers via a waitlist. By structurally modeling the committee's decision problem. I develop the first framework in the economics graduate admissions literature, from Ehrenberg and Mavros (1995) to Bai et al. (2022), that links application information to admissions decisions and post-program success. This framework makes optimality testable, and deviation gains from switching to the optimal decision rule quantifiable.

Before estimating structural models, I start with reduced-form regressions and lasso of a binary acceptance variable on information in application files. I find that relative to the university's other social science programs, economics admissions appears more data-driven, as the economics regressions have a higher goodness of fit both in and out of sample. I also find evidence of network effects, as having a recommender who wrote letters for other applicants in the sample is predictive of acceptance. This result holds across the social science programs. Moreover, I construct Wald tests that can determine whether the determinants of acceptance or success are the same across programs. Using simulated data, I show that these tests reject the null only on a dataset where I set the data-generating parameters to make the determinants different.

However, regressions and Wald tests cannot determine whether an admissions committee is behaving optimally. If I were to test whether the average marginal effects (AMEs) of variables in application files, such as GRE scores, are the same in an acceptance regression vs. a success regression, I would not be testing whether the committee is optimizing. As I show at the end of Section 4, equal AMEs are neither necessary nor sufficient for optimization. Therefore, any optimality tests must be based on structural models of the committee's selection problem that account for evaluation costs and multi-stage decision-making. For this reason, I develop and estimate structural models of the committee's problem in order of increasing complexity.

I start with static, dynamic, and myopic two-round models. In these models, there is a Round 1 filter based on only objective information like GRE scores, followed by a Round 2 final decision based on objective and subjective information, such as variables created by textual analysis of recommendation letters. I then proceed to three-round waitlist-hybrid and waitlist-dynamic models. Round 1 in these models is the same as in the two-round models, but in Round 2, the committee can accept or waitlist the survivors of Round 1. Finally, in Round 3, the committee gives additional acceptances off the waitlist and rejects the remaining waitlisted applicants. In Round 2, the committee is more likely, all else equal, to give early acceptances to applicants with higher matriculation probabilities. The reason for doing so is to ensure that the program can fill its cohort without having to give too many late acceptances to weaker applicants. Ideally, the committee will be able to uncover variables, such as referencing a specific professor in the application essay, that predict matriculation without also indicating a lower success probability. This situation is analogous to the small-market Oakland Athletics baseball team in *Moneyball* using analytics to find good players *they could afford* (Lewis 2004).

After estimating the structural models, I use the estimates to compare each two-round model's decision rule to the committee's decision rule. In a simulated dataset where the committee is making mistakes, such as discriminating against a demographic, the CDF of the average success probabilities of the model's acceptance set (across perturbations of the estimates about their asymptotic distribution) stochastically dominates the CDF of the average success probabilities of the committee's acceptance set. However, in a simulated dataset where the committee is optimizing, a Kolmogorov-Smirnov fails to reject the null of equal CDFs. I repeat this process for the three-round models, this time comparing the average success probabilities of the matriculation sets because accepting a good applicant is only beneficial if that applicant matriculates. In the suboptimal dataset, the deviation gains are over four times larger than in the optimal dataset.

In addition to its contributions to the graduate admissions literature and to methodological literatures on dynamic selection and costly evaluation, this paper also contributes to the economics literature on synthetic data and privacy protection. Recent economics papers use synthetic U.S. Census Board data and compare the synthetic data estimates to estimates on actual data, but this is the first economics paper to design and implement a comprehensive synthetic data protocol from the construction of the synthetic data to the final estimation results. As a result, it serves as a template for researchers who wish to work with confidential data that they are not allowed to access directly, which can help researchers make discoveries they could not have otherwise.

More broadly, this paper provides a tractable framework for studying selection under costly evaluation. By doing so, it advances the existing literature from identifying variables that predict success to testing, conditional on time and effort constraints, whether decision rules are optimal. This framework can uncover undervalued signals, sources of potential discrimination, and efficiency losses from misweighted criteria. For example, in the simulated dataset where the committee was undervaluing advanced math and discriminating against a demographic, the likelihood-ratio and deviation gains tests reject the null of optimality. Moreover, by computing the counterfactual of following the optimal decision rule, I was able to determine how much stronger the acceptance and matriculation sets would be if not for the misweightings.

This study is based on a single institution, so any specific empirical reuslts reflect that setting. However, the framework is designed to apply broadly to environments where due to evaluation costs, decision-makers must employ a multi-stage process to select among their available options. Such environments range from hiring committees to clinical trials of medicines. That said, a worthwhile avenue for future research would be getting data from multiple universities to model the applicant's decision problem and compute equilibrium admissions outcomes for applicants across universities. Meanwhile, the simulation exercises establish identification and the statistical power of the optimality tests while implementation on the full admissions data under the synthetic data protocol is ongoing. In addition, admissions committees may pursue objectives beyond maximizing success probabilities, such as diversity or research field balance. The framework can be extended to incorporate such objectives directly into the committee's problem.

Ultimately, the models in this paper can help real-world decision-makers make better selection decisions based on *ex-ante* information when faced with uncertainty and costly evaluation. Crucially, decision-makers can improve selection outcomes without having to increase their evaluation effort, or resort to satisficing instead of optimizing. Instead, they can use any of the paper's structural framework and optimality tests to refine how they use that information, and consequently do the best possible job at selecting from their set of available options.

References

Abowd, J. M., & Schmutte, I. M. (2015). Economic Analysis and Statistical Disclosure Limitation. *Brookings Papers on Economic Activity*, 221–267.

Anderson, A., Rosen, J., Rust, J., & Wong, K.-P. (2025). Disequilibrium Play in Tennis. *Journal of Political Economy*, *133*(1), 190–251.

Athey, S., Katz, L. F., Krueger, A. B., Levitt, S., & Poterba, J. (2007). What Does Performance in Graduate School Predict? Graduate Economics Education and Student Outcomes. *American Economic Review*, *97*(2), 512–520.

Athey, S., & Luca, M. (2019). Economists (and Economics) in Tech Companies. *Journal of Economic Perspectives*, *33*(1), 209–230.

Attiyeh, G., & Attiyeh, R. (1997). Testing for Bias in Graduate School Admissions. *The Journal of Human Resources*, *32*(3), 524–548.

Bai, J., Esche, M., MacLeod, W. B., & Shi, Y. (2022). Subjective Evaluations and Stratification in Graduate Education. *NBER Working Paper 30677*.

Benedetto, G., Stanley, J. C., & Totty, E. (2018). The Creation and Use of the SIPP Synthetic Beta v7.0. *Census Working Papers*.

Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender Identity and Relative Income within Households. *The Quarterly Journal of Economics*, *130*(2), 571–614.

Caplin, A., & Dean, M. (2015). Revealed Preference, Rational Inattention, and Costly Information Acquisition. *American Economic Review*, *105*(7), 2183–2203.

Carr, M. D., Wiemers, E. E., & Moffitt, R. A. (2025). Using Synthetic Data to Estimate Earnings Dynamics: Evidence from the Survey of Income and Program Participation Gold Standard File and Synthetic Beta. *Harvard Data Science Review, Special Issue 6*.

Chapelle, R., & Falissard, B. (2023). Statistical properties and privacy guarantees of an original distance-based fully synthetic data generation method. *arXiv:2310.06571*.

Chetty, R., Deming, D. J., & Friedman, J. N. (2026). Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges. *The Quarterly Journal of Economics*, *141*(1), 51–145.

Ehrenberg, R. G., & Mavros, P. G. (1995). Do Doctoral Students' Financial Support Patterns

Affect Their Times-to-Degree and Completion Probabilities? *The Journal of Human Resources*, *30*(3), 581–609.

Ferreyra, M. M., Garriga, C., Martín-Ocampo, J. D., & Sánchez-Díaz, A. M. (2023). Cows Don't Give Milk: An Effort Model of College Graduation. *EdWorkingPaper No. 23-713*.

García Guzmán, P. (2026). General citation. *The Econ PhD placements dataset*.

Grove, W. A., Dutkowsky, D. H., & Grodner, A. (2007). Survive Then Thrive: Determinants of Success in the Economics PH.D. Program. *Economic Inquiry*, *45*(4), 864–871.

Grove, W. A., & Wu, S. (2007). The Search for Economics Talent: Doctoral Completion and Research Productivity. *American Economic Review*, *97*(2), 506–511.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, *47*(1), 153–161.

Jones, A., Schuhmann, P., Soques, D., & Witman, A. (2020). So you want to go to graduate school? Factors that influence admissions to economics PhD programs. *The Journal of Economic Education*, *51*(2), 177–190.

Koenecke, A., & Varian, H. (2020). Synthetic Data Generation for Economists. *American Economic Association Annual Meeting*.

Krueger, A. B., & Wu, S. (2000). Forecasting Job Placements of Economics Graduate Students. *The Journal of Economic Education*, *31*(1), 81–94.

Lewis, M. (2004). *Moneyball: The Art of Winning an Unfair Game*. W.W. Norton & Company.

Little, R. J. (1993). Statistical Analysis of Masked Data. *Journal of Official Statistics*, *9*(2), 407–426.

Madera, J. M., Hebl, M. R., & Martin, R. C. (2009). Gender and Letters of Recommendation for Academia: Agentic and Communal Differences. *Journal of Applied Psychology*, *94*(6), 1591–1599.

Matějka, F., & McKay, A. (2015). Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review*, *105*(1), 272–298.

McCall, J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics*, *84*(1), 113–126.

Mood, C. (2010). Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. *European Sociological Review*, *26*(1), 67–82.

Mortensen, D. T. (1970). Job Search, the Duration of Unemployment, and the Phillips Curve. *American Economic Review*, *60*(5), 847–862.

Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

PandaInUniv. (2026). General citation. *PandaInUniv*.

Patki, N., Wedge, R., & Veeramachaneni, K. (2016). The Synthetic data vault. *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 399–410.

Rubin, D. B. (1993). Discussion: Statistical Disclosure Limitation. *Journal of Official Statistics*, *9*(2), 461–468.

Schlauch, G., & Startz, R. (2018). The path to an economics PhD. *Economics Bulletin*, *38*(4), 1864–1876.

Siegfried, J. J., & Stock, W. A. (2007). The Undergraduate Origins of PhD Economists. *The Journal of Economic Education*, *38*(4), 461–482.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.

Stanley, J. C., & Totty, E. S. (2024). Synthetic Data and Social Science Research: Accuracy Assessments and Practical Considerations from the SIPP Synthetic Beta. *NBER Working Paper 32979*.

Stock, W. A., Finegan, T. A., & Siegfried, J. J. (2006). Attrition in Economics Ph.D. Programs. *American Economic Review*, *96*(2), 458–466.

Stock, W. A., Finegan, T. A., & Siegfried, J. J. (2009a). Can you earn a Ph.D. in economics in five years? *Economics of Education Review*, *28*(5), 523–537.

Stock, W. A., Finegan, T. A., & Siegfried, J. J. (2009b). Completing an Economics PhD in Five Years. *American Economic Review*, *99*(2), 624–629.

Stock, W. A., & Siegfried, J. J. (2014). Fifteen Years of Research on Graduate Education in Economics: What Have we Learned? *The Journal of Economic Education*, *45*(4), 287–303.

Stock, W. A., & Siegfried, J. J. (2015). The Undergraduate Origins of PhD Economics Revisited. *The Journal of Economic Education*, *46*(2), 150–165.

Stock, W. A., Siegfried, J. J., & Finegan, T. A. (2011). Completion Rates and Time-to-Degree in Economics PhD Programs. *American Economic Review*, *101*(3), 176–188.

Vuong, Q. H. (1989). Likelihood Ratio Tests of Model Selection and Non-Nested Hypotheses. *Econometrica*, *57*(2), 307–333.

Weitzman, M. L. (1979). Optimal Search for the Best Alternative. *Econometrica*, *47*(3), 641–654.

Young, L., & Soroka, S. (2012). Affective News: The Automated Coding of Sentiment in Political Texts. *Political Communications*, *29*(2), 205–231.

Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B*, *68*(1), 49–67.

Appendices

A Variables, Rankings, and Dictionaries

This table contains a list of the variables in this paper. Variables available only for the economics program are bold, and those available for all but economics are italic. *Number of Math Courses* and $\mathbb{1}$(*Advanced Math*) became objective variables starting in 2020 due to a change in the university's application system. Program rankings are from U.S. News, and placement rankings are from IDEAS/RePEc. The dictionaries for the subjective independent variables are from Madera, Hebl, and Martin (2009), Young and Soroka (2012), and Bai et al. (2022). When using the dictionaries, I implement the Porter stemmer to reduce words to their roots.

Table 8: Variables

(a) Dependent

Admission	Success
$\mathbb{1}$(Transition)	**$\mathbb{1}$(Complete PhD Program)**
$\mathbb{1}$(Accept)	**$\mathbb{1}$(Academic Placement)**
$\mathbb{1}$(Waitlist)	**$\mathbb{1}$(Government Placement)**
$\mathbb{1}$(Open House)	**$\mathbb{1}$(Consulting Placement)**
$\mathbb{1}$(Matriculate)	**$\mathbb{1}$(Tech Placement)**
$\mathbb{1}$(Attend PhD Program)	**Placement U.S. Ranking**
PhD Program U.S. Ranking	**Placement Global Ranking**
PhD Program Global Ranking	**$\mathbb{1}$(Good Placement)**

(b) Objective Independent

Demographic	Performance	Economics-Only Performance
$\mathbb{1}$(Covid Year)	GRE Verbal Percentile	**Masters GPA**
Winsorized Age	GRE Quantitative Percentile	**$\mathbb{1}$(Missing Masters GPA)**
$\mathbb{1}$(Female)	GRE Analytic Percentile	**$\mathbb{1}$(Elite U.S. Masters)**
$\mathbb{1}$(U.S. African/Hispanic/Native)	$\mathbb{1}$(Missing GRE Percentiles)	**$\mathbb{1}$(Elite International Masters)**
$\mathbb{1}$(U.S. Asian)	Undergraduate GPA	**$\mathbb{1}$(Other International Masters)**
$\mathbb{1}$(International Asian)	$\mathbb{1}$(Missing Undergraduate GPA)	**$\mathbb{1}$(Research Experience)**
$\mathbb{1}$(International Other)	TOEFL/IELTS Score	**Objective Score**
	$\mathbb{1}$(Missing TOEFL/IELTS Score)	**Relative Committee Score**
	$\mathbb{1}$(Elite U.S. University)	
	$\mathbb{1}$(Elite U.S. Liberal Arts College)	
	$\mathbb{1}$(Elite International University)	
	$\mathbb{1}$(Other International University)	
	$\mathbb{1}$(Graduate Degree)	
	$\mathbb{1}$(Work Experience)	
	Number of Professor Recommenders	
	Number of Prolific Recommenders	
	Concentration Dummy Variables	

(c) Subjective Independent

Recommendation Letter	Essay/CV/Supplemental Form
Relative Letter Length	**Essay Length (Essay)**
Relative Letter Features	**$\mathbb{1}$(Specific Topic) (Essay)**
Positive Words (YS 2012)	**$\mathbb{1}$(Specific Professor) (Essay)**
Negative Words (YS 2012)	**$\mathbb{1}$(Economics Work Experience) (CV)**
Standout Words (BEMS 2022)	**$\mathbb{1}$(Working Paper) (CV)**
Ability Words (BEMS 2022)	**$\mathbb{1}$(Publication) (CV)**
Research Words (BEMS 2022)	**$\mathbb{1}$(Coding) (CV)**
Grindstone Words (BEMS 2022)	**Number of Math Courses (Supplemental)**
Teaching Words (BEMS 2022)	**$\mathbb{1}$(Advanced Math) (Supplemental)**
Communal Words (BEMS 2022)	**$\mathbb{1}$(Theory Field) (Essay/CV)**
Agentic Words (MHM 2009)	**$\mathbb{1}$(Applied Field) (Essay/CV)**

Academia-Equivalent Rankings (Government)

- **European Institutions:** European Central Bank (80), HM Treasury (120)
- **Foreign Central Banks:** Banco Central do Brasil (120), Banco de México (120), Bank of Canada (80), Bank of England (80), Bank of Israel (120), Bank of Italy (120), Bank of Japan (80), Bank of Spain (120), Banque de France (80), Central Bank of Chile (120), Central Bank of Colombia (120), Central Bank of Korea (80), Central Bank of Turkey (120), Norges Bank (80), Reserve Bank of Australia (80), Reserve Bank of India (120), Reserve Bank of New Zealand (80), Sveriges Riksbank (80), Swiss National Bank (80)
- **Multilaterals/International:** African Development Bank (120), Asian Development Bank (120), Bank for International Settlements (80), European Bank for Reconstruction and Development (120), European Investment Bank (120), Inter-American Development Bank (120), International Monetary Fund (40), Statistics Canada (120), World Bank (40)
- **Policy Research Institutes:** American Enterprise Institute (120), American Institutes for Research (135), Becker Friedman Institute (80), Brookings Institution (80), Center for Global Development (120), Center for Migration Studies in New York (135), Center for Strategic and International Studies (135), Cowles Foundation (80), Harris School Policy Labs (120), Hoover Institution (80), IDinsight (135), Innovations for Poverty Action (135), J-PAL (80), MDRC (135), National Bureau of Economic Research (80), Peterson Institute for International Economics (80), Pew Research Center (135), RAND Corporation (120), Resources for the Future (120), Urban Institute (120)
- **U.S. Executive/Cabinet:** Department of Commerce (120), Department of Justice (80), Department of Labor (120), Department of the Treasury (120)
- **U.S. Executive Policy Offices:** Council of Economic Advisers (80), Office of Management and Budget (80)
- **U.S. Federal Reserve System:** Board of Governors (40), Atlanta, Boston, Chicago, Cleveland, Dallas, Kansas City, Minneapolis, New York, Philadelphia, Richmond, San Francisco, St. Louis (80)
- **U.S. GSE/Housing Finance:** Fannie Mae (120), Freddie Mac (120)
- **U.S. Independent Agencies:** Commodity Futures Trading Commission (120), Federal Deposit Insurance Corporation (120), Federal Energy Regulatory Commission (120), Federal Housing Finance Agency (120), Federal Trade Commission (80), Office of Financial Research (80), Securities and Exchange Commission (80)
- **U.S. Legislative Committees:** Congressional Budget Office (80), Congressional Research Service (120), Joint Committee on Taxation (80), Joint Economic Committee (120)

- **U.S. Science and Research:** National Science Foundation (80)
- **U.S. Statistical Agencies:** Bureau of Economic Analysis (120), Bureau of Labor Statistics (120), Census Bureau (120)

Academia-Equivalent Rankings (Consulting/Technology/Other)

- **Consulting:** Aecon Consulting (135), AlixPartners (135), Analysis Group (120), Bain & Compnay (120), Bates White (135), Berkeley Research Group (120), Boston Consulting Group (120), Brattle Group (120), Charles River Associates (120), Coleman Research (135), Compass Lexecon (120), Cornerstone Research (120), Deloitte Economic and Financial Advisory (135), Econic Partners (135), Economists Incorporated (135), Edgeworth Economics (135), Ernst & Young Economic Advisory (135), Frontier Economics (135), FTI Consulting (135), Guidehouse (135), KPMG Economics & Valuation (135), Matrix Economics (135), McKinsey & Company (120), Mercer (135), NERA Economic Consulting (120), Oliver Wyman (135), PwC Economics, Roland Berger (135), Willis Towers Watson (135)
- **Technology (Athey & Luca 2019):** Airbnb (135), Alibaba (120), Amazon (120), AppNexus (135), Apple (120), Block/Square (135), ByteDance/TikTok (120), CoreLogic (135), Coursera (135), DStillery (135), Didichuxing (120), Digonex (135), DoorDash (135), eBay (120), ECONorthwest (135), Expedia (135), Forkcast (135), Glassdoor (135), Google (120), Granular (135), Groupon (135), Houzz (135), Huawei (135), IBM (120), Indeed (135), ING (135), Instacart (135), Intel (135), Kensho (135), Lending Club (135), LinkedIn (120), Lyft (120), Meta/Facebook (120), Microsoft (120), Netflix (120), Nuna (135), Nvidia (120), Pandora (135), PayPal (135), Pinterest (135), Prattle (135), Quantco (135), Quora (135), Redfin (135), Ripple (135), Roblox (135), Rover (135), Salesforce (135), Shopify (135), Spotify (135), Stripe (135), Trulia (135), Uber (120), Upwork (135), Visa (135), Walmart (135), Wealthfront (135), Yahoo (120), Yelp (135), Zillow (135)
- **Technology Laboratories:** Amazon Science (80), Apple Machine Learning Research (80), DeepMind (40), Google Research (40), Meta Research (80), Microsoft Research (40), Netflix Research (80), OpenAI Research (80), Uber Research (80)
- **Other:** AIG (135), Bank of America, (120), Bloomberg (120), Boeing (135) Capital One (135), Exelon (135), ExxonMobil (135), Ford (135), General Electric (135), General Motors (135), Goldman Sachs (120), JPMorgan Chase (120), Liberty Mutual (135), Moody's Analytics (120), Morningstar (135), Nielsen (135), PNC (135), Progressive (135), S&P Global (120), Shell (135), Swiss Re (135), Toyota (135), Other (135), No Placement (150)

Dictionaries (Recommendation Letters)

- **Positive/Negative Words:** Young and Soroka (2012)'s Lexicoder Sentiment Dictionary
- **Standout, Ability, Research, Grindstone, Communal Words:** Bai et al. (2022)
- **Agentic Words:** assertive, confident, aggressive, ambitious, dominant, forceful, independent, daring, outspoken, intellectual (Madera et al. 2009); **AND** audacious, bold, command, compete, decisive, determine, drive, enterprise, fearless, influential, leader, pioneer, powerful, proactive, resilient, resourceful, self-assured, strategic, tenacious, vision
- **Teaching Words:** academic, assess, clarity, coach, collaborate, communicate, cultivate, curriculum, demonstrate, develop, educate, encourage, engage, enlighten, evaluate, explain, facilitate, foster, guide, impart, inspire, insight, instruct, knowledge, learn, lesson, mentor, nurture, participate, pedagogy, present, scholar, support, teach, train, tutor, workshop. (Note: I constructed this dictionary because it is unavailable in Bai et al. 2022.)

Dictionaries (Application Essays/CVs)

- **Specific Topic:** applied, asset, behavior, climate, computation, data, development, dynamic programming, economics, education, empirical, environment, experiment, finance, fiscal, game theory, health, immigration, industrial, inequality, international, io, labor, linear programming, macro, math, metrics, micro, monetary, policy, political, poverty, pricing, probability, public, sports, statistic, theory, trade, urban
- **Economics Work Experience:** assistant, fed, imf, international monetary fund, predoc, ra, research, treasury, world bank
- **Working Paper:** arxiv, center for economic policy research, cepr, econpapers, national bureau of economic research, nber, research papers in economics, repec, social science research network, ssrn, working paper
- **Publication:** Journal names from IDEAS/RePEc Aggregate Rankings for Journals
- **Coding:** c/c++, eviews, fortran, java, jax, julia, latex, matlab, python, r, sas, stata, webscrape

B Additional Reduced-Form Results

Table 9: Means by Program

(a) All Applicants

	Economics	Government	History	Linguistics	Psychology
Accept	0.1121	0.0430	0.0986	0.0678	0.0535
CovidYear	0.2478	0.2792	0.3029	0.3460	0.3004
WinsorAge	26.2782	27.9633	26.8478	28.0692	25.8230
Female	0.4201	0.4182	0.4232	0.6133	0.7490
USAfrHispNat	0.0289	0.0766	0.1159	0.0665	0.1728
USAsian	0.0313	0.0421	0.0304	0.0353	0.0576
IntAsian	0.5322	0.2510	0.1058	0.2714	0.1523
IntOther	0.2517	0.2967	0.2203	0.3256	0.1584
GREVerbalPct	71.9000	78.7244	78.2719	75.0315	75.6655
GREQuantPct	85.9263	67.9672	64.2361	68.4772	65.1577
GREAnalyticPct	52.8672	67.3694	66.7717	62.3525	65.7340
UgradGPA	3.6252	3.5884	3.6401	3.6304	3.5798
TOEFL/IELTS	7.5200	7.5234	7.5035	7.5239	7.4941
EliteUSUni	0.1516	0.1847	0.2565	0.1954	0.2840
EliteUSLA	0.0423	0.0672	0.0696	0.0231	0.0535
EliteForeign	0.0717	0.0578	0.0478	0.0488	0.0206
OtherForeign	0.5736	0.3953	0.2333	0.5102	0.1852
GradDegree	0.7478	0.7553	0.6652	0.7544	0.4259
WorkExper	0.5938	0.7463	0.6638	0.7544	0.7593
#ProfRecom	2.5173	2.3469	2.6261	2.4355	2.1831
#ProlificRecom	1.0760	0.6912	0.5043	0.7069	0.3128
#MathCourses	6.6396				
AdvMath	0.4827				
MissingGRE	0.0255	0.4612	0.6217	0.6839	0.4650
MissingUgradGPA	0.6607	0.4827	0.3072	0.5726	0.2490
MissingTOEFL/IELTS	0.5958	0.7714	0.8449	0.7707	0.8827
Observations	2078	2231	690	737	486

(b) Accept = 1

	Economics	Government	History	Linguistics	Psychology
CovidYear	0.1888	0.2188	0.2059	0.2200	0.2308
WinsorAge	25.0815	27.7292	27.2794	27.1000	25.7692
Female	0.4850	0.5000	0.5588	0.5800	0.6154
USAfrHispNat	0.0386	0.1458	0.1765	0.0400	0.0769
USAsian	0.0429	0.0312	0.0147	0.0600	0.1154
IntAsian	0.4163	0.1667	0.0882	0.3600	0.0769
IntOther	0.2575	0.2292	0.3235	0.2000	0.1538
GREVerbalPct	84.4311	87.2304	82.2780	79.1768	85.7920
GREQuantPct	91.8188	71.3019	70.0275	73.5675	65.3062
GREAnalyticPct	69.1136	79.6735	69.6724	67.7679	73.1803
UgradGPA	3.7127	3.6539	3.7073	3.6783	3.6421
TOEFL/IELTS	7.6383	7.5835	7.5520	7.6035	7.4965
EliteUSUni	0.2833	0.3125	0.3382	0.2200	0.5385
EliteUSLA	0.1030	0.1458	0.1176	0.0200	0.1154
EliteForeign	0.1030	0.0521	0.0588	0.1000	0.0385
OtherForeign	0.3863	0.2396	0.2647	0.4000	0.1154
GradDegree	0.5837	0.6771	0.7206	0.7400	0.4615
WorkExper	0.6309	0.8125	0.7647	0.7000	0.9231
#ProfRecom	2.5494	2.3750	2.8235	2.7200	2.2308
#ProlificRecom	1.2833	0.9167	0.6765	1.0800	0.4615
#MathCourses	7.2575				
AdvMath	0.6781				
MissingGRE	0.0086	0.2917	0.6324	0.5800	0.1923
MissingUgradGPA	0.4893	0.3125	0.3088	0.4800	0.1923
MissingTOEFL/IELTS	0.6567	0.8854	0.8529	0.7600	0.8846
Observations	233	96	68	50	26

Table 10: Correlation Matrix

	Covid	Age	Fem	USAHN	USAsn	IntAsn	IntOth	GREV	GREQ	GREA	GPA	T/I	EUSUni	EUSLA	EFor	OthFor	Grad	Work	Prof	Pro	MGRE	MGPA	MT/I
CovidYear	1.00																						
WinsorAge	0.01	1.00																					
Female	0.02	-0.13	1.00																				
USAfrHispNat	0.00	-0.04	0.04	1.00																			
USAsian	-0.00	-0.06	0.05	-0.05	1.00																		
IntAsian	-0.05	-0.07	0.02	-0.19	-0.14	1.00																	
IntOther	0.01	0.18	-0.06	-0.17	-0.12	-0.42	1.00																
GREVerbalPct	0.03	-0.07	-0.03	-0.01	0.04	-0.09	-0.17	1.00															
GREQuantPct	0.01	-0.14	-0.08	-0.15	-0.00	0.39	-0.06	0.22	1.00														
GREAnalyticPct	0.04	-0.08	0.02	0.06	0.08	-0.34	-0.11	0.59	-0.14	1.00													
UgradGPA	-0.02	-0.28	0.07	-0.12	-0.02	-0.00	0.00	0.09	0.09	0.11	1.00												
TOEFL/IELTS	-0.01	-0.10	0.03	0.00	0.01	0.01	-0.02	0.26	0.06	0.23	0.00	1.00											
EliteUSUni	0.01	-0.12	0.02	0.12	0.14	-0.12	-0.22	0.16	-0.00	0.20	-0.00	0.00	1.00										
EliteUSLA	-0.02	-0.02	0.04	0.04	0.04	-0.08	-0.09	0.12	-0.03	0.14	-0.01	0.00	-0.11	1.00									
EliteForeign	-0.00	-0.04	-0.00	-0.06	-0.04	0.09	0.09	0.02	0.08	-0.01	0.00	0.04	-0.12	-0.06	1.00								
OtherForeign	-0.02	0.20	-0.04	-0.23	-0.14	0.34	0.39	-0.25	0.19	-0.38	0.00	-0.02	-0.43	-0.21	-0.22	1.00							
GradDegree	0.00	0.41	-0.08	-0.13	-0.09	0.21	0.15	-0.14	0.04	-0.21	-0.24	-0.01	-0.21	-0.10	0.05	0.37	1.00						
WorkExper	0.01	0.33	0.05	0.05	0.01	-0.18	0.11	-0.00	-0.14	0.09	-0.07	-0.03	-0.02	0.03	-0.03	-0.01	0.10	1.00					
#ProfRecom	-0.02	-0.19	-0.04	-0.04	0.01	0.11	-0.09	0.02	0.08	-0.03	0.13	0.02	-0.02	0.02	-0.00	-0.02	-0.03	-0.18	1.00				
#ProlificRecom	0.01	-0.06	-0.02	-0.07	-0.01	0.24	-0.10	0.01	0.20	-0.07	0.02	0.07	-0.01	-0.01	0.06	0.08	0.17	-0.09	0.19	1.00			
MissingGRE	0.09	0.15	0.05	0.11	-0.01	-0.22	0.17	-0.00	0.00	0.00	-0.07	-0.08	-0.03	-0.04	-0.03	-0.01	0.02	0.10	-0.08	-0.20	1.00		
MissingUgradGPA	-0.01	0.18	-0.04	-0.24	-0.15	0.36	0.42	-0.23	0.21	-0.38	0.00	-0.00	-0.46	-0.23	0.23	0.83	0.37	-0.02	-0.03	0.10	-0.01	1.00	
MissingTOEFL/IELTS	0.05	-0.07	0.03	0.16	0.11	-0.32	-0.22	0.16	-0.21	0.30	-0.00	-0.00	0.29	0.14	-0.04	-0.58	-0.20	0.05	-0.07	-0.05	0.12	-0.58	1.00
Observations	6222																						

Table 11: Elastic-Net Logistic Regression Results

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.3324	-0.3007	-0.3759	-0.2686	-0.1099
WinsorAge	-0.0769	0.0227	0.0235		0.0006
Female	0.3281	0.4176	0.4424		-0.3185
USAfrHispNat	0.2607	0.7009	0.6120		-0.3049
USAsian	-0.2566	-0.4079	-0.3659	0.2440	0.5626
IntAsian	-0.6105	-0.1800	0.1744	0.1584	-0.2834
IntOther	0.5106	0.1863	0.6030	-0.1492	0.1598
GREVerbalPct	0.0144	0.0190	0.0200	0.0009	0.0217
GREQuantPct	0.1248	0.0107	0.0144	0.0082	-0.0021
GREAnalyticPct	0.0161	0.0154	0.0053	0.0097	0.0029
UgradGPA	1.9459	0.5195	0.7539	0.4325	0.4685
TOEFL/IELTS	0.4013	0.4021	0.7215	0.5309	-0.1384
EliteUSUni	0.6536	0.3087	0.3937	-0.0267	0.5678
EliteUSLA	0.8353	0.5333	0.6518		0.8373
EliteForeign	0.4059	0.0256	0.3949	0.5323	0.4238
OtherForeign		0.0185	0.1569		-0.1016
GradDegree		0.0300	0.2005	0.0162	0.1647
WorkExper	0.2789	0.3206	0.3222		0.5590
#ProfRecom		-0.0179	0.3350	0.2201	0.0029
#ProlificRecom	0.2273	0.2140	0.1707	0.2159	0.2230
#MathCourses	0.1443				
AdvMath	0.2411				
MissingGRE	0.5471	-0.1982	0.0177	-0.2056	-0.5202
MissingUgradGPA		-0.1124	-0.1046		-0.0050
MissingTOEFL/IELTS	-0.2530	0.2993	0.2326		-0.2940
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486
Pseudo-R^2	0.2617	0.1418	0.1645	0.1504	0.1702
CV Pseudo-R^2	0.2280	0.0999	0.0778	0.0827	0.0567
α^*	0.8	0.0	0.0	0.2	0.0

Table 12: Post-Lasso Logistic Regression Results

(a) Coefficients

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.3971**	-0.4277	-0.5714*	-0.5742	
	(0.1989)	(0.2610)	(0.3424)	(0.3825)	
WinsorAge	-0.0981***	0.0483*	0.0440		
	(0.0333)	(0.0249)	(0.0323)		
Female	0.4127**	0.6628***	0.6605**		-0.3204
	(0.1668)	(0.2286)	(0.2928)		(0.4908)
USAfrHispNat	0.4317	0.9947***	1.0910***		
	(0.5163)	(0.3544)	(0.4169)		
USAsian	-0.4978	-0.5507		0.6986	0.7934
	(0.4124)	(0.6452)		(0.6907)	(0.6164)
IntAsian	-0.7555***	-0.1117		0.4011	
	(0.2655)	(0.3783)		(0.4516)	
IntOther	0.5780**	0.4017	1.0844**	-0.1716	
	(0.2828)	(0.3297)	(0.4353)	(0.4510)	
GREVerbalPct	0.0161***	0.0299**	0.0511***		0.0498***
	(0.0057)	(0.0146)	(0.0171)		(0.0185)
GREQuantPct	0.1505***	0.0151**	0.0298***	0.0156	
	(0.0176)	(0.0077)	(0.0099)	(0.0120)	
GREAnalyticPct	0.0173***	0.0215**		0.0158	
	(0.0041)	(0.0089)		(0.0121)	
UgradGPA	2.1546***	0.6897	1.2804**	0.9282	
	(0.6249)	(0.5477)	(0.5178)	(0.9753)	
TOEFL/IELTS	0.3894	0.6451	1.0434*	0.8607*	
	(0.2887)	(0.4645)	(0.5521)	(0.5143)	
EliteUSUni	0.7856***	0.3693	0.5764		0.7688
	(0.2305)	(0.2746)	(0.3516)		(0.4721)
EliteUSLA	0.9982***	0.6323*	1.0380**		1.3415
	(0.3347)	(0.3555)	(0.5138)		(0.8540)
EliteForeign	0.5045*		0.5997	0.9800	
	(0.2713)		(0.5682)	(0.5974)	
GradDegree			0.2882		
			(0.3219)		
WorkExper	0.3793**	0.4156	0.5011		1.6031**
	(0.1881)	(0.3000)	(0.3279)		(0.6992)
#ProfRecom			0.5701**	0.4703*	
			(0.2603)	(0.2698)	
#ProlificRecom	0.2778***	0.2966**	0.2160	0.3356**	0.4341
	(0.0843)	(0.1232)	(0.1595)	(0.1455)	(0.2885)
#MathCourses	0.1658***				
	(0.0592)				
AdvMath	0.2434				
	(0.1780)				
MissingGRE	1.3915			-0.3006	-0.9035
	(0.8977)			(0.3833)	(0.5562)
MissingTOEFL/IELTS	-0.4119*	0.5876	0.4731		
	(0.2295)	(0.4731)	(0.5151)		
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486
Pseudo-R^2	0.2656	0.1514	0.1878	0.1724	0.1844
CV Pseudo-R^2	0.2166	0.0908	0.0772	0.0114	0.0445

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) AMEs

DV = Accept	(1) Economics	(2) Government	(3) History	(4) Linguistics	(5) Psychology
CovidYear	-0.0305**	-0.0162	-0.0432*	-0.0315	
	(0.0152)	(0.0100)	(0.0260)	(0.0212)	
WinsorAge	-0.0075***	0.0018*	0.0033		
	(0.0025)	(0.0010)	(0.0024)		
Female	0.0317**	0.0251***	0.0499**		-0.0144
	(0.0128)	(0.0087)	(0.0219)		(0.0222)
USAfrHispNat	0.0331	0.0377***	0.0824***		
	(0.0396)	(0.0137)	(0.0316)		
USAsian	-0.0382	-0.0209		0.0383	0.0357
	(0.0317)	(0.0245)		(0.0382)	(0.0270)
IntAsian	-0.0580***	-0.0042		0.0220	
	(0.0203)	(0.0143)		(0.0249)	
IntOther	0.0444**	0.0152	0.0819**	-0.0094	
	(0.0217)	(0.0125)	(0.0330)	(0.0248)	
GREVerbalPct	0.0012***	0.0011**	0.0039***		0.0022**
	(0.0004)	(0.0006)	(0.0013)		(0.0009)
GREQuantPct	0.0116***	0.0006*	0.0023***	0.0009	
	(0.0013)	(0.0003)	(0.0007)	(0.0007)	
GREAnalyticPct	0.0013***	0.0008**		0.0009	
	(0.0003)	(0.0003)		(0.0007)	
UgradGPA	0.1654***	0.0261	0.0968**	0.0510	
	(0.0479)	(0.0208)	(0.0403)	(0.0536)	
TOEFL/IELTS	0.0299	0.0244	0.0788*	0.0472	
	(0.0220)	(0.0177)	(0.0420)	(0.0287)	
EliteUSUni	0.0603***	0.0140	0.0436*		0.0346
	(0.0176)	(0.0104)	(0.0262)		(0.0214)
EliteUSLA	0.0766***	0.0239*	0.0784**		0.0604
	(0.0254)	(0.0135)	(0.0386)		(0.0398)
EliteForeign	0.0387*		0.0453	0.0538	
	(0.0208)		(0.0425)	(0.0331)	
GradDegree			0.0218		
			(0.0244)		
WorkExper	0.0291**	0.0157	0.0379		0.0722**
	(0.0144)	(0.0114)	(0.0249)		(0.0335)
#ProfRecom			0.0431**	0.0258*	
			(0.0195)	(0.0147)	
#ProlificRecom	0.0213***	0.0112**	0.0163	0.0184**	0.0195
	(0.0064)	(0.0047)	(0.0121)	(0.0080)	(0.0127)
#MathCourses	0.0127***				
	(0.0045)				
AdvMath	0.0187				
	(0.0137)				
MissingGRE	0.1068			-0.0165	-0.0407
	(0.0691)			(0.0210)	(0.0253)
MissingTOEFL/IELTS	-0.0316*	0.0223	0.0357		
	(0.0178)	(0.0180)	(0.0388)		
Concentration FEs	No	Yes	Yes	Yes	Yes
Observations	2078	2231	690	737	486

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 13: Wald Test Summary Statistics

(a) Different Determinants

	Economics		Non-Economics 1		Non-Economics 2	
	Mean	SD	Mean	SD	Mean	SD
Accept	0.1335	0.3401	0.0452	0.2078	0.0543	0.2266
Success	0.3000	0.4583	0.3126	0.4636	0.1314	0.3379
GREQuantPct	84.8390	7.5405	65.1343	7.4495	65.6114	7.6706
Gender	0.3935	0.4885	0.3952	0.4889	0.5914	0.4916
Demographic1	0.5005	0.5000	0.3091	0.4621	0.2600	0.4386
Demographic2	0.2505	0.4333	0.2957	0.4563	0.3414	0.4742
AdvMath	0.5095	0.4999				
ComScore	4.9325	1.9877	6.0191	2.0205	6.0243	2.0038
Observations	2000		2300		700	

(b) Same Determinants

	Economics		Non-Economics 1		Non-Economics 2	
	Mean	SD	Mean	SD	Mean	SD
Accept	0.1335	0.3401	0.0900	0.2862	0.0786	0.2691
Success	0.3000	0.4583	0.4057	0.4910	0.1157	0.3199
GREQuantPct	84.8390	7.5405	65.1343	7.4495	65.6114	7.6706
Gender	0.3935	0.4885	0.3952	0.4889	0.5914	0.4916
Demographic1	0.5005	0.5000	0.3091	0.4621	0.2600	0.4386
Demographic2	0.2505	0.4333	0.2957	0.4563	0.3414	0.4742
AdvMath	0.5095	0.4999				
ComScore	4.9325	1.9877	6.0191	2.0205	6.0243	2.0038
Observations	2000		2300		700	

Table 14: Wald Test Correlation Matrix

(a) Different Determinants

	Acc	Suc	GRE	Gen	Dem1	Dem2	Com	NEcon1	NEcon2
Accept	1								
Success	0.0885	1							
GREQuantPct	0.2376	0.1167	1						
Gender	0.0582	-0.0220	-0.0795	1					
Demographic1	0.0152	-0.0004	0.3377	-0.0136	1				
Demographic2	0.0078	0.0003	-0.0876	-0.0209	-0.4918	1			
ComScore	0.1437	0.1500	-0.0770	0.0980	-0.1567	-0.0359	1		
NonEconomics1	-0.1232	0.0624	-0.6017	-0.0501	-0.1326	0.0238	0.1930	1	
NonEconomics2	-0.0405	-0.1352	-0.2472	0.1384	-0.0988	0.0514	0.0854	-0.3724	1
Observations	5000								

(b) Same Determinants

	Acc	Suc	GRE	Gen	Dem1	Dem2	Com	NEcon1	NEcon2
Accept	1								
Success	0.1603	1							
GREQuantPct	0.2772	0.1665	1						
Gender	0.0484	-0.0581	-0.0795	1					
Demographic1	-0.0233	0.0199	0.3377	-0.0136	1				
Demographic2	0.0487	0.0148	-0.0876	-0.0209	-0.4918	1			
ComScore	0.2045	0.1486	-0.0770	0.0980	-0.1567	-0.0359	1		
NonEconomics1	-0.0474	0.1636	-0.6017	-0.0501	-0.1326	0.0238	0.1930	1	
NonEconomics2	-0.0357	-0.1787	-0.2472	0.1384	-0.0988	0.0514	0.0854	-0.3724	1
Observations	5000								

C Additional Structural Results

Table 15: Structural Summary Statistics

(a) Suboptimal Committee

	Mean	SD
Transition	0.3910	0.4880
Accept	0.1350	0.3417
Success	0.3040	0.4600
GREQuantPct	84.9310	7.5286
Gender	0.3970	0.4893
Demographic1	0.5410	0.4983
Demographic2	0.2600	0.4386
AdvMath	0.5300	0.4991
ComScore	5.1470	2.0277
YearTransMore	0.5000	0.5000
Observations	1000	

(b) Optimal Committee

	Mean	SD
Transition	0.3940	0.4886
Accept	0.1400	0.3470
Success	0.3040	0.4600
GREQuantPct	84.9310	7.5286
Gender	0.3970	0.4893
Demographic1	0.5410	0.4983
Demographic2	0.2600	0.4386
AdvMath	0.5300	0.4991
ComScore	5.1470	2.0277
YearTransMore	0.5000	0.5000
Observations	1000	

Table 16: Structural Correlation Matrix

(a) Suboptimal Committee

	Trans	Acc	Suc	GRE	Fem	Asn	Oth	Math	Com	Year
Transition	1									
Accept	0.4930	1								
Success	0.1164	0.0379	1							
GREQuantPct	0.2550	0.2081	0.3225	1						
Gender	0.0200	0.0503	-0.0164	-0.0398	1					
Demographic1	-0.0351	-0.0472	0.0023	0.2688	-0.0032	1				
Demographic2	0.0670	0.0394	0.0642	-0.0579	-0.0476	-0.6435	1			
AdvMath	0.0894	0.1140	0.1868	0.2799	-0.0181	0.1941	-0.0996	1		
ComScore	0.0915	0.1662	0.1204	0.2029	0.0783	-0.0827	-0.0396	0.2372	1	
YearTransMore	0.1865	-0.0497	0.0217	-0.0410	-0.0225	-0.0341	0.0593	0.0521	0.0192	1
Observations	1000									

(b) Optimal Committee

	Trans	Acc	Suc	GRE	Fem	Asn	Oth	Math	Com	Year
Transition	1									
Accept	0.5004	1								
Success	0.1300	0.0905	1							
GREQuantPct	0.2876	0.3038	0.3225	1						
Gender	-0.0185	0.0084	-0.0164	-0.0398	1					
Demographic1	-0.0006	0.0189	0.0023	0.2688	-0.0032	1				
Demographic2	0.0493	0.0105	0.0642	-0.0579	-0.0476	-0.6435	1			
AdvMath	0.0991	0.1605	0.1868	0.2799	-0.0181	0.1941	-0.0996	1		
ComScore	0.0980	0.1271	0.1204	0.2029	0.0783	-0.0827	-0.0396	0.2372	1	
YearTransMore	0.1678	-0.0692	0.0217	-0.0410	-0.0225	-0.0341	0.0593	0.0521	0.0192	1
Observations	1000									

Table 17: Year-Dynamic Unrestricted Estimates

(a) Suboptimal Committee

	FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-6.1586***	-1.3913***	0.5480	-8.6113***	-1.8072***	-4.6409***	-0.8906***	-8.4142***	-1.7374***
	(0.8297)	(0.1654)	(0.7283)	(0.9390)	(0.1610)	(1.6387)	(0.2989)	(1.7756)	(0.3129)
GREQuantPct	0.0710***	0.0160***	0.0581***	0.0923***	0.0194***	0.0386**	0.0074**	0.0846***	0.0175***
	(0.0099)	(0.0020)	(0.0088)	(0.0110)	(0.0020)	(0.0191)	(0.0036)	(0.0211)	(0.0039)
Gender	-0.0381	-0.0086	0.3310***	0.1796	0.0377	0.3556	0.0682	-0.0468	-0.0097
	(0.1363)	(0.0308)	(0.1204)	(0.1418)	(0.0297)	(0.2307)	(0.0441)	(0.2303)	(0.0475)
Demographic1	0.5198***	0.1174***	-1.2830***	-0.4878**	-0.1024**	-0.5489	-0.1053*	-0.2506	-0.0517
	(0.1800)	(0.0400)	(0.1529)	(0.1990)	(0.0412)	(0.3363)	(0.0638)	(0.3422)	(0.0700)
Demographic2	-0.0415	-0.0094	-0.9475***	0.0313	0.0066	-0.0869	-0.0167	0.5529*	0.1142*
	(0.2000)	(0.0452)	(0.1759)	(0.2102)	(0.0441)	(0.3517)	(0.0675)	(0.3207)	(0.0659)
AdvMath			0.8895***			0.4169	0.0800*	0.6588***	0.1360***
			(0.1239)			(0.2549)	(0.0482)	(0.2507)	(0.0504)
ComScore						0.2201***	0.0422***	0.0048	0.0010
						(0.0666)	(0.0120)	(0.0584)	(0.0121)
YearTransMore				0.8992***	0.1887***	-1.1992***	-0.2301***		
				(0.1418)	(0.0273)	(0.2380)	(0.0398)		
$\sqrt{\text{MeanSqError}}$			1.8844***						
			(0.0398)						
Obs | Pseudo-R^2	1000	0.0596							
LL | Param | AIC	-3758.1	33	7582.1						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-6.1391***	-1.3875***	0.5474	-9.2064***	-1.9245***	-8.8857***	-1.6359***	-8.7036***	-1.8098***
	(0.8288)	(0.1654)	(0.7284)	(0.9500)	(0.1576)	(1.7825)	(0.2813)	(1.8073)	(0.3173)
GREQuantPct	0.0708***	0.0160***	0.0581***	0.0998***	0.0209***	0.0937***	0.0173***	0.0879***	0.0183***
	(0.0099)	(0.0020)	(0.0088)	(0.0112)	(0.0019)	(0.0204)	(0.0033)	(0.0213)	(0.0039)
Gender	-0.0384	-0.0087	0.3312***	0.0009	0.0002	0.1963	0.0361	-0.0426	-0.0089
	(0.1363)	(0.0308)	(0.1204)	(0.1424)	(0.0298)	(0.2429)	(0.0448)	(0.2313)	(0.0481)
Demographic1	0.5189***	0.1173***	-1.2807***	-0.3536*	-0.0739*	-0.4632	-0.0853	-0.3130	-0.0651
	(0.1800)	(0.0400)	(0.1529)	(0.2007)	(0.0416)	(0.3599)	(0.0659)	(0.3443)	(0.0707)
Demographic2	-0.0429	-0.0097	-0.9447***	0.0434	0.0091	-0.1716	-0.0316	0.5339	0.1110
	(0.1999)	(0.0452)	(0.1759)	(0.2140)	(0.0447)	(0.3749)	(0.0691)	(0.3263)	(0.0676)
AdvMath			0.8894***			0.7280***	0.1340***	0.5854**	0.1217**
			(0.1239)			(0.2611)	(0.0463)	(0.2519)	(0.0512)
ComScore						0.0908	0.0167	0.0227	0.0047
						(0.0623)	(0.0113)	(0.0574)	(0.0119)
YearTransMore				0.8263***	0.1727***	-1.2645***	-0.2328***		
				(0.1421)	(0.0276)	(0.2443)	(0.0379)		
$\sqrt{\text{MeanSqError}}$			1.8844***						
			(0.0398)						
Obs | Pseudo-R^2	1000	0.0628							
LL | Param | AIC	-3753.7	33	7573.4						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 18: Year-Dynamic Restricted Estimates

(a) Suboptimal Committee

	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-6.1931***	0.5166	-0.4477	0.6391	-0.6118***	35.17%	-8.4870	
	(0.8371)	(0.7436)	(0.6919)	(0.1647, 2.4802)	(0.1802)	(27.59%, 43.57%)	(12.5846)	
GREQuantPct	0.0715***	0.0586***					0.0799	
	(0.0100)	(0.0090)					(0.1379)	
Gender	-0.0410	0.3267***					0.1457	
	(0.1361)	(0.1202)					(0.1426)	
Demographic1	0.5215***	-1.2825***					-0.3775	
	(0.1802)	(0.1530)					(0.5271)	
Demographic2	-0.0349	-0.9447***					0.2483	
	(0.1999)	(0.1760)					(0.5059)	
AdvMath		0.8948***					0.4141	
		(0.1239)					(0.4431)	
ComScore							0.1340**	
							(0.0628)	
YearTransMore			-0.0096	0.6330	0.8913	56.94%		
			(0.0626)	(0.1709, 2.3444)	(1.3618)	(9.04%, 94.62%)		
$\sqrt{\text{MeanSqError}}$		1.8840***						
		(0.0399)						
Sigma								0.1819
								(0.2909)
Observations	1000							
LL | Param | AIC	-3785.0	24	7618.0					
LR Stat | DF | P-Val	53.8131	9	0.0000					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-6.1768***	0.5240	-0.5344***	0.5860	-0.5960***	35.52%	-8.7931***	
	(0.8229)	(0.7313)	(0.1440)	(0.4419, 0.7771)	(0.1541)	(28.94%, 42.70%)	(3.0648)	
GREQuantPct	0.0713***	0.0584***					0.0876***	
	(0.0099)	(0.0089)					(0.0334)	
Gender	-0.0391	0.3305***					0.0238	
	(0.1359)	(0.1204)					(0.1132)	
Demographic1	0.5230***	-1.2810***					-0.2855	
	(0.1803)	(0.1531)					(0.2036)	
Demographic2	-0.0371	-0.9441***					0.1727	
	(0.1999)	(0.1760)					(0.1999)	
AdvMath		0.8913***					0.5028**	
		(0.1239)					(0.1953)	
ComScore							0.0659*	
							(0.0384)	
YearTransMore			0.0245	0.6006	0.7803**	54.59%		
			(0.0515)	(0.4481, 0.8049)	(0.3189)	(40.35%, 68.12%)		
$\sqrt{\text{MeanSqError}}$		1.8849***						
		(0.0398)						
Sigma								0.1473***
								(0.0541)
Observations	1000							
LL | Param | AIC	-3773.5	24	7595.0					
LR Stat | DF | P-Val	39.6494	9	0.0000					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Table 19: Year-Dynamic Unrestricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

	FStage1		FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-5.1788***	-1.7932***	-2.9927*	-1.0548*	2.7507	-8.5772***	-1.8014***	-4.8978***	-0.9362***	-8.2901***	-1.7168***
	(0.5553)	(0.1624)	(1.7053)	(0.5911)	(2.3336)	(0.9369)	(0.1610)	(1.6572)	(0.2992)	(1.7618)	(0.3127)
GREQuantPct	0.0554***	0.0192***	0.0409**	0.0144**	0.0355	0.0918***	0.0193***	0.0415**	0.0079**	0.0832***	0.0172***
	(0.0066)	(0.0020)	(0.0167)	(0.0057)	(0.0226)	(0.0110)	(0.0020)	(0.0193)	(0.0036)	(0.0209)	(0.0039)
Gender	0.1094	0.0379	-0.1976	-0.0696	0.1937	0.1789	0.0376	0.3418	0.0653	-0.0589	-0.0122
	(0.0858)	(0.0296)	(0.1380)	(0.0482)	(0.1947)	(0.1417)	(0.0297)	(0.2318)	(0.0441)	(0.2296)	(0.0475)
Demographic1	-0.2859**	-0.0990**	0.2982	0.1051	-1.1999***	-0.4842**	-0.1017**	-0.5561	-0.1063*	-0.2415	-0.0500
	(0.1186)	(0.0406)	(0.1976)	(0.0691)	(0.2734)	(0.1988)	(0.0412)	(0.3383)	(0.0638)	(0.3411)	(0.0700)
Demographic2	0.0241	0.0083	-0.1464	-0.0516	-0.9202***	0.0339	0.0071	-0.0826	-0.0158	0.5526*	0.1144*
	(0.1266)	(0.0438)	(0.2000)	(0.0703)	(0.2915)	(0.2101)	(0.0441)	(0.3531)	(0.0675)	(0.3201)	(0.0660)
AdvMath					0.8786***			0.4132	0.0790	0.6481***	0.1342***
					(0.2057)			(0.2558)	(0.0482)	(0.2499)	(0.0504)
ComScore								0.2220***	0.0424***	0.0053	0.0011
								(0.0668)	(0.0120)	(0.0582)	(0.0121)
YearTransMore	0.5446***	0.1886***				0.8980***	0.1886***	-1.2072***	-0.2307***		
	(0.0850)	(0.0274)				(0.1416)	(0.0273)	(0.2391)	(0.0397)		
InvMills			-0.4189	-0.1476	-0.1894						
			(0.3872)	(0.1361)	(0.5612)						
$\sqrt{\text{MeanSqError}}$					1.8925***						
					(0.0662)						
Obs | Pseudo R-^{2}	1000	0.0760									
LL | Param | AIC	-2715.9	41	5513.9								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FStage1		FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-5.5333***	-1.9103***	-2.6740	-0.9412	2.9426	-9.1587***	-1.9166***	-9.4836***	-1.7254***	-8.2448***	-1.7310***
	(0.5668)	(0.1614)	(1.9361)	(0.6738)	(2.6765)	(0.9470)	(0.1576)	(1.8420)	(0.2807)	(1.7584)	(0.3170)
GREQuantPct	0.0599***	0.0207***	0.0373**	0.0131**	0.0342	0.0992***	0.0208***	0.0998***	0.0182***	0.0824***	0.0173***
	(0.0067)	(0.0020)	(0.0188)	(0.0065)	(0.0258)	(0.0111)	(0.0019)	(0.0210)	(0.0033)	(0.0208)	(0.0039)
Gender	0.0025	0.0009	-0.1822	-0.0641	0.1863	0.0002	0.0000	0.1969	0.0358	-0.0434	-0.0091
	(0.0860)	(0.0297)	(0.1380)	(0.0482)	(0.1918)	(0.1422)	(0.0298)	(0.2459)	(0.0448)	(0.2292)	(0.0481)
Demographic1	-0.2056*	-0.0710*	0.2973	0.1047	-1.2445***	-0.3519*	-0.0736*	-0.4210	-0.0766	-0.2654	-0.0557
	(0.1193)	(0.0409)	(0.1954)	(0.0683)	(0.2697)	(0.2004)	(0.0416)	(0.3655)	(0.0662)	(0.3411)	(0.0709)
Demographic2	0.0312	0.0108	-0.1242	-0.0437	-1.0245***	0.0433	0.0091	-0.1066	-0.0194	0.5579*	0.1171*
	(0.1280)	(0.0442)	(0.2051)	(0.0721)	(0.2997)	(0.2137)	(0.0447)	(0.3807)	(0.0693)	(0.3250)	(0.0679)
AdvMath					0.8964***			0.7077***	0.1288***	0.5718**	0.1200**
					(0.2088)			(0.2634)	(0.0463)	(0.2495)	(0.0513)
ComScore								0.0968	0.0176	0.0233	0.0049
								(0.0631)	(0.0113)	(0.0570)	(0.0120)
YearTransMore	0.4989***	0.1722***				0.8249***	0.1726***	-1.2813***	-0.2331***		
	(0.0852)	(0.0277)				(0.1419)	(0.0276)	(0.2479)	(0.0378)		
InvMills			-0.4587	-0.1615	-0.2035						
			(0.4255)	(0.1495)	(0.6179)						
$\sqrt{\text{MeanSqError}}$					1.9059***						
					(0.0667)						
Obs | Pseudo R-^{2}	1000	0.0811									
LL | Param | AIC	-2720.3	41	5522.6								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 20: Year-Dynamic Restricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

	FStage1	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-5.1849***	-3.0077*	2.7390	-0.4529	0.6358	-0.6156***	35.08%	-8.3629	
	(0.5600)	(1.7125)	(2.3203)	(0.6380)	(0.1821, 2.2201)	(0.1808)	(27.49%, 43.51%)	(11.5276)	
GREQuantPct	0.0554***	0.0418**	0.0363					0.0786	
	(0.0066)	(0.0168)	(0.0226)					(0.1266)	
Gender	0.1106	-0.1963	0.1960					0.1511	
	(0.0858)	(0.1363)	(0.1938)					(0.1374)	
Demographic1	-0.2846**	0.2917	-1.2040***					-0.3901	
	(0.1188)	(0.1960)	(0.2717)					(0.5285)	
Demographic2	0.0225	-0.1595	-0.9260***					0.2394	
	(0.1268)	(0.1973)	(0.2899)					(0.4445)	
AdvMath			0.8890***					0.4501	
			(0.2057)					(0.4458)	
ComScore								0.1300**	
								(0.0579)	
YearTransMore	0.5474***			-0.0042	0.6331	0.8901	56.82%		
	(0.0861)			(0.0616)	(0.1890, 2.1209)	(1.2635)	(10.79%, 93.47%)		
InvMills		-0.4754	-0.2427						
		(0.3913)	(0.5594)						
$\sqrt{\text{MeanSqError}}$			1.8935***						
			(0.0666)						
Sigma									0.1805
									(0.2674)
Observations	1000								
LL | Param | AIC	-2742.5	32	5549.0						
LR Stat | DF | P-Val	53.1755	9	0.0000						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	FStage1	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-5.5371***	-2.6870	2.9441	-0.5459***	0.5793	-0.5863***	35.75%	-8.3634***	
	(0.5732)	(1.9557)	(2.6682)	(0.1161)	(0.4615, 0.7273)	(0.1500)	(29.31%, 42.74%)	(2.3193)	
GREQuantPct	0.0599***	0.0382**	0.0345					0.0826***	
	(0.0068)	(0.0190)	(0.0257)					(0.0250)	
Gender	0.0032	-0.1802	0.1850					0.0447	
	(0.0860)	(0.1357)	(0.1912)					(0.1096)	
Demographic1	-0.2061*	0.2902	-1.2455***					-0.2408	
	(0.1194)	(0.1936)	(0.2693)					(0.1838)	
Demographic2	0.0295	-0.1360	-1.0281***					0.1880	
	(0.1281)	(0.2022)	(0.2992)					(0.1935)	
AdvMath			0.9026***					0.4942***	
			(0.2088)					(0.1764)	
ComScore								0.0643*	
								(0.0372)	
YearTransMore	0.4995***			0.0281	0.5958	0.7546***	54.20%		
	(0.0857)			(0.0503)	(0.4697, 0.7558)	(0.2627)	(42.41%, 65.54%)		
InvMills		-0.5095	-0.2301						
		(0.4343)	(0.6188)						
$\sqrt{\text{MeanSqError}}$			1.9062***						
			(0.0667)						
Sigma									0.1418***
									(0.0414)
Observations	1000								
LL | Param | AIC	-2739.7	32	5543.4						
LR Stat | DF | P-Val	38.8065	9	0.0000						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Table 21: Year-Myopic Unrestricted Estimates

(a) Suboptimal Committee

	FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	0.1201*	0.5300***	4.6362***	-8.5778***	-1.8014***	-4.7827***	-0.9153***	-8.3592***	-1.7287***
	(0.0634)	(0.0158)	(0.0879)	(0.9369)	(0.1610)	(1.6494)	(0.2990)	(1.7684)	(0.3127)
GREQuantPct				0.0919***	0.0193***	0.0401**	0.0077**	0.0839***	0.0174***
				(0.0110)	(0.0020)	(0.0193)	(0.0036)	(0.0210)	(0.0039)
Gender				0.1788	0.0376	0.3374	0.0646	-0.0594	-0.0123
				(0.1417)	(0.0297)	(0.2315)	(0.0442)	(0.2300)	(0.0476)
Demographic1				-0.4846**	-0.1018**	-0.5415	-0.1036	-0.2355	-0.0487
				(0.1988)	(0.0412)	(0.3380)	(0.0639)	(0.3420)	(0.0701)
Demographic2				0.0335	0.0070	-0.0695	-0.0133	0.5612*	0.1161*
				(0.2101)	(0.0441)	(0.3531)	(0.0676)	(0.3209)	(0.0661)
AdvMath			0.9638***			0.4166	0.0797*	0.6464***	0.1337***
			(0.1243)			(0.2555)	(0.0482)	(0.2502)	(0.0504)
ComScore						0.2219***	0.0425***	0.0059	0.0012
						(0.0668)	(0.0120)	(0.0583)	(0.0121)
YearTransMore				0.8980***	0.1886***	-1.2106***	-0.2317***		
				(0.1416)	(0.0273)	(0.2388)	(0.0397)		
$\sqrt{\text{MeanSqError}}$			1.9698***						
			(0.0413)						
Obs | Pseudo-R^2	1000	0.0363							
LL | Param | AIC	-3851.0	25	7752.1						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	0.1201*	0.5300***	4.6362***	-9.1582***	-1.9165***	-9.2354***	-1.6874***	-8.4285***	-1.7639***
	(0.0634)	(0.0158)	(0.0879)	(0.9469)	(0.1576)	(1.8175)	(0.2808)	(1.7765)	(0.3171)
GREQuantPct				0.0992***	0.0208***	0.0972***	0.0178***	0.0846***	0.0177***
				(0.0111)	(0.0019)	(0.0207)	(0.0033)	(0.0210)	(0.0039)
Gender				0.0002	0.0000	0.1906	0.0348	-0.0419	-0.0088
				(0.1422)	(0.0298)	(0.2448)	(0.0448)	(0.2299)	(0.0481)
Demographic1				-0.3522*	-0.0737*	-0.4229	-0.0773	-0.2773	-0.0580
				(0.2004)	(0.0416)	(0.3637)	(0.0662)	(0.3423)	(0.0709)
Demographic2				0.0431	0.0090	-0.1152	-0.0211	0.5521*	0.1155*
				(0.2137)	(0.0447)	(0.3789)	(0.0693)	(0.3257)	(0.0678)
AdvMath			0.9638***			0.7138***	0.1304***	0.5680**	0.1189**
			(0.1243)			(0.2625)	(0.0463)	(0.2501)	(0.0513)
ComScore						0.0953	0.0174	0.0237	0.0050
						(0.0628)	(0.0113)	(0.0572)	(0.0119)
YearTransMore				0.8248***	0.1726***	-1.2820***	-0.2342***		
				(0.1419)	(0.0276)	(0.2467)	(0.0378)		
$\sqrt{\text{MeanSqError}}$			1.9698***						
			(0.0413)						
Obs | Pseudo-R^2	1000	0.0396							
LL | Param | AIC	-3846.6	25	7743.2						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 22: Year-Myopic Restricted Estimates

(a) Suboptimal Committee

	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	0.1291**	4.6418***	-0.4609	0.6307	-0.6100***	35.21%	-8.4650	
	(0.0636)	(0.0884)	(0.6357)	(0.1814, 2.1926)	(0.1827)	(32.88%, 37.60%)	(10.9416)	
GREQuantPct							0.0819	
							(0.1199)	
Gender							0.1398	
							(0.1284)	
Demographic1							-0.4103	
							(0.4518)	
Demographic2							0.2056	
							(0.4868)	
AdvMath		0.9652***					0.3557	
		(0.1243)					(0.4351)	
ComScore							0.1123**	
							(0.0529)	
YearTransMore			-0.0046	0.6279	0.8797	56.70%		
			(0.0615)	(0.1876, 2.1019)	(1.2631)	(10.93%, 93.32%)		
$\sqrt{\text{MeanSqError}}$		1.9675***						
		(0.0412)						
Sigma								0.1803
								(0.2674)
Observations	1000							
LL | Param | AIC	-3881.2	16	7794.4					
LR Stat | DF | P-Val	60.3013	9	0.0000					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	0.1317**	4.6389***	-0.5593***	0.5716	-0.5909***	35.64%	-8.5595***	
	(0.0635)	(0.0879)	(0.1114)	(0.4595, 0.7111)	(0.1494)	(34.08%, 37.24%)	(2.1793)	
GREQuantPct							0.0869***	
							(0.0239)	
Gender							0.0343	
							(0.1085)	
Demographic1							-0.2998*	
							(0.1787)	
Demographic2							0.1207	
							(0.1762)	
AdvMath		0.9645***					0.4241**	
		(0.1244)					(0.1693)	
ComScore							0.0498	
							(0.0351)	
YearTransMore			0.0304	0.5892	0.7487***	53.94%		
			(0.0510)	(0.4695, 0.7394)	(0.2510)	(42.65%, 64.83%)		
$\sqrt{\text{MeanSqError}}$		1.9718***						
		(0.0414)						
Sigma								0.1403***
								(0.0391)
Observations	1000							
LL | Param | AIC	-3869.1	16	7770.3					
LR Stat | DF | P-Val	45.0826	9	0.0000					

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Table 23: Year-Myopic Unrestricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

	FStage1		FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-0.5244***	-0.1957***	0.3641	0.1419	4.8485***	-8.5775***	-1.8014***	-4.7843***	-0.9156***	-8.3574***	-1.7284***
	(0.0589)	(0.0192)	(0.3839)	(0.1490)	(0.5941)	(0.9369)	(0.1610)	(1.6495)	(0.2990)	(1.7682)	(0.3127)
GREQuantPct						0.0919***	0.0193***	0.0401**	0.0077**	0.0839***	0.0173***
						(0.0110)	(0.0020)	(0.0193)	(0.0036)	(0.0210)	(0.0039)
Gender						0.1788	0.0376	0.3374	0.0646	-0.0592	-0.0122
						(0.1417)	(0.0297)	(0.2315)	(0.0442)	(0.2300)	(0.0476)
Demographic1						-0.4842**	-0.1017**	-0.5415	-0.1036	-0.2351	-0.0486
						(0.1988)	(0.0412)	(0.3380)	(0.0639)	(0.3420)	(0.0701)
Demographic2						0.0339	0.0071	-0.0695	-0.0133	0.5614*	0.1161*
						(0.2101)	(0.0441)	(0.3531)	(0.0676)	(0.3210)	(0.0661)
AdvMath					0.9097***			0.4165	0.0797*	0.6464***	0.1337***
					(0.1967)			(0.2555)	(0.0482)	(0.2502)	(0.0504)
ComScore								0.2219***	0.0425***	0.0059	0.0012
								(0.0668)	(0.0120)	(0.0583)	(0.0121)
YearTransMore	0.4793***	0.1788***				0.8980***	0.1886***	-1.2106***	-0.2317***		
	(0.0814)	(0.0287)				(0.1416)	(0.0273)	(0.2388)	(0.0397)		
InvMills			-0.1547	-0.0603	-0.0030						
			(0.3964)	(0.1543)	(0.6083)						
$\sqrt{\text{MeanSqError}}$					1.9491***						
					(0.0676)						
Obs | Pseudo R-^{2}	1000	0.0489									
LL | Param | AIC	-2795.4	29	5648.9								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FStage1		FAdvMath		FComScore	Transition		Accept		Success	
	Coef	AME	Coef	AME	Coef	Coef	AME	Coef	AME	Coef	AME
Intercept	-0.4902***	-0.1843***	0.4285	0.1664	4.9569***	-9.1577***	-1.9164***	-9.2245***	-1.6855***	-8.4374***	-1.7657***
	(0.0586)	(0.0195)	(0.4250)	(0.1643)	(0.6578)	(0.9469)	(0.1576)	(1.8167)	(0.2808)	(1.7772)	(0.3171)
GREQuantPct						0.0992***	0.0208***	0.0971***	0.0177***	0.0847***	0.0177***
						(0.0111)	(0.0019)	(0.0207)	(0.0033)	(0.0210)	(0.0039)
Gender						0.0002	0.0000	0.1897	0.0347	-0.0412	-0.0086
						(0.1422)	(0.0298)	(0.2448)	(0.0448)	(0.2299)	(0.0481)
Demographic1						-0.3519*	-0.0736*	-0.4248	-0.0776	-0.2785	-0.0583
						(0.2004)	(0.0416)	(0.3637)	(0.0662)	(0.3423)	(0.0709)
Demographic2						0.0434	0.0091	-0.1174	-0.0214	0.5504*	0.1152*
						(0.2137)	(0.0447)	(0.3788)	(0.0693)	(0.3256)	(0.0678)
AdvMath					0.9270***			0.7136***	0.1304***	0.5665**	0.1185**
					(0.1981)			(0.2626)	(0.0463)	(0.2501)	(0.0513)
ComScore								0.0952	0.0174	0.0237	0.0050
								(0.0628)	(0.0113)	(0.0572)	(0.0119)
YearTransMore	0.4300***	0.1617***				0.8248***	0.1726***	-1.2833***	-0.2345***		
	(0.0811)	(0.0292)				(0.1419)	(0.0276)	(0.2467)	(0.0378)		
InvMills			-0.2068	-0.0803	-0.1171						
			(0.4399)	(0.1706)	(0.6787)						
$\sqrt{\text{MeanSqError}}$					1.9628***						
					(0.0678)						
Obs | Pseudo R-^{2}	1000	0.0522									
LL | Param | AIC	-2805.9	29	5669.8								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 24: Year-Myopic Restricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

	FStage1	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-0.5219***	0.3660	4.8574***	-0.4584	0.6323	-0.6106***	35.19%	-8.4597	
	(0.0589)	(0.3889)	(0.6005)	(0.4992)	(0.2377, 1.6820)	(0.1783)	(27.68%, 43.51%)	(8.6873)	
GREQuantPct								0.0814	
								(0.0950)	
Gender								0.1396	
								(0.1268)	
Demographic1								-0.4047	
								(0.3758)	
Demographic2								0.2066	
								(0.3980)	
AdvMath			0.9145***					0.3654	
			(0.1964)					(0.3542)	
ComScore								0.1172**	
								(0.0505)	
YearTransMore	0.4753***			-0.0079	0.6274	0.8704	56.46%		
	(0.0813)			(0.0590)	(0.2422, 1.6247)	(0.9919)	(16.77%, 89.30%)		
InvMills		-0.1441	0.0077						
		(0.4020)	(0.6141)						
$\sqrt{\text{MeanSqError}}$			1.9501***						
			(0.0677)						
Sigma									0.1785
									(0.2089)
Observations	1000								
LL | Param | AIC	-2824.9	20	5689.7						
LR Stat | DF | P-Val	58.8213	9	0.0000						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	FStage1	FAdvMath	FComScore	Cutoff1	$L(\text{Cutoff})_1$	Cutoff2	$P(\text{Cutoff})_2$	Success	Sigma
Intercept	-0.4903***	0.4347	4.9623***	-0.5519***	0.5758	-0.5852***	35.77%	-8.5581***	
	(0.0585)	(0.4277)	(0.6606)	(0.1068)	(0.4671, 0.7099)	(0.1489)	(29.38%, 42.72%)	(2.0776)	
GREQuantPct								0.0861***	
								(0.0226)	
Gender								0.0384	
								(0.1075)	
Demographic1								-0.2591	
								(0.1750)	
Demographic2								0.1470	
								(0.1796)	
AdvMath			0.9310***					0.4198**	
			(0.1981)					(0.1642)	
ComScore								0.0566	
								(0.0356)	
YearTransMore	0.4302***			0.0282	0.5923	0.7481***	54.06%		
	(0.0810)			(0.0501)	(0.4762, 0.7368)	(0.2432)	(43.09%, 64.65%)		
InvMills		-0.1978	-0.1119						
		(0.4424)	(0.6804)						
$\sqrt{\text{MeanSqError}}$			1.9632***						
			(0.0678)						
Sigma									0.1401***
									(0.0370)
Observations	1000								
LL | Param | AIC	-2827.7	20	5695.3						
LR Stat | DF | P-Val	43.5305	9	0.0000						

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Table 25: Waitlist Summary Statistics

(a) Suboptimal Committee

	Mean	SD
Transition	0.3700	0.4828
Waitlist	0.3120	0.4633
Accept	0.1510	0.3580
Matriculate	0.0310	0.1733
Success	0.3100	0.4625
GREQuantPct	84.9310	7.5286
Gender	0.3970	0.4893
Demographic1	0.5500	0.4975
Demographic2	0.2620	0.4397
AdvMath	0.5300	0.4991
ComScore	5.1470	2.0277
MatScore	4.9370	2.2426
YearTransMat	0.5000	0.5000
Observations	1000	

(b) Optimal Committee

	Mean	SD
Transition	0.3650	0.4814
Waitlist	0.3240	0.4680
Accept	0.1500	0.3571
Matriculate	0.0390	0.1936
Success	0.3150	0.4645
GREQuantPct	84.9310	7.5286
Gender	0.3970	0.4893
Demographic1	0.5500	0.4975
Demographic2	0.2620	0.4397
AdvMath	0.5300	0.4991
ComScore	5.1470	2.0277
MatScore	4.9370	2.2426
YearTransMat	0.5000	0.5000
Observations	1000	

Table 26: Waitlist Correlation Matrix

(a) Suboptimal Committee

	Trans	Wait	Acc	IM	Suc	GRE	Gen	Dem1	Dem2	Math	Com	Mat	Year
Transition	1												
Watilist	0.8787	1											
Accept	0.5503	0.2766	1										
Matriculate	0.2334	-0.0084	0.4241	1									
Success	0.0685	0.0760	-0.0049	-0.0700	1								
GREQuantPct	0.3041	0.2745	0.2057	0.0675	0.2520	1							
Gender	0.0428	0.0138	0.0288	-0.0390	-0.0533	-0.0398	1						
Demographic1	-0.0146	0.0017	-0.0340	0.0574	0.0413	0.2483	0.0068	1					
Demographic2	0.0804	0.0553	0.0472	-0.0016	-0.0208	-0.0918	-0.0605	-0.6587	1				
AdvMath	0.0660	0.0417	0.0502	0.0413	0.1243	0.2799	-0.0181	0.1953	-0.1042	1			
ComScore	0.0364	0.0268	0.0493	-0.1154	0.1146	0.2029	0.0783	-0.0960	-0.0869	0.2372	1		
MatScore	0.0622	0.0420	-0.0778	0.1774	-0.0496	-0.0945	-0.0383	0.0562	0.0370	-0.0479	-0.3571	1	
YearTransMat	0.1947	0.2029	-0.1480	0.0058	0.0000	-0.0410	-0.0225	-0.0241	0.0318	0.0521	0.0192	0.4383	1
Observations	1000												

(b) Optimal Committee

	Trans	Wait	Acc	IM	Suc	GRE	Gen	Dem1	Dem2	Math	Com	Mat	Year
Transition	1												
Watilist	0.9131	1											
Accept	0.5541	0.3614	1										
Matriculate	0.2657	0.0923	0.4796	1									
Success	0.0448	0.0135	0.0528	0.0302	1								
GREQuantPct	0.2503	0.1908	0.2389	0.0801	0.3210	1							
Gender	-0.0038	0.0191	-0.0489	-0.0262	-0.0310	-0.0398	1						
Demographic1	0.0136	0.0034	0.0478	0.0161	-0.0444	0.2483	0.0068	1					
Demographic2	0.0065	0.0005	-0.0146	0.0327	0.0513	-0.0918	-0.0605	-0.6587	1				
AdvMath	0.0647	0.0397	0.0645	0.0034	0.0994	0.2799	-0.0181	0.1953	-0.1042	1			
ComScore	0.0813	0.0710	0.0704	-0.1063	0.1557	0.2029	0.0783	-0.0960	-0.0869	0.2372	1		
MatScore	0.0167	-0.0215	-0.0869	0.2130	-0.0280	-0.0945	-0.0383	0.0562	0.0370	-0.0479	-0.3571	1	
YearTransMat	0.1641	0.1624	-0.1736	0.0052	-0.0624	-0.0410	-0.0225	-0.0241	0.0318	0.0521	0.0192	0.4383	1
Observations	1000												

Table 27: Waitlist-Dynamic Unrestricted Estimates

(a) Suboptimal Committee

	FAdvMath		FComScore	F$\dot{m}_2$	F$\overline{m_{2A}}$	Transition		Accept2		Accept3		Success	
	Coef	AME	Coef	Coef	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-6.2079***	-1.4002***	0.9526	5.4979***	0.1489***	-10.7210***	-2.1071***	1.0375	0.1332	3.6817	0.4845	-8.2484***	-1.7415***
	(0.8279)	(0.1646)	(0.7367)	(0.3634)	(0.0018)	(1.0460)	(0.1548)	(2.3164)	(0.2962)	(2.3327)	(0.3079)	(1.7721)	(0.3282)
GREQuantPct	0.0707***	0.0160***	0.0566***	-0.0082*		0.1109***	0.0218***	-0.0180	-0.0023	0.0471*	0.0062*	0.0844***	0.0178***
	(0.0099)	(0.0020)	(0.0088)	(0.0045)		(0.0119)	(0.0019)	(0.0253)	(0.0032)	(0.0271)	(0.0034)	(0.0210)	(0.0040)
Gender	-0.0397	-0.0090	0.3124***	-0.0468		0.3479**	0.0684**	0.3888	0.0499	-0.7248**	-0.0954**	-0.0186	-0.0039
	(0.1367)	(0.0308)	(0.1188)	(0.0637)		(0.1474)	(0.0286)	(0.2995)	(0.0386)	(0.3422)	(0.0442)	(0.2390)	(0.0505)
Demographic1	0.6068***	0.1369***	-1.5750***	0.0087		-0.1970	-0.0387	-0.1599	-0.0205	-1.1811**	-0.1554**	-0.0753	-0.0159
	(0.1818)	(0.0401)	(0.1549)	(0.0902)		(0.2044)	(0.0401)	(0.4602)	(0.0592)	(0.5507)	(0.0706)	(0.3793)	(0.0801)
Demographic2	0.0513	0.0116	-1.3568***	0.0814		0.4583**	0.0901**	0.2797	0.0359	-1.0884**	-0.1432**	-0.1579	-0.0333
	(0.2027)	(0.0457)	(0.1784)	(0.0969)		(0.2215)	(0.0431)	(0.4728)	(0.0606)	(0.5058)	(0.0661)	(0.3844)	(0.0811)
AdvMath			0.9118***	0.1320**				0.4780	0.0613	-0.0258	-0.0034	0.1268	0.0268
			(0.1212)	(0.0669)				(0.3239)	(0.0415)	(0.3566)	(0.0470)	(0.2574)	(0.0543)
ComScore				0.0112				0.0051	0.0007	0.1221	0.0161	0.0349	0.0074
				(0.0166)				(0.0676)	(0.0087)	(0.0787)	(0.0103)	(0.0579)	(0.0122)
YearTransMat						1.0061***	0.1977***						
						(0.1492)	(0.0264)						
$\dot{m}_2$					1.0524***			-0.3224**	-0.0414**				
					(0.0003)			(0.1515)	(0.0192)				
$\overline{m_{2A}}$										-1.4912***	-0.1962***		
										(0.1656)	(0.0081)		
$\sqrt{\text{MeanSqError}}$			1.8530***	0.9790***	0.0041***								
			(0.0392)	(0.0022)	(0.0002)								
Obs | Pseudo-R^2	1000	0.3899											
LL | Param | AIC	-3656.6	52	7417.1										

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FAdvMath		FComScore	F$\dot{m}_2$	F$\overline{m_{2A}}$	Transition		Accept2		Accept3		Success	
	Coef	AME	Coef	Coef	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-6.4278***	-1.4433***	0.9765	5.5054***	-1.2247***	-7.8688***	-1.6527***	-5.4487**	-0.5308**	-0.1602	-0.0213	-10.4819***	-2.0411***
	(0.8383)	(0.1646)	(0.7368)	(0.3634)	(0.0038)	(0.9204)	(0.1634)	(2.3463)	(0.2269)	(2.0596)	(0.2735)	(1.8211)	(0.2882)
GREQuantPct	0.0732***	0.0164***	0.0564***	-0.0083*		0.0825***	0.0173***	0.0465	0.0045	0.0621**	0.0082**	0.1047***	0.0204***
	(0.0100)	(0.0020)	(0.0088)	(0.0045)		(0.0107)	(0.0020)	(0.0287)	(0.0028)	(0.0266)	(0.0034)	(0.0211)	(0.0035)
Gender	-0.0348	-0.0078	0.3118***	-0.0469		0.0553	0.0116	-0.5162	-0.0503	-0.1677	-0.0223	-0.0423	-0.0082
	(0.1373)	(0.0308)	(0.1188)	(0.0637)		(0.1410)	(0.0296)	(0.3697)	(0.0362)	(0.3034)	(0.0404)	(0.2493)	(0.0486)
Demographic1	0.6128***	0.1376***	-1.5764***	0.0091		-0.2800	-0.0588	0.2920	0.0284	0.4073	0.0541	-0.5056	-0.0984
	(0.1826)	(0.0401)	(0.1549)	(0.0902)		(0.1945)	(0.0407)	(0.6142)	(0.0600)	(0.4949)	(0.0654)	(0.3708)	(0.0716)
Demographic2	0.0632	0.0142	-1.3590***	0.0816		-0.0781	-0.0164	0.5062	0.0493	0.1398	0.0186	0.5141	0.1001
	(0.2037)	(0.0457)	(0.1784)	(0.0969)		(0.2140)	(0.0449)	(0.6069)	(0.0592)	(0.5084)	(0.0675)	(0.3855)	(0.0743)
AdvMath			0.9127***	0.1322**				0.3109	0.0303	-0.0573	-0.0076	-0.1314	-0.0256
			(0.1212)	(0.0669)				(0.3770)	(0.0367)	(0.3646)	(0.0484)	(0.2671)	(0.0520)
ComScore				0.0112				-0.0207	-0.0020	0.0648	0.0086	0.1494**	0.0291**
				(0.0166)				(0.0991)	(0.0096)	(0.0888)	(0.0118)	(0.0692)	(0.0133)
YearTransMat						0.7931***	0.1666***						
						(0.1403)	(0.0277)						
$\dot{m}_2$					1.4277***			-0.1868	-0.0182				
					(0.0007)			(0.1655)	(0.0158)				
$\overline{m_{2A}}$										-1.0984***	-0.1459***		
										(0.1172)	(0.0034)		
$\sqrt{\text{MeanSqError}}$			1.8530***	0.9790***	0.0057***								
			(0.0392)	(0.0022)	(0.0005)								
Obs | Pseudo-R^2	1000	0.3770											
LL | Param | AIC	-3786.1	52	7676.3										

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 28: Waitlist-Dynamic Restricted Estimates

(a) Suboptimal Committee

	FAdvMath	FComScore	F$\ddot{m}_2$	F$\overline{m_{2A}}$	Cutoff1	L(Cutoff)$_1$	Cutoff3	P(Cutoff)$_3$	Success	Beta	Sigma
Intercept	-6.2714***	0.8833	5.4571***	0.1489***	0.3481***	1.4164	-15.2416***	13.00%	-8.5337*		
	(0.8340)	(0.7393)	(0.3633)	(0.0018)	(0.1172)	(1.1258, 1.7820)	(3.1088)	(3.20%, 40.28%)	(4.8648)		
GREQuantPct	0.0715***	0.0575***	-0.0076*						0.0886		
	(0.0099)	(0.0089)	(0.0045)						(0.0549)		
Gender	-0.0377	0.3141***	-0.0582						0.1445		
	(0.1366)	(0.1187)	(0.0637)						(0.2817)		
Demographic1	0.6097***	-1.5742***	0.0121						-0.4893*		
	(0.1817)	(0.1549)	(0.0902)						(0.2779)		
Demographic2	0.0603	-1.3544***	0.0733						-0.2386		
	(0.2024)	(0.1782)	(0.0969)						(0.3301)		
AdvMath		0.9105***	0.1148*						0.3908		
		(0.1212)	(0.0669)						(0.2501)		
ComScore			0.0124						0.0421		
			(0.0166)						(0.0642)		
YearTransMat					-0.2401***	1.1140		98.90%			
					(0.0613)	(0.9659, 1.2848)		(94.76%, 99.78%)			
$\ddot{m}_2$				1.0524***							
				(0.0003)							
$\overline{m_{2A}}$							3.0950***				
							(0.5798)				
$\sqrt{\text{MeanSqError}}$		1.8548***	0.9791***	0.0041***							
		(0.0394)	(0.0022)	(0.0002)							
Beta										0.9506***	
										(0.2236)	
Sigma											0.3633***
											(0.0992)
Observations	1000										
LL | Param | AIC	-3726.6	36	7525.3								
LR Stat | DF | P-Val	140.1329	16	0.0000								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

	FAdvMath	FComScore	F$\ddot{m}_2$	F$\overline{m_{2A}}$	Cutoff1	L(Cutoff)$_1$	Cutoff3	P(Cutoff)$_3$	Success	Beta	Sigma
Intercept	-6.3855***	0.8895	5.4674***	-1.2247***	0.2728***	1.3137	-10.8886***	8.54%	-10.6373***		
	(0.8428)	(0.7565)	(0.3635)	(0.0038)	(0.0348)	(1.2270, 1.4065)	(3.3289)	(0.78%, 52.65%)	(1.6709)		
GREQuantPct	0.0729***	0.0575***	-0.0080*						0.1086***		
	(0.0100)	(0.0091)	(0.0045)						(0.0189)		
Gender	-0.0394	0.3093***	-0.0477						-0.3655		
	(0.1370)	(0.1191)	(0.0637)						(0.2349)		
Demographic1	0.5960***	-1.5884***	0.0172						-0.2786		
	(0.1823)	(0.1549)	(0.0902)						(0.3409)		
Demographic2	0.0490	-1.3595***	0.0894						0.3797		
	(0.2034)	(0.1787)	(0.0969)						(0.3233)		
AdvMath		0.9255***	0.1356**						-0.1450		
		(0.1216)	(0.0669)						(0.3226)		
ComScore			0.0115						0.1492*		
			(0.0166)						(0.0862)		
YearTransMat					-0.1716***	1.1066		95.42%			
					(0.0352)	(1.0468, 1.1699)		(87.56%, 98.40%)			
$\ddot{m}_2$				1.4277***							
				(0.0007)							
$\overline{m_{2A}}$							1.9267***				
							(0.4829)				
$\sqrt{\text{MeanSqError}}$		1.8550***	0.9791***	0.0057***							
		(0.0393)	(0.0022)	(0.0005)							
Beta										0.9538***	
										(0.0337)	
Sigma											0.2945***
											(0.0281)
Observations	1000										
LL | Param | AIC	-3818.2	36	7708.3								
LR Stat | DF | P-Val	64.0402	16	0.0000								

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

Table 29: Waitlist-Dynamic Unrestricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

	FStage1		FAdvMath		FComScore	F$\ddot{m}_2$	F$\overline{m_{2A}}$	Transition		Accept2		Accept3		Success	
	Coef	AME	Coef	AME	Coef	Coef	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-6.3804***	-2.0906***	-3.4981*	-1.2239*	2.8114	7.2395***	0.3539***	-10.7082***	-2.1055***	1.3743	0.1762	5.3172**	0.6942**	-7.8194***	-1.6624***
	(0.6106)	(0.1572)	(1.8860)	(0.6473)	(2.6732)	(0.6887)	(0.0038)	(1.0451)	(0.1549)	(2.2965)	(0.2928)	(2.3592)	(0.3037)	(1.7334)	(0.3276)
GREQuantPct	0.0661***	0.0217***	0.0396**	0.0139**	0.0394	-0.0260***		0.1110***	0.0218***	-0.0219	-0.0028	0.0295	0.0038	0.0799***	0.0170***
	(0.0070)	(0.0019)	(0.0184)	(0.0063)	(0.0258)	(0.0083)		(0.0119)	(0.0019)	(0.0253)	(0.0032)	(0.0266)	(0.0034)	(0.0206)	(0.0040)
Gender	0.2030**	0.0665**	0.1626	0.0569	0.6119***	-0.2034**		0.3464**	0.0681**	0.3963	0.0508	-0.7544**	-0.0985**	-0.0159	-0.0034
	(0.0880)	(0.0286)	(0.1475)	(0.0513)	(0.2188)	(0.1028)		(0.1473)	(0.0286)	(0.2997)	(0.0386)	(0.3455)	(0.0441)	(0.2373)	(0.0504)
Demographic1	-0.1332	-0.0437	0.6245***	0.2185***	-1.2588***	0.0513		-0.2141	-0.0421	-0.1504	-0.0193	-1.0906**	-0.1424**	-0.0979	-0.0208
	(0.1213)	(0.0396)	(0.2084)	(0.0703)	(0.2771)	(0.1586)		(0.2039)	(0.0400)	(0.4588)	(0.0589)	(0.5549)	(0.0711)	(0.3742)	(0.0795)
Demographic2	0.2572*	0.0843*	0.0475	0.0166	-1.2957***	-0.0137		0.4383**	0.0862**	0.2732	0.0350	-1.0075**	-0.1315**	-0.1905	-0.0405
	(0.1321)	(0.0430)	(0.2234)	(0.0781)	(0.3259)	(0.1620)		(0.2209)	(0.0431)	(0.4721)	(0.0604)	(0.5084)	(0.0663)	(0.3793)	(0.0805)
AdvMath					0.7506***	0.0656				0.4923	0.0631	0.0398	0.0052	0.1438	0.0306
					(0.2176)	(0.1073)				(0.3247)	(0.0415)	(0.3583)	(0.0467)	(0.2554)	(0.0543)
ComScore						0.0369				0.0048	0.0006	0.1290	0.0168*	0.0328	0.0070
						(0.0261)				(0.0678)	(0.0087)	(0.0788)	(0.0102)	(0.0577)	(0.0122)
YearTransMat	0.5997***	0.1965***						1.0061***	0.1978***						
	(0.0885)	(0.0266)						(0.1491)	(0.0264)						
$\ddot{m}_2$							1.0263***			-0.3288**	-0.0422**				
							(0.0007)			(0.1479)	(0.0187)				
$\overline{m_{2A}}$												-1.5304***	-0.1998***		
												(0.1678)	(0.0074)		
InvMills			-0.2289	-0.0801	-0.7304										
			(0.3545)	(0.1241)	(0.5361)										
$\sqrt{\text{MeanSqError}}$					1.8994***	0.9581***	0.0112***								
					(0.0682)	(0.0145)	(0.0005)								
Obs | Pseudo-R^2	1000	0.4934													
LL | Param | AIC	-2023.1	60	4166.3												

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

(b) Optimal Committee

	FStage1		FAdvMath		FComScore	F$\ddot{m}_2$	F$\overline{m_{2A}}$	Transition		Accept2		Accept3		Success	
	Coef	AME	Coef	AME	Coef	Coef	Coef	Coef	AME	Coef	AME	Coef	AME	Coef	AME
Intercept	-4.8625***	-1.6784***	-3.3658*	-1.1857*	1.8384	6.2905***	-0.3093***	-8.0678***	-1.6875***	-7.7385***	-0.7415***	-0.1808	-0.0240	-9.8861***	-1.9462***
	(0.5540)	(0.1650)	(1.8728)	(0.6491)	(2.4312)	(0.7189)	(0.0043)	(0.9324)	(0.1635)	(2.5841)	(0.2347)	(2.0591)	(0.2734)	(1.7554)	(0.2871)
GREQuantPct	0.0510***	0.0176***	0.0419**	0.0148**	0.0420*	-0.0174**		0.0848***	0.0177***	0.0682**	0.0065**	0.0620**	0.0082**	0.0976***	0.0192***
	(0.0065)	(0.0020)	(0.0177)	(0.0061)	(0.0231)	(0.0086)		(0.0109)	(0.0020)	(0.0312)	(0.0029)	(0.0266)	(0.0034)	(0.0203)	(0.0035)
Gender	0.0357	0.0123	0.0639	0.0225	0.2914	0.0890		0.0591	0.0124	-0.5395	-0.0517	-0.1356	-0.0180	-0.0596	-0.0117
	(0.0858)	(0.0296)	(0.1456)	(0.0513)	(0.1923)	(0.1128)		(0.1415)	(0.0296)	(0.3783)	(0.0364)	(0.3031)	(0.0403)	(0.2466)	(0.0486)
Demographic1	-0.1721	-0.0594	0.4192**	0.1477**	-1.1334***	0.1101		-0.2879	-0.0602	0.2533	0.0243	0.4261	0.0566	-0.4539	-0.0894
	(0.1174)	(0.0404)	(0.1970)	(0.0683)	(0.2557)	(0.1627)		(0.1955)	(0.0407)	(0.6433)	(0.0619)	(0.4952)	(0.0654)	(0.3662)	(0.0716)
Demographic2	-0.0380	-0.0131	-0.0492	-0.0173	-0.8501***	0.1106		-0.0799	-0.0167	0.5085	0.0487	0.1735	0.0230	0.5555	0.1093
	(0.1294)	(0.0446)	(0.2143)	(0.0755)	(0.2867)	(0.1773)		(0.2150)	(0.0449)	(0.6333)	(0.0608)	(0.5088)	(0.0675)	(0.3812)	(0.0742)
AdvMath					1.0421***	0.1118				0.2795	0.0268	-0.0581	-0.0077	-0.1045	-0.0206
					(0.1961)	(0.1240)				(0.3842)	(0.0368)	(0.3645)	(0.0484)	(0.2639)	(0.0520)
ComScore						0.0053				-0.0174	-0.0017	0.0645	0.0086	0.1478**	0.0291**
						(0.0312)				(0.0989)	(0.0095)	(0.0889)	(0.0118)	(0.0684)	(0.0133)
YearTransMat	0.4862***	0.1678***						0.8004***	0.1674***						
	(0.0847)	(0.0277)						(0.1410)	(0.0277)						
$\ddot{m}_2$							1.2910***			-0.1179	-0.0113				
							(0.0008)			(0.1517)	(0.0144)				
$\overline{m_{2A}}$												-1.0993***	-0.1460***		
												(0.1173)	(0.0034)		
InvMills			-0.3619	-0.1275	-0.0012										
			(0.4395)	(0.1545)	(0.5745)										
$\sqrt{\text{MeanSqError}}$					1.7743***	1.0532***	0.0145***								
					(0.0631)	(0.0129)	(0.0006)								
Obs | Pseudo-R^2	1000	0.4716													
LL | Param | AIC	-2143.9	60	4407.8												

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01.

Table 30: Waitlist-Dynamic Restricted Estimates, z Missing when $d_1 = R$

(a) Suboptimal Committee

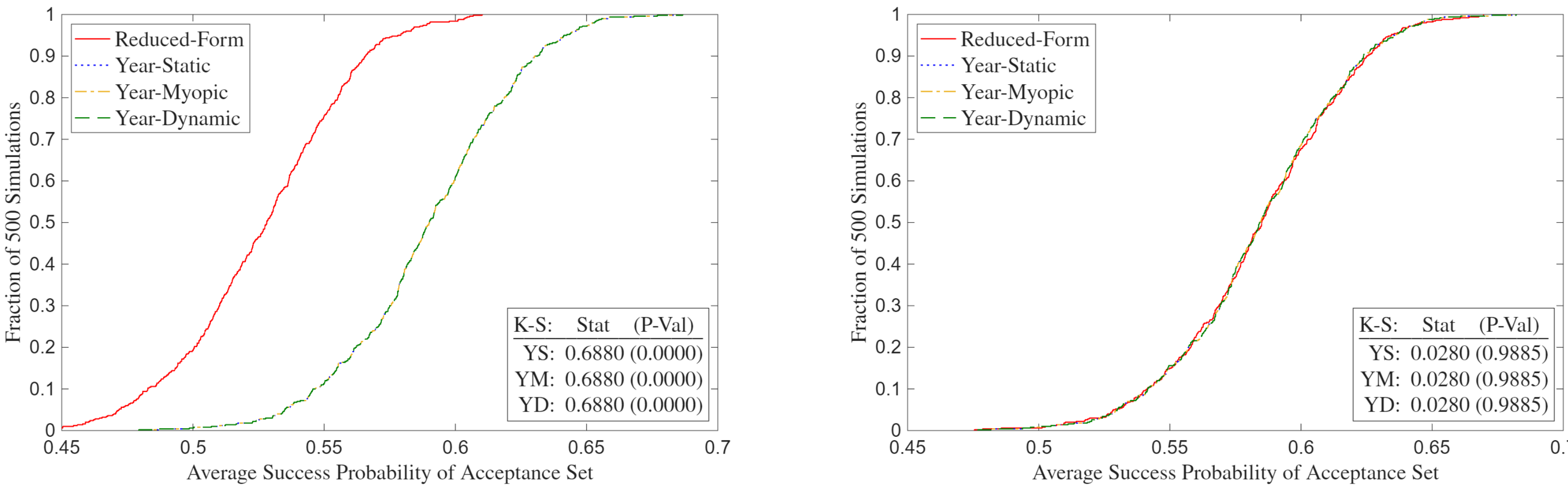
	FStage1	FAdvMath	FComScore	F$\hat{m}_2$	F$\overline{m_{2A}}$	Cutoff1	$L(\text{Cutoff})_1$	Cutoff3	$P(\text{Cutoff})_3$	Success	Beta	Sigma
Intercept	-6.3854***	-3.4955*	2.8010	7.1897***	0.3539***	0.3447	1.4115	-15.2825**	12.63%	-8.1530		
	(0.6222)	(1.8891)	(2.7427)	(0.6891)	(0.0038)	(0.2369)	(0.8873, 2.2456)	(6.4479)	(1.76%, 53.89%)	(7.7494)		
GREQuantPct	0.0662***	0.0397**	0.0402	-0.0254***						0.0841		
	(0.0071)	(0.0185)	(0.0265)	(0.0083)						(0.0876)		
Gender	0.2028**	0.1634	0.6180***	-0.2247**						0.1718		
	(0.0882)	(0.1475)	(0.2190)	(0.1027)						(0.3545)		
Demographic1	-0.1360	0.6339***	-1.2499***	0.0656						-0.4460		
	(0.1214)	(0.2084)	(0.2776)	(0.1587)						(0.2837)		
Demographic2	0.2547*	0.0562	-1.2917***	-0.0198						-0.2234		
	(0.1322)	(0.2229)	(0.3267)	(0.1621)						(0.3332)		
AdvMath			0.7526***	0.0283						0.2449		
			(0.2176)	(0.1074)						(0.2493)		
ComScore				0.0398						0.0517		
				(0.0261)						(0.0637)		
YearTransMat	0.6046***					-0.2388**	1.1117		98.87%			
	(0.0900)					(0.1113)	(0.8544, 1.4464)		(70.77%, 99.97%)			
$\hat{m}_2$					1.0263***							
					(0.0007)							
$\overline{m_{2A}}$								3.0968**				
								(1.2801)				
InvMills		-0.2419	-0.7991									
		(0.3534)	(0.5481)									
$\sqrt{\text{MeanSqError}}$			1.9101***	0.9584***	0.0112***							
			(0.0695)	(0.0145)	(0.0005)							
Beta											0.9496***	
											(0.0051)	
Sigma												0.3621*
												(0.2142)
Observations	1000											
LL | Param | AIC	-2093.9	44	4275.8									
LR Stat | DF | P-Val	141.5488	16	0.0000									

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.

(b) Optimal Committee

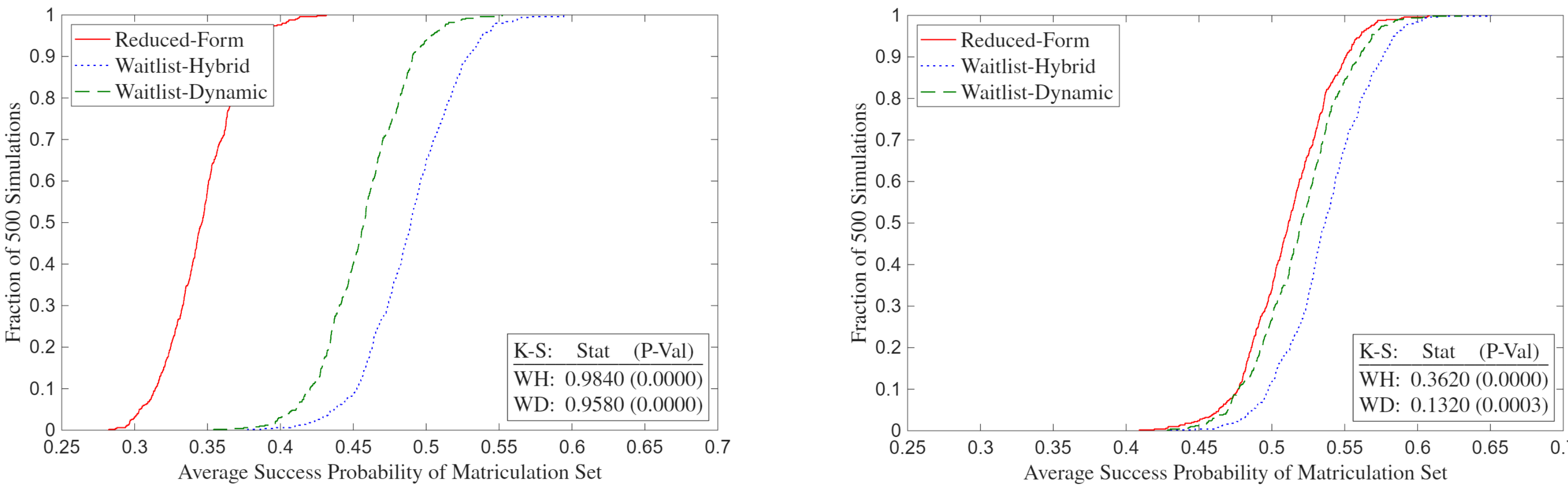
	FStage1	FAdvMath	FComScore	F$\hat{m}_2$	F$\overline{m_{2A}}$	Cutoff1	$L(\text{Cutoff})_1$	Cutoff3	$P(\text{Cutoff})_3$	Success	Beta	Sigma
Intercept	-4.8616***	-3.3608*	1.8102	6.2650***	-0.3093***	0.2697***	1.3096	-10.5162***	9.90%	-10.1044***		
	(0.5602)	(1.8772)	(2.4663)	(0.7187)	(0.0043)	(0.0343)	(1.2245, 1.4005)	(2.4227)	(1.81%, 39.53%)	(1.7608)		
GREQuantPct	0.0510***	0.0420**	0.0439*	-0.0171**						0.1033***		
	(0.0065)	(0.0177)	(0.0235)	(0.0086)						(0.0203)		
Gender	0.0360	0.0631	0.2795	0.0906						-0.3537		
	(0.0858)	(0.1459)	(0.1928)	(0.1128)						(0.2316)		
Demographic1	-0.1719	0.4208**	-1.1529***	0.1064						-0.2621		
	(0.1174)	(0.1973)	(0.2563)	(0.1627)						(0.3374)		
Demographic2	-0.0379	-0.0463	-0.8508***	0.1103						0.3731		
	(0.1294)	(0.2148)	(0.2883)	(0.1773)						(0.3240)		
AdvMath			1.0446***	0.1079						-0.1879		
			(0.1986)	(0.1240)						(0.3391)		
ComScore				0.0055						0.1416*		
				(0.0312)						(0.0826)		
YearTransMat	0.4856***					-0.1701***	1.1047		95.54%			
	(0.0851)					(0.0347)	(1.0451, 1.1677)		(86.61%, 98.61%)			
$\hat{m}_2$					1.2910***							
					(0.0008)							
$\overline{m_{2A}}$								1.8792***				
								(0.3680)				
InvMills		-0.3751	-0.1080									
		(0.4435)	(0.5806)									
$\sqrt{\text{MeanSqError}}$			1.7658***	1.0534***	0.0145***							
			(0.0622)	(0.0129)	(0.0006)							
Beta											0.9513***	
											(0.0058)	
Sigma												0.2923***
												(0.0262)
Observations	1000											
LL | Param | AIC	-2177.3	44	4442.5									
LR Stat | DF | P-Val	66.7052	16	0.0000									

Robust standard errors in parentheses: *p<0.10, **p<0.05, ***p<0.01. Confidence intervals are 95%.



(a) Suboptimal Committee (b) Optimal Committee

Figure 7: Deviation Gains, z Missing when $d_1 = R$



(a) Suboptimal Committee (b) Optimal Committee

Figure 8: Waitlist Deviation Gains, z Missing when $d_1 = R$

D Fully-Specified Waitlist Models

In Section 5, I present simplified versions of equations in the waitlist models. However, the full equations include additional state variables: n_3 is a given year's number of waitlisted applicants, $\overline{q_3}$ is that year's average waitlisted applicant quality, and $\overline{m_3}$ is that year's average waitlisted applicant matriculation score. The full equations are as follows, starting with Round 3:

$$v_3(x,z,\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3},\varepsilon_3) = \max_{d_3\in\{A,R\}} u_3(x,z,\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3},d_3) + \varepsilon_3(d_3), \text{ such that:} \tag{27}$$

$$u_3(x,z,\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3},d_3) = P(y=S|x,z)\mathbb{1}(d_3=A) + P(y_3^*=S|\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3})\mathbb{1}(d_3=R).$$

The Round 2 value function is:

$$v_2(x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2},\varepsilon_2) = \max_{d_2\in\{A,W\}} u_2(x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2},d_2) + \varepsilon_2(d_2), \text{ such that:} \tag{28}$$

$$\begin{aligned} u_2(x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2},d_2) = {} & P(y=S|x,z)\mathbb{1}(d_2=A) \\ & + \beta\mathbb{E}[v_3(x,z,\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3},\varepsilon_3)|x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2}]\mathbb{1}(d_2=W), \text{ such that:} \end{aligned}$$

$$\begin{aligned} & \mathbb{E}[v_3(x,z,\overline{m_{2A}},n_3,\overline{q_3},\overline{m_3},\varepsilon_3)|x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2}] = \\ & \quad \int_{\widetilde{\overline{m_3}}}\int_{\widetilde{\overline{q_3}}}\int_{\widetilde{n_3}}\int_{\widetilde{\overline{m_{2A}}}} \sigma\log\left(\sum_{\widetilde{d_3}\in\{A,R\}} e^{u_3(x,z,\widetilde{\overline{m_{2A}}},\widetilde{n_3},\widetilde{\overline{q_3}},\widetilde{\overline{m_3}},\widetilde{d_3})/\sigma}\right) \\ & \quad dF_{\overline{m_{2A}}|\ddot{m}_2}(\widetilde{\overline{m_{2A}}}|\ddot{m}_2)dF_{n_3|n_2}(\widetilde{n_3}|n_2)dF_{\overline{q_3}|\overline{q_2}}(\widetilde{\overline{q_3}}|\overline{q_2})dF_{\overline{m_3}|\overline{m_2}}(\widetilde{\overline{m_3}}|\overline{m_2}). \end{aligned}$$

Finally, the Round 1 value function is:

$$v_1(x,t,n_1,\overline{q_1},\overline{m_1},\varepsilon_1) = \max_{d_1\in\{T,R\}} u_1(x,t,n_1,\overline{q_1},\overline{m_1},d_1) + \varepsilon_1(d_1), \text{ such that:} \tag{29}$$

$$\begin{aligned} u_1(x,t,n_1,\overline{q_1},\overline{m_1},d_1) = {} & \mathbb{E}[v_2(x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2},\varepsilon_2)|x,n_1,\overline{q_1},\overline{m_1}]\mathbb{1}(d_1=T) \\ & + L(y_1^*=S|t)\mathbb{1}(d_1=R), \text{ such that:} \end{aligned}$$

$$\begin{aligned} & \mathbb{E}[v_2(x,z,\ddot{m}_2,n_2,\overline{q_2},\overline{m_2},\varepsilon_2)|x,n_1,\overline{q_1},\overline{m_1}] = \int_{\widetilde{\overline{m_2}}}\int_{\widetilde{\overline{q_2}}}\int_{\widetilde{n_2}}\int_{\widetilde{z}}\int_{\widetilde{\ddot{m}_2}} \sigma\log\left(\sum_{\widetilde{d_2}\in\{A,R\}} e^{u_2(x,\widetilde{z},\widetilde{\ddot{m}_2},\widetilde{n_2},\widetilde{\overline{q_2}},\widetilde{\overline{m_2}},\widetilde{d_2})/\sigma}\right) \\ & \quad dF_{\ddot{m}_2|x,z}(\widetilde{\ddot{m}_2}|x,\widetilde{z})dF_{z|x}(\widetilde{z}|x)dF_{n_2|n_1}(\widetilde{n_2}|n_1)dF_{\overline{q_2}|\overline{q_1}}(\widetilde{\overline{q_2}}|\overline{q_1})dF_{\overline{m_2}|\overline{m_1}}(\widetilde{\overline{m_2}}|\overline{m_1}). \end{aligned}$$

In the full waitlist-hybrid model, the Round 1 utility function is instead:

$$u_1(x,t,d_1) = P(y=S|x)\mathbb{1}(d_1=T) + P(y_1^*=S|t)\mathbb{1}(d_1=R). \tag{30}$$

Expressed as simply as possible, the restricted CCPs are:

$$P(d_r|\text{states}) = \frac{\exp(u_r(\text{states}, d_r)/\sigma)}{\sum_{\widetilde{d}_r \in D_r} \exp(u_r(\text{states}, \widetilde{d}_r)/\sigma)}. \tag{31}$$

The unrestricted CCPs, success probabilities, and cutoff probabilities are $P_{d_1} = P\big(d_1 = T\,\big|x, t, n_1, \overline{q_1}, \overline{m_1}; \theta_T\big)$, $P_{d_2} = P\big(d_2 = A\big|x, z, \ddot{m}_2, n_2, \overline{q_2}, \overline{m_2}; \theta_{A_2}\big)$, $P_{d_3} = P\big(d_3 = A\big|x, z, \overline{m_{2A}}, n_3, \overline{q_3}, \overline{m_3}; \theta_{A_3}\big)$, $P_y = P\big(y = S\big|x, z; \theta_S\big)$, $L_{y_1^*} = L\big(y_1^* = S\big|t; \theta_{S_1^*}\big)$, and $P_{y_3^*} = P\big(y_3^* = S\big|\overline{m_{2A}}, n_3, \overline{q_3}, \overline{m_3}; \theta_{S_3^*}\big)$. They have the same logit specifications as in the year-static and dynamic models; see Equations (9)–(11).

Let applicants 1 to N_3 be waitlisted, N_3+1 to N_2 be accepted in Round 2, and N_2+1 to N_1 be rejected in Round 1. Then the unrestricted full waitlist-dynamic likelihood is:

$$\begin{aligned}
\mathcal{L}^U_{N_1}(\theta) = \prod_{n=1}^{N_3} & f_{z|x}(z_n|x_n; \theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n, z_n; \theta_{F_{\ddot{m}_2|x,z}}) f_{n_2|n_1}(n_{2,n}|n_{1,n}; \theta_{F_{n_2|n_1}}) f_{\overline{q_2}|\overline{q_1}}(\overline{q_{2,n}}|\overline{q_{1,n}}; \theta_{F_{\overline{q_2}|\overline{q_1}}}) \\
& f_{\overline{m_2}|\overline{m_1}}(\overline{m_{2,n}}|\overline{m_{1,n}}; \theta_{F_{\overline{m_2}|\overline{m_1}}}) f_{\overline{m_{2A}}|\ddot{m}_2}(\overline{m_{2A,n}}|\ddot{m}_{2,n}; \theta_{F_{\overline{m_{2A}}|\ddot{m}_2}}) f_{n_3|n_2}(n_{3,n}|n_{2,n}; \theta_{F_{n_3|n_2}}) f_{\overline{q_3}|\overline{q_2}}(\overline{q_{3,n}}|\overline{q_{2,n}}; \theta_{F_{\overline{q_3}|\overline{q_2}}}) \\
& f_{\overline{m_3}|\overline{m_2}}(\overline{m_{3,n}}|\overline{m_{2,n}}; \theta_{F_{\overline{m_3}|\overline{m_2}}}) P_{d_{1,n}}(\theta_T)(1 - P_{d_{2,n}}(\theta_{A_2})) \\
& P_{d_{3,n}}(\theta_{A_3})^{\mathbb{1}(d_{3,n}=A)}(1 - P_{d_{3,n}}(\theta_{A_3}))^{\mathbb{1}(d_{3,n}=R)} P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1 - P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)} \\
\prod_{n=N_3+1}^{N_2} & f_{z|x}(z_n|x_n; \theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n, z_n; \theta_{F_{\ddot{m}_2|x,z}}) f_{n_2|n_1}(n_{2,n}|n_{1,n}; \theta_{F_{n_2|n_1}}) f_{\overline{q_2}|\overline{q_1}}(\overline{q_{2,n}}|\overline{q_{1,n}}; \theta_{F_{\overline{q_2}|\overline{q_1}}}) \\
& f_{\overline{m_2}|\overline{m_1}}(\overline{m_{2,n}}|\overline{m_{1,n}}; \theta_{F_{\overline{m_2}|\overline{m_1}}}) f_{\overline{m_{2A}}|\ddot{m}_2}(\overline{m_{2A,n}}|\ddot{m}_{2,n}; \theta_{F_{\overline{m_{2A}}|\ddot{m}_2}}) f_{n_3|n_2}(n_{3,n}|n_{2,n}; \theta_{F_{n_3|n_2}}) f_{\overline{q_3}|\overline{q_2}}(\overline{q_{3,n}}|\overline{q_{2,n}}; \theta_{F_{\overline{q_3}|\overline{q_2}}}) \\
& f_{\overline{m_3}|\overline{m_2}}(\overline{m_{3,n}}|\overline{m_{2,n}}; \theta_{F_{\overline{m_3}|\overline{m_2}}}) P_{d_{1,n}}(\theta_T) P_{d_{2,n}}(\theta_{A_2}) P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1 - P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)} \\
\prod_{n=N_2+1}^{N_1} & f_{z|x}(z_n|x_n; \theta_{F_{z|x}}) f_{\ddot{m}_2|x,z}(\ddot{m}_{2,n}|x_n, z_n; \theta_{F_{\ddot{m}_2|x,z}}) f_{n_2|n_1}(n_{2,n}|n_{1,n}; \theta_{F_{n_2|n_1}}) f_{\overline{q_2}|\overline{q_1}}(\overline{q_{2,n}}|\overline{q_{1,n}}; \theta_{F_{\overline{q_2}|\overline{q_1}}}) \\
& f_{\overline{m_2}|\overline{m_1}}(\overline{m_{2,n}}|\overline{m_{1,n}}; \theta_{F_{\overline{m_2}|\overline{m_1}}})(1 - P_{d_{1,n}}(\theta_T)).
\end{aligned} \tag{32}$$

For the unrestricted and restricted likelihoods, θ^U_S is estimated with the first-stage likelihood $\mathcal{L}^P_{N_2}(\theta_S) = \prod_{n=1}^{N_2} P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1 - P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)}$ to get $\overline{q}$ = a given year's $\overline{P_{y_n}}$.

E Multiyear Models

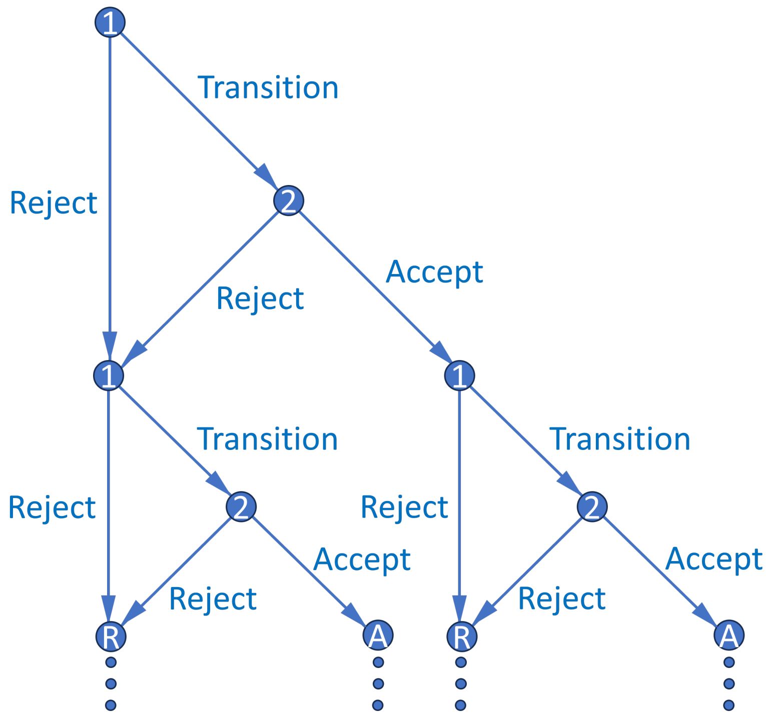


Figure 9: Multiyear Decision Graph

There are two different multiyear models. In the first model, the memory variable is the size of the previous year's cohort. The committee accepts applicants more likely to matriculate because it wants to ensure that next year's memory variable, which is this year's cohort size, is sufficiently large that it has the option to avoid accepting too many relatively weak applicants next year. In the second model, the memory is a diversity variable that increases an applicant's success probability; for example, the previous year's cohort's main area of interest. Applicants with a diverse area of interest may be more likely to succeed because their advisor, with fewer other students to advise, can give them more individualized attention. The committee will therefore accept applicants with a less common area of interest to help balance out the program for next year's class.

E.1 Matriculation Model

In the waitlist models in Section 5.4, the committee benefits from giving early acceptances to applicants who are more likely to matriculate. Doing so helps prevent the program from losing out on good applicants who would have matriculated if given an early offer but end up matriculating elsewhere while stuck on the waitlist. However, the waitlist models do not imply that the committee should also give more late acceptances to likely matriculants. To model that phenomenon, a multiyear model is necessary, as there must be a period after the final round in any given year for dynamics to matter in that round. Figure 9 is a decision graph of the simplest possible multiyear model, in which each year has only two rounds, as in the year-dynamic model (see Section 5.3). This model is infinite horizon, and future years are discounted with factor $\beta \in (0,1)$.

In each year's Round 2 value function, I specify the cutoff probability as $P(y_2^* = S|\overline{q}, M, c_{-1})$ instead of $P(y_2^* = S|t)$. $\overline{q} =$ is the average quality (i.e. $\overline{P(\text{Success})}$) of the current year's applicants who survive Round 1, M is the current year's number of surviving applicants, and c_{-1} is the previous year's cohort size. The cutoff probability increases in all three variables; in particular, it increases in c_{-1} because the committee must keep the total number of students in the program across years relatively constant. So if c_{-1} is large, the committee must now accept fewer students.

Given that, what information available at the start of Round 2 determines c? The current year's average surviving applicant's matriculation score $\overline{m}$ always does. Furthermore, if and only if the applicant is accepted, their own matriculation score m does as well. Now suppose the committee maximizes discounted "lifetime" cohort success by solving an infinite-horizon dynamic programming problem across years. As in the waitlist models, the optimal dynamic decision may not match the optimal myopic one. Instead, the committee may accept an applicant with a high m even if their expected success probability is less than what the Round 2 cutoff level (which need not be between 0 and 1) would otherwise be. The reason is that accepting a high m applicant increases c's distribution, which increases next year's Round 2 cutoff and value.

Intuitively, even if the committee can simply accept more applicants next year to make up for a smaller cohort this year, they will do so out of necessity, for example, to avoid future TA shortages. This situation will be undesirable if there is any degree of risk aversion with respect to the quality of next year's applicants. Also, as in the waitlist models, even if the data show that the committee does not think dynamically, if c_{-1} affects the Round 2 cutoff, perhaps it should. Data on success measures for recent years is not necessary to identify this effect; admission data will suffice. That said, estimating this model is more computationally challenging since it does not have a closed-form solution. Nested fixed point algorithms suffer from the curse of dimensionality, and the number of (x,z) state variables in application files is large. But lasso and factor analysis (see Grove et al. 2007) can help reduce the size of the state space.

I now present the value functions, starting with each year's Round 2 value:

$$v_2(x,z,m,\overline{m},\overline{q},M,c_{-1},\varepsilon_2) = \max_{d_2 \in \{A,R\}} u_2(x,z,m,\overline{m},\overline{q},M,c_{-1},d_2) + \varepsilon_2(d_2), \text{ such that:} \tag{33}$$

$$\begin{aligned} u_2(x,z,m,\overline{m},\overline{q},M,c_{-1},d_2) = &\{P(y=S|x,z) + \beta\mathbb{E}[v_1(x',t',c,\varepsilon_1{}')|m,\overline{m}]\}\mathbb{1}(d_2=A) \\ &+ \{L(y_2^*=S|\overline{q},M,c_{-1}) + \beta\mathbb{E}[v_1(x',t',c,\varepsilon_1{}')|\overline{m}]\}\mathbb{1}(d_2=R), \text{ such that:} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[v_1(x',t',c,\varepsilon_1{}')|(m),\overline{m}] &= \int_c \int_{t'} \int_{x'} \sigma \log\left(\sum_{d_1{}' \in \{T,R\}} e^{u_1(x',t',c,d_1{}')/\sigma}\right) dF_x(x')dF_t(t')dF_{c|(m),\overline{m}}(c|(m),\overline{m}) \\ &= \int_c \int_{t_1{}'} \cdots \int_{t_{J_t}{}'} \int_{x_1{}'} \cdots \int_{x_{J_x}{}'} \left[\,\cdot\,\right] dF_{x,1}(x_1{}'|x_2{}',\ldots,x_{J_x}{}') \cdots dF_{x,J_x}(x_{J_x}{}') \\ &\qquad dF_{t,1}(t_1{}'|t_2{}',\ldots,t_{J_t}{}') \cdots dF_{t,J_t}(t_{J_t}{}')dF_{c|(m),\overline{m}}(c|(m),\overline{m}). \end{aligned}$$

As such, Round 2's conditional choice probabilities are:

$$P(d_2|x,z,m,\overline{m},\overline{q},M,c_{-1}) = \frac{\exp(u_2(x,z,m,\overline{m},\overline{q},M,c_{-1},d_2)/\sigma)}{\sum_{\widetilde{d_2}\in\{A,R\}}\exp(u_2(x,z,m,\overline{m},\overline{q},M,c_{-1},\widetilde{d_2})/\sigma)}. \tag{34}$$

Each year's Round 1 value function is:

$$v_1(x,t,c_{-1},\varepsilon_1) = \max_{d_1\in\{T,R\}} u_1(x,t,c_{-1},d_1)+\varepsilon_1(d_1), \text{ such that:} \tag{35}$$

$$u_1(x,t,c_{-1},d_1) = \mathbb{E}[v_2(x,z,m,\overline{m},\overline{q},M,c_{-1},\varepsilon_2)|x,c_{-1}]\mathbb{1}(d_1=T) \\ +\{L(y_1^*=S|t)+\beta\mathbb{E}[v_1(x',t',c,\varepsilon_1')|\overline{m}]\}\mathbb{1}(d_1=R), \text{ such that:}$$

$$\mathbb{E}[v_2(x,z,m,\overline{m},\overline{q},M,c_{-1},\varepsilon_2)|x,c_{-1}] = \int_{\widetilde{M}}\int_{\widetilde{\overline{q}}}\int_{\widetilde{\overline{m}}}\int_{\widetilde{m}}\int_{\widetilde{z}}\sigma\log\left(\sum_{\widetilde{d_2}\in\{A,R\}} e^{u_2(x,\widetilde{z},\widetilde{m},\widetilde{\overline{m}},\widetilde{\overline{q}},\widetilde{M},c_{-1},\widetilde{d_2})/\sigma}\right) \\ dF_{z|x}(\widetilde{z}|x)dF_{m|x}(\widetilde{m}|x)dF_{\overline{m}}(\widetilde{\overline{m}})dF_{\overline{q}}(\widetilde{\overline{q}})dF_M(\widetilde{M}) \\ = \int_{\widetilde{M}}\int_{\widetilde{\overline{q}}}\int_{\widetilde{\overline{m}}}\int_{\widetilde{m}}\int_{\widetilde{z_1}}\cdots\int_{\widetilde{z_{J_z}}}\Big[\cdot\Big]dF_{z|x,1}(\widetilde{z_1}|\widetilde{z_2},\ldots,\widetilde{z_{J_z}},x)\cdots dF_{z|x,J_z}(\widetilde{z_{J_z}}|x)dF_{m|x}(\widetilde{m}|x)dF_{\overline{m}}(\widetilde{\overline{m}})dF_{\overline{q}}(\widetilde{\overline{q}})dF_M(\widetilde{M}).$$

And as such, Round 1's conditional choice probabilities (CCPs) are:

$$P(d_1|x,t,c_{-1}) = \frac{\exp(u_1(x,t,c_{-1},d_1)/\sigma)}{\sum_{\widetilde{d_1}\in\{T,R\}}\exp(u_1(x,t,c_{-1},\widetilde{d_1})/\sigma)}. \tag{36}$$

The unrestricted CCPs, success probabilities, and cutoff levels are $P_{d_1} = P\big(d_1=T\,\big|x,t,c_{-1};\theta_T\big)$, $P_{d_2} = P\big(d_2=A\big|x,z,m,\overline{m},\overline{q},M,c_{-1};\theta_A\big)$, $P_y = P\big(y=S\big|x,z;\theta_S\big)$, $L_{y_1^*} = L\big(y_1^*=S\big|t;\theta_{S_1^*}\big)$, and $L_{y_2^*} = L\big(y_2^*=S\big|\overline{q},M,c_{-1};\theta_{S_2^*}\big)$, which aside from the cutoffs have the usual logit specification. In this model, the cutoffs are not encased in a function that restricts their range.

Given all that, the unrestricted likelihood is:

$$\mathcal{L}_N^U(\theta) = \prod_{n=1}^{N_T} f_x(x_n;\theta_{F_x})f_t(t_n;\theta_{F_t})f_{z|x}(z_n|x_n;\theta_{F_{z|x}})f_{m|x}(m_n|x_n;\theta_{F_{m|x}})f_{\overline{m}}(\overline{m}_n;\theta_{F_{\overline{m}}})f_{\overline{q}}(\overline{q}_n;\theta_{F_{\overline{q}}})f_M(M_n;\theta_{F_M})f_{c|\overline{m}}(c_n|\overline{m}_n;\theta_{F_{c|\overline{m}}}) \\ f_{c|m,\overline{m}}(c_n|m_n,\overline{m}_n;\theta_{F_{c|m,\overline{m}}})P_{d_{1,n}}(\theta_T)P_{d_{2,n}}(\theta_A)^{\mathbb{1}(d_{2,n}=A)}(1-P_{d_{2,n}}(\theta_A))^{\mathbb{1}(d_{2,n}=R)}P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)} \\ \prod_{n=N_T+1}^{N} f_x(x_n;\theta_{F_x})f_t(t_n;\theta_{F_t})f_{z|x}(z_n|x_n;\theta_{F_{z|x}})f_{m|x}(m_n|x_n;\theta_{F_{m|x}})(1-P_{d_{1,n}}(\theta_T)). \tag{37}$$

As with the year and waitlist-dynamic models, I use (θ_F^U,θ_S^U) as an initial guess of (θ_F^R,θ_S^R). Also, for the unrestricted and restricted likelihoods, I estimate θ_S^U with the first-stage likelihood $\mathcal{L}_{N_T}^P(\theta_S) = \prod_{n=1}^{N_T} P_{y_n}(\theta_S)^{\mathbb{1}(y_n=S)}(1-P_{y_n}(\theta_S))^{\mathbb{1}(y_n=F)}$ to get $\overline{q}$ = a given year's $\overline{P_{y_n}}$.

To show that the dynamics of this model work as intended, I prove the following theorem:

Theorem 4. *Success probability equal, an applicant's Round 2 acceptance probability is strictly increasing in their matriculation score.*

Proof. See Appendix F.

E.2 Diversity Model

I remove the state variables $\overline{q}$, M, and c_{-1} and add in their place two new states a and $\overline{a}_{-1}$. Let $a = \mathbb{1}(\text{Diverse})$, where a diverse characteristic is one whose probability of appearing among next year's applicants is bounded above by some $p \in (0,1)$. For example, I might have $a = \mathbb{1}(\text{TheoryField})$, where in most years, theory is less common than applied. Also, let $\overline{a}_{-1}$ equal the percentage of last year's acceptance set that is theory. If $|a - \overline{a}_{-1}|$ shrinks, then the applicant's expected success probability decreases, as their future advisor may have less time for them.[14]

As such, the optimal dynamic decision may not match the optimal myopic one. In particular, the committee may accept a diverse applicant it would not have otherwise because doing so increases the success probability of next year's applicant pool. Accepting a diverse applicant increases the distribution of $\overline{a}$ conditional on a, which because diverse applicants are less common increases the distribution of $|a' - \overline{a}|$, which increases next year's expected success probability, which increases next year's expected value. That said, I can also estimate a myopic model with memory variable $\overline{a}_{-1}$. The committee will still prioritize diverse applicants because diversity increases applicants' success probabilities, but there will be no forward-looking dynamic effect.

I now present the value functions, starting with each year's Round 2 value:

$$v_2(x,z,a,\overline{a}_{-1},t,\varepsilon_2) = \max_{d_2 \in \{A,R\}} u_2(x,z,a,\overline{a}_{-1},t,d_2) + \varepsilon_2(d_2), \text{ such that:} \tag{38}$$

$$\begin{aligned} u_2(x,z,a,\overline{a}_{-1},t,d_2) = &\{P(y=S|x,z,|a-\overline{a}_{-1}|) + \beta\mathbb{E}[v_1(x',\overline{a},t',\varepsilon_1')|a]\}\mathbb{1}(d_2 = A) \\ &+ \{L(y_2^* = S|t) + \beta\mathbb{E}[v_1(x',\overline{a},t',\varepsilon_1')]\}\mathbb{1}(d_2 = R), \text{ such that:} \end{aligned}$$

$$\mathbb{E}[v_1(x',\overline{a},t',\varepsilon_1')|a] = \int_{\widetilde{t'}}\int_{\widetilde{\overline{a}}}\int_{\widetilde{x'}} \sigma \log\left(\sum_{d_1' \in \{T,R\}} e^{u_1(\widetilde{x'},\widetilde{\overline{a}},\widetilde{t'},d_1')/\sigma}\right) dF_{x'}(\widetilde{x'}) dF_{\overline{a}|a}(\widetilde{\overline{a}}|a) dF_{t'}(\widetilde{t'})$$

$$= \int_{\widetilde{t_1'}} \cdots \int_{\widetilde{t_{J_t}'}} \int_{\widetilde{\overline{a}}} \int_{\widetilde{x_1'}} \cdots \int_{\widetilde{x_{J_x}'}} \left[\cdot\right] F_{x,1}(x_1'|x_2',\ldots,x_{J_x}') \cdots dF_{x,J_x}(x_{J_x}') dF_{\overline{a}|a}(\widetilde{\overline{a}}|a) dF_{t,1}(t_1'|t_2',\ldots,t_{J_t}') \cdots dF_{t,J_t}(t_{J_t}').$$

As such, Round 2's conditional choice probabilities are:

$$P(d_2|x,z,a,\overline{a}_{-1},t) = \frac{\exp(u_2(x,z,a,\overline{a}_{-1},t,d_2)/\sigma)}{\sum_{\widetilde{d_2} \in \{A,R\}} \exp(u_2(x,z,a,\overline{a}_{-1},t,\widetilde{d_2})/\sigma)}. \tag{39}$$

Each year's Round 1 value function is:

$$v_1(x,\overline{a}_{-1},t,\varepsilon_1) = \max_{d_1 \in \{T,R\}} u_1(x,\overline{a}_{-1},t,d_1) + \varepsilon_1(d_1), \text{ such that:} \tag{40}$$

[14] If the program never gets too imbalanced, this relationship will not be identifiable, in which case estimation will not work for this choice of a. However, there might be a different suitable diversity variable.

$$u_1(x,\overline{a}_{-1},t,d_1) = \mathbb{E}[v_2(x,z,a,\overline{a}_{-1},t,\varepsilon_2)|x,\overline{a}_{-1},t]\mathbb{1}(d_1=T)$$
$$+\{L(y_1^*=S|t)+\beta\mathbb{E}[v_1(x',\overline{a},t',\varepsilon_1')]\}\mathbb{1}(d_1=R), \text{ such that:}$$

$$\mathbb{E}[v_2(x,z,a,\overline{a}_{-1},t,\varepsilon_2)|x,\overline{a}_{-1},t] = \int_{\widetilde{a}}\int_{\widetilde{z}}\sigma\log\left(\sum_{\widetilde{d}_2\in A,R} e^{u_2(x,\widetilde{z},\widetilde{a},\overline{a}_{-1},t,\widetilde{d}_2)/\sigma}\right)dF_{z|x}(\widetilde{z}|x)dF_a(\widetilde{a})$$
$$= \int_{\widetilde{a}}\int_{\widetilde{z_1}}\cdots\int_{\widetilde{z_J}}\left[\cdot\right]dF_1(\widetilde{z_1}|\widetilde{z_2},\ldots\widetilde{z_J},x)\cdots dF_J(\widetilde{z_J}|x)dF_a(\widetilde{a})$$

$$\mathbb{E}[v_1(x',\overline{a},t',\varepsilon_1')] = \int_{\widetilde{t'}}\int_{\widetilde{\overline{a}}}\int_{\widetilde{x'}}\sigma\log\left(\sum_{d_1'\in\{T,R\}} e^{u_1(\widetilde{x'},\widetilde{\overline{a}},\widetilde{t'},d_1')/\sigma}\right)dF_{x'}(\widetilde{x'})dF_{\overline{a}}(\widetilde{\overline{a}})dF_{t'}(\widetilde{t'})$$
$$= \int_{\widetilde{t_1'}}\cdots\int_{\widetilde{t_{J_t}'}}\int_{\widetilde{\overline{a}}}\int_{\widetilde{x_1'}}\cdots\int_{\widetilde{x_{J_x}'}}\left[\cdot\right]F_{x,1}({x_1}'|{x_2}',\ldots,{x_{J_x}}')\cdots dF_{x,J_x}({x_{J_x}}')dF_{\overline{a}}(\widetilde{\overline{a}})dF_{t,1}({t_1}'|{t_2}',\ldots,{t_{J_t}}')\cdots dF_{t,J_t}({t_{J_t}}').$$

And as such, Round 1's conditional choice probabilities are:

$$P(d_1|x,\overline{a}_{-1},t) = \frac{\exp(u_1(x,\overline{a}_{-1},t,d_1)/\sigma)}{\sum_{\widetilde{d}_1\in\{T,R\}}\exp(u_1(x,\overline{a}_{-1},t,\widetilde{d}_1)/\sigma)}. \tag{41}$$

The unrestricted CCPs, success probabilities, and cutoff levels are $P_{d_1} = P\big(d_1=T\big|x,\overline{a}_{-1},t;\theta_T\big)$, $P_{d_2} = P\big(d_2=A\big|x,z,a,\overline{a}_{-1},t;\theta_A\big)$, $P_y = P\big(y=S\big|x,z,|a-\overline{a}_{-1}|;\theta_S\big)$, $L_{y_1^*} = L\big(y_1^*=S\big|t;\theta_{S_1^*}\big)$, and $L_{y_2^*} = L\big(y_2^*=S\big|t;\theta_{S_2^*}\big)$; all but the cutoffs are logit. Given all that, the unrestricted likelihood is:

$$\mathcal{L}_N^U(\theta) = \prod_{n=1}^{N_A} f_x(x_n;\theta_{F_x})f_t(t_n;\theta_{F_t})f_{z|x}(z_n|x_n;\theta_{F_{z|x}})f_a(a)f_{\overline{a}}(\overline{a})f_{\overline{a}|a}(\overline{a}|a)P_{d_{2,n}}P_{y_n}^{\mathbb{1}(y_n=S)}(1-P_{y_n})^{\mathbb{1}(y_n=F)}$$
$$\prod_{n=N_A+1}^{N_{T,R}} f_x(x_n;\theta_{F_x})f_t(t_n;\theta_{F_t})f_{z|x}(z_n|x_n;\theta_{F_{z|x}})f_a(a)f_{\overline{a}}(\overline{a})(1-P_{d_{2,n}})P_{y_n}^{\mathbb{1}(y_n=S)}(1-P_{y_n})^{\mathbb{1}(y_n=F)}$$
$$\prod_{n=N_{T,R}+1}^{N} f_x(x_n;\theta_{F_x})f_t(t_n;\theta_{F_t})f_{z|x}(z_n|x_n;\theta_{F_{z|x}})f_a(a)f_{\overline{a}}(\overline{a})(1-P_{d_{1,n}}). \tag{42}$$

I estimate $f_{\overline{a}|a}(\overline{a}|a)$ on the N_A accepted applicants only (those from N_A+1 to $N_{T,R}$ are rejected in Round 2) because only accepted applicants can influence the acceptance set's area of interest. Also, as with the year and waitlist-dynamic models, I use (θ_F^U,θ_S^U) as an initial guess of (θ_F^R,θ_S^R).

To show that the dynamics of this model work as intended, I prove the following theorem:

Theorem 5. *Consider any two applicants, one with* $a=1$ *and one with* $a=0$*, but with equal success probabilities and therefore different values of x or z. Suppose* $P(a'=1)=p>0$*. Then for all sufficiently small p, the* $a=1$ *applicant has a strictly greater Round 2 acceptance probability.*

Proof. See Appendix F.

F Proofs of Theorems

Theorem 1. *For simplicity, suppose there is one round. The maximum likelihood estimate $\hat{P}_{y^*} = \hat{P}(y^* = S|t)$ is such that the number of acceptances N_A equals the expected N_A. Also, for any $\hat{P}_{y^*}$ and N_A, $\mathcal{L}^R$ is highest when the committee accepts the N_A highest $P_{y_n} = P(y_n = S|x_n, z_n)$.*

Proof. Supposing one round, the restricted likelihood is:

$$\mathcal{L}^R = \prod_{n=1}^{N} \left[\frac{e^{P_{y_n}/\sigma}}{e^{P_{y_n}/\sigma} + e^{P_{y^*}/\sigma}}\right]^{\mathbb{1}(d_n=A)} \left[\frac{e^{P_{y^*}/\sigma}}{e^{P_{y_n}/\sigma} + e^{y^*/\sigma}}\right]^{\mathbb{1}(d_n=R)} \left[\frac{e^{P_{y_n}}}{1+e^{P_{y_n}}}\right]^{\mathbb{1}(y_n=S)} \left[\frac{1}{1+e^{P_{y_n}}}\right]^{\mathbb{1}(y_n=F)} \tag{43}$$

When an applicant is accepted, the resulting term in $\mathcal{L}^R$ decreases in P_{y^*}, providing downward pressure on $\hat{P}_{y^*}$. Also, when one is rejected, the term increases in P_{y^*}, providing upward pressure.

Since $\mathbb{1}(d_n = A) + \mathbb{1}(d_n = R) = \mathbb{1}(y_n = S) + \mathbb{1}(y_n = F) = 1$, the restricted log-likelihood is:

$$\begin{aligned}
\ell^R = \sum_{n=1}^{N} \Bigg[& \mathbb{1}(d_n = A) \log\left(\frac{e^{P_{y_n}/\sigma}}{e^{P_{y_n}/\sigma} + e^{P_{y^*}/\sigma}}\right) + \mathbb{1}(d_n = R) \log\left(\frac{e^{P_{y^*}/\sigma}}{e^{P_{y_n}/\sigma} + e^{P_{y^*}/\sigma}}\right) \\
& + \mathbb{1}(y_n = S) \log\left(\frac{e^{P_{y_n}}}{1+e^{P_{y_n}}}\right) + \mathbb{1}(y_n = F) \log\left(\frac{1}{1+e^{P_{y_n}}}\right)\Bigg] \\
= \sum_{n=1}^{N} \Bigg[& \frac{1}{\sigma}\left(\mathbb{1}(d_n = A)P_{y_n} + \mathbb{1}(d_n = R)P_{y^*}\right) - \log\left(e^{P_{y_n}/\sigma} + e^{P_{y^*}/\sigma}\right) + \mathbb{1}(y_n = S)P_{y_n} - \log\left(1 + e^{P_{y_n}}\right)\Bigg].
\end{aligned} \tag{44}$$

Since ℓ^R is strictly concave in P_{y^*}, I find the unique maximizer $\hat{P}_{y^*}$ by the first-order condition:

$$\frac{1}{\sigma}\left[N - N_A - \sum_{n=1}^{N} \frac{e^{\hat{P}_{y^*}/\sigma}}{e^{P_{y_n}/\sigma} + e^{\hat{P}_{y^*}/\sigma}}\right] = 0 \quad \Rightarrow \quad \sum_{n=1}^{N} \frac{e^{P_{y_n}/\sigma}}{e^{P_{y_n}/\sigma} + e^{\hat{P}_{y^*}/\sigma}} = N_A. \tag{45}$$

The left side of the latter equation is each applicant's acceptance probability summed over the number of applicants; that is, expected N_A. Therefore, $\hat{P}_{y^*}$ is such that the expected N_A equals N_A.

Given that $\hat{P}_{y^*}$ is identified, I will show that ℓ^R, and therefore $\mathcal{L}^R$, are highest when the committee accepts the N_A applicants with the highest P_{y_n}. The only part of ℓ^R that depends on d_n is $\sum_{n=1}^{N}[\mathbb{1}(d_n = A)P_{y_n} + \mathbb{1}(d_n = R)P_{y^*}] = NP_{y^*} + \sum_{n=1}^{N} \mathbb{1}(d_n = A)(P_{y_n} - P_{y^*})$; this equality holds since $\mathbb{1}(d_n = A) + \mathbb{1}(d_n = R) = 1$. As such, the only part depending on d_n is $\sum_{n=1}^{N} \mathbb{1}(d_n = A)(P_{y_n} - P_{y^*})$.

This expression is highest when $d_n = A$ for the N_A highest values of P_{y_n}, so the likelihood is highest when the committee accepts the N_A highest P_{y_n}. Intuitively, the committee's selectivity identifies $\hat{P}_{y^*}$, and it is not necessary for identification for the committee to accept applicants with higher P_{y_n}. However, if the committee engages in positive sorting, the model's fit will improve.

Theorem 2. *In the waitlist-hybrid and dynamic models, success probability equal, an applicant's Round 2 acceptance probability is strictly increasing in their matriculation score.*

Proof. Suppose that there are two arbitrary applicants in year t indexed by H and L, such that $P(y=S \mid x^H, z^H) = P(y=S \mid x^L, z^L)$, but $m_2^H > m_2^L$. That is, the applicants have equal success probabilities, but H has the higher matriculation score. Because $\ddot{m}_2$ is defined as the average matriculation score of all the applicants in t with m_2 excluded, $\ddot{m}_2^H < \ddot{m}_2^L$ by construction. Given this inequality, I will show that $P(d_2 = A \mid x^H, z^H, \ddot{m}_2^H) > P(d_2 = A \mid x^L, z^L, \ddot{m}_2^L)$. It is the case that:

$$P(d_2 = A \mid x, z, \ddot{m}_2) = \frac{\exp(P(y=S \mid x,z)/\sigma)}{\exp(P(y=S \mid x,z)/\sigma) + \exp(\beta \mathbb{E}[v_3(x,z,\overline{m}_{2A},\varepsilon_3) \mid x,z,\ddot{m}_2]/\sigma)}, \text{ where}$$
$$\mathbb{E}[v_3(x,z,\overline{m}_{2A},\varepsilon_3) \mid x,z,\ddot{m}_2] = \int_{\widetilde{\overline{m}_{2A}}} \sigma \log\left(e^{P(y=S|x,z)/\sigma} + e^{P(y_3^*=S|\widetilde{\overline{m}_{2A}})/\sigma}\right) dF_{\overline{m}_{2A}|\ddot{m}_2}(\widetilde{\overline{m}_{2A}} \mid \ddot{m}_2). \tag{46}$$

The only place $\ddot{m}_2$ appears in $P(d_2 = A \mid x, z, \ddot{m}_2)$ is in the distribution of $\overline{m}_{2A}$ conditional on $\ddot{m}_2$. I parameterize $\overline{m}_{2A} \mid \ddot{m}_2 \sim \mathcal{N}(\ddot{m}_2 \beta_{\ddot{m}_2}, \sigma^2_{\ddot{m}_2})$, where by assumption $\beta_{\ddot{m}_2} > 0$. $\beta_{\ddot{m}_2}$ is positive because years when $\ddot{m}_2$ is high are years when matriculation scores are high in general. Exploiting inter-year variation, this relationship results in a greater expected $\overline{m}_{2A}$ in those years. Then:

$$F_{\overline{m}_{2A}|\ddot{m}_2}(\overline{m}_{2A} \mid \ddot{m}_2^H; \beta_{\ddot{m}_2}, \sigma^2_{\ddot{m}_2}) = \Phi\left(\frac{\overline{m}_{2A} - \ddot{m}_2^H \beta_{\ddot{m}_2}}{\sigma_{\ddot{m}_2}}\right) > \Phi\left(\frac{\overline{m}_{2A} - \ddot{m}_2^L \beta_{\ddot{m}_2}}{\sigma_{\ddot{m}_2}}\right) = F_{\overline{m}_{2A}|\ddot{m}_2}(\overline{m}_{2A} \mid \ddot{m}_2^L; \beta_{\ddot{m}_2}, \sigma^2_{\ddot{m}_2}). \tag{47}$$

Therefore, $\overline{m}_{2A} \mid \ddot{m}_2^L$ first-order stochastically dominates $\overline{m}_{2A} \mid \ddot{m}_2^H$. To apply this stochastic dominance, I must show that $\sigma \log\left(e^{P(y=S|x,z)/\sigma} + e^{P(y_3^*=S|\overline{m}_{2A})/\sigma}\right)$ is strictly increasing in $\overline{m}_{2A}$. But the only place $\overline{m}_{2A}$ appears in this expression is in $P(y_3^* = S \mid \overline{m}_{2A}; \theta_{S_3^*}) = \frac{\exp\left(f_{S_3^*}(\overline{m}_{2A})\theta_{S_3^*}\right)}{1+\exp\left(f_{S_3^*}(\overline{m}_{2A})\theta_{S_3^*}\right)}$, where by assumption $f_{S_3^*}(\cdot)$ is strictly increasing and $\theta_{S_3^*} > 0$. $\theta_{S_3^*}$ is positive because the expected Round 3 cutoff probability is higher when early acceptees are more likely on average to matriculate. As such, I have that $P(y_3^* = S \mid \overline{m}_{2A})$, and in turn $\sigma \log\left(e^{P(y=S|x,z)/\sigma} + e^{P(y_3^*=S|\overline{m}_{2A})/\sigma}\right)$, are strictly increasing in $\overline{m}_{2A}$. Combining this result with $(\overline{m}_{2A} \mid \ddot{m}_2^L) \succ_{\text{FOSD}} (\overline{m}_{2A} \mid \ddot{m}_2^H)$, I obtain:

$$\begin{aligned}
\mathbb{E}[v_3(x^L,z^L,\overline{m}_{2A},\varepsilon_3) \mid x^L,z^L,\ddot{m}_2^L] &= \int_{\widetilde{\overline{m}_{2A}}} \sigma \log\left(e^{P(y=S|x^L,z^L)/\sigma} + e^{P(y_3^*=S|\widetilde{\overline{m}_{2A}})/\sigma}\right) dF_{\overline{m}_{2A}|\ddot{m}_2}(\widetilde{\overline{m}_{2A}} \mid \ddot{m}_2^L) \\
&= \int_{\widetilde{\overline{m}_{2A}}} \sigma \log\left(e^{P(y=S|x^H,z^H)/\sigma} + e^{P(y_3^*=S|\widetilde{\overline{m}_{2A}})/\sigma}\right) dF_{\overline{m}_{2A}|\ddot{m}_2}(\widetilde{\overline{m}_{2A}} \mid \ddot{m}_2^L) \\
&> \int_{\widetilde{\overline{m}_{2A}}} \sigma \log\left(e^{P(y=S|x^H,z^H)/\sigma} + e^{P(y_3^*=S|\widetilde{\overline{m}_{2A}})/\sigma}\right) dF_{\overline{m}_{2A}|\ddot{m}_2}(\widetilde{\overline{m}_{2A}} \mid \ddot{m}_2^H) \\
&= \mathbb{E}[v_3(x^H,z^H,\overline{m}_{2A},\varepsilon_3) \mid x^H,z^H,\ddot{m}_2^H].
\end{aligned} \tag{48}$$

The second equality in this chain holds because $P(y=S \mid x^H, z^H) = P(y=S \mid x^L, z^L)$. Therefore, by Equation (46) and Inequality (48), I have that $P(d_2 = A \mid x^H, z^H, \ddot{m}_2^H) > P(d_2 = A \mid x^L, z^L, \ddot{m}_2^L)$. In the full version of the waitlist-hybrid model, there are six additional states: $(n_r, \overline{q}_r, \overline{m}_r)$ for $r \in (2,3)$, and in the full version of the waitlist-dynamic model, there are nine: $(n_r, \overline{q}_r, \overline{m}_r)$ for

$r \in (1,2,3)$. However, these states do not vary across applicants within a year, so the proof that $P(d_2 = A \mid x^H, z^H, \ddot{m}_2^H, n_2, \bar{q}_2, \overline{m}_2) > P(d_2 = A \mid x^L, z^L, \ddot{m}_2^L, n_2, \bar{q}_2, \overline{m}_2)$ is analogous to this proof.

Theorem 3. *Let d^R be the model's decision rule with $\sigma = 0$ in the CCPs, and let d^U be the committee's decision rule. By optimality:*

$$\frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S \middle| x_n, z_n; \hat{\theta}_S^U\right) \mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^R\left[P_{y_n}\left(\hat{\theta}_S^U\right)\right] = A\right\}$$
$$\geq \frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S \middle| x_n, z_n; \hat{\theta}_S^U\right) \mathbb{1}\left\{d_{1,n}\left[P_{y_n}\left(\hat{\theta}^U\right)\right] = T, d_{2,n}^U\left[P_{y_n}\left(\hat{\theta}_A^U\right)\right] = A\right\} \quad (49)$$

Proof. This theorem states that conditional on a fixed Round 1 decsion rule $d_1[P_y(\hat{\theta}^U)]$ and corresponding transition set of exogenously-determined size N_T, the sample average success probability of the also-exogenous N_A applicants that the model accepts in Round 2 weakly exceeds the sample average success probability of the N_A applicants that any actual committee accepts in Round 2. The inequality holds since the model's Round 2 decision rule is to accept the N_A applicants with the highest success probabilities. Therefore, the sample average success probability of the model's acceptances must weakly exceed that of any other set of N_A acceptances.

However, the inequality would not always hold if it were written as follows (changes in bold):

$$\frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S \middle| x_n, z_n; \hat{\theta}_S^U\right) \mathbb{1}\left\{\mathbf{d_{1,n}^R}\left[\mathbf{P_{y_n}}\left(\hat{\boldsymbol{\theta}}_{\mathbf{S}}^{\mathbf{U}}\right)\right] = \mathbf{T}, d_{2,n}^R\left[P_{y_n}\left(\hat{\theta}_S^U\right)\right] = A\right\}$$
$$\textbf{vs. } \frac{1}{N_A}\sum_{n=1}^{N} P\left(y_n = S \middle| x_n, z_n; \hat{\theta}_S^U\right) \mathbb{1}\left\{\mathbf{d_{1,n}^U}\left[\mathbf{P_{y_n}}\left(\hat{\boldsymbol{\theta}}_{\mathbf{T}}^{\mathbf{U}}\right)\right] = \mathbf{T}, d_{2,n}^U\left[P_{y_n}\left(\hat{\theta}_A^U\right)\right] = A\right\} \quad (50)$$

The difference is that the Round 1 transition set is no longer fixed. Now, the model's transition set is based on choosing the N_T applicants in each year with the highest $\mathbb{E}[v_2(x,z,\varepsilon_2)|x] = \int_{\tilde{z}} \sigma \log(e^{P(y=S|x,z)/\sigma} + e^{P(y_2^*=S)/\sigma}) dF(\tilde{z}|x)$. In contrast, the actual committee may have a different Round 1 decision rule. Crucially, although the model is choosing applicants akin to those with the highest expected success probabilities, it is ignoring any value added from upside potential.

For example, suppose $N_A = 1$, and there is an applicant in Round 1 with a low expected success probability but a known high degree of variance in the distribution of z conditional on their x_n. That is, $F(z|x)$ is heteroskedastic. Also suppose that their low expected success probability yields an $\mathbb{E}[v_2(x,z,\varepsilon_2)|x]$ too low to survive Round 1, even though their actual success probability ends up being the highest of any applicant once z is realized in Round 2. If the actual committee's decision rule considers upside potential, then the actual committee may accept this applicant, while the model will not. In most cases, the model's Round 1 decision rule should lead to a better acceptance set than an actual committee's rule. But such an outcome is not guaranteed.

Theorem 4. *In the multiyear-dynamic matriculation model, success probability equal, an applicant's Round 2 acceptance probability is strictly increasing in their matriculation score.*

Proof. Suppose there are two arbitrary applicants in year t indexed by H and L, such that $P(y=S \mid x^H, z^H) = P(y=S \mid x^L, z^L)$, but $m^H > m^L$. That is, H has the higher matriculation score. I will show that $P(d_2 = A \mid x^H, z^H, m^H, \overline{m}, \overline{q}, M, c_{-1}) > P(d_2 = A \mid x^L, z^L, m^L, \overline{m}, \overline{q}, M, c_{-1})$. The states $(\overline{m}, \overline{q}, M, c_{-1})$ do not vary across applicants within a year.

Define for each round the choice-specific payoffs for acceptance (A) and rejection (R):

$$\pi_2^A \equiv P(y=S \mid x, z) + \beta \mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m, \overline{m}], \quad \pi_2^R \equiv L(y_2^* = S \mid \overline{q}, M, c_{-1}) + \beta \mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid \overline{m}],$$

$$\pi_1^A \equiv \mathbb{E}[v_2(x', z', m', \overline{m}', \overline{q}', M', c, \varepsilon_2') \mid x', c], \quad \pi_1^R \equiv L(y_1^{*\prime} = S \mid t') + \mathbb{E}[v_1(x'', t'', c', \varepsilon_1'') \mid \overline{m}']. \quad (51)$$

It is the case that:

$$P(d_2 = A \mid x, z, m, \overline{m}, \overline{q}, M, c_{-1}) = \frac{\exp(\pi_2^A/\sigma)}{\exp(\pi_2^A/\sigma) + \exp(\pi_2^R/\sigma)}, \text{ where:}$$

$$\mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m, \overline{m}] = \int_c \int_{t'} \int_{x'} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(x')\, dF_t(t')\, dF_{c|m,\overline{m}}(c \mid m, \overline{m}), \text{ where:}$$

$$\mathbb{E}[v_2(x', z', m', \overline{m}', \overline{q}', M', c, \varepsilon_2') \mid x', c] = \int_{\widetilde{M'}} \int_{\widetilde{\overline{q}'}} \int_{\widetilde{\overline{m}'}} \int_{\widetilde{m'}} \int_{\widetilde{z'}} \sigma \log\left(e^{\pi_2^{A\prime}/\sigma} + e^{\pi_2^{R\prime}/\sigma}\right)$$
$$dF_{z|x}(\widetilde{z'} \mid x') dF_{m|x}(\widetilde{m'} \mid x')\, dF_{\overline{m}}(\widetilde{\overline{m}'})\, dF_{\overline{q}}(\widetilde{\overline{q}'})\, dF_M(\widetilde{M'}), \quad (52)$$

where $\pi_2^{A\prime}$ and $\pi_2^{R\prime}$ are defined analogously to Equation (51) at the next year's states.

Per-period utilities are bounded and $\beta \in (0,1)$, so the value functions exist and are unique. The only place m appears in $P(d_2 = A \mid x, z, m, \overline{m}, \overline{q}, M, c_{-1})$ is in the distribution of c conditional on $(m, \overline{m})$. I parameterize $c \mid (m, \overline{m}) \sim \mathcal{N}\left(\begin{bmatrix} m & \overline{m} \end{bmatrix} \beta_{m,\overline{m}}, \sigma_{m,\overline{m}}^2\right)$, where by assumption $\beta_m > 0$. β_m is positive because higher matriculation scores lead to a greater expected cohort size. Multiple years of data are necessary to identify β_m, since c varies only across years. Then:

$$F_{c|(m,\overline{m})}(c \mid (m^H, \overline{m}); \beta_{m,\overline{m}}, \sigma_{m,\overline{m}}^2) = \Phi\left(\frac{c - \begin{bmatrix} m^H & \overline{m} \end{bmatrix} \beta_{m,\overline{m}}}{\sigma_{m,\overline{m}}}\right)$$
$$< \Phi\left(\frac{c - \begin{bmatrix} m^L & \overline{m} \end{bmatrix} \beta_{m,\overline{m}}}{\sigma_{m,\overline{m}}}\right) = F_{c|(m,\overline{m})}(c \mid (m^L, \overline{m}); \beta_{m,\overline{m}}, \sigma_{m,\overline{m}}^2). \quad (53)$$

Therefore, $c \mid (m^H, \overline{m})$ first-order stochastically dominates $c \mid (m^L, \overline{m})$. To apply this stochastic dominance, I must show that $\mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m, \overline{m}]$'s integrand consisting of all but the outermost integral is strictly increasing in c. But the only place c appears there is in $\mathbb{E}[v_2(x', z', m', \overline{m}', \overline{q}', M', c, \varepsilon_2') | x', c]$, and the only place c appears in that expectation is in $L(y_2^{*\prime} = S \mid \overline{q}', M', c; \theta_{S_2^*}) = f_{S_2^*}\left(\begin{bmatrix} \overline{q}' & \overline{M}' & c \end{bmatrix}\right) \theta_{S_2^*}$.

By assumption, $f_{S_2^*}(\cdot)$ is strictly increasing and $\theta_{S_2^*} > 0$. $\theta_{S_2^*, c}$ is positive because a greater cohort size this year means that the committee, needing to keep the total number of students in the program relatively constant, can afford to be more selective next year. As such, I have that $P(y_2^{*\prime} = S \mid \overline{q}', M', c)$, and in turn $\mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m, \overline{m}]$'s integrand, are strictly increasing in c. Combining this result with $[c \mid (m^H, \overline{m})] \succ_{FOSD} [c \mid (m^L, \overline{m})]$, I obtain the inequality:

$$\begin{aligned}
\mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m^H, \overline{m}] &= \int_c \int_{t'} \int_{x'} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(x')\, dF_t(t')\, dF_{c|m,\overline{m}}(c \mid m^H, \overline{m}) \\
&> \int_c \int_{t'} \int_{x'} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(x')\, dF_t(t')\, dF_{c|m,\overline{m}}(c \mid m^L, \overline{m}) \\
&= \mathbb{E}[v_1(x', t', c, \varepsilon_1') \mid m^L, \overline{m}].
\end{aligned} \tag{54}$$

Therefore, by Equation (52), Inequality (54), and $P(y = S \mid x^H, z^H) = P(y = S \mid x^L, z^L)$, I have that $P(d_2 = A \mid x^H, z^H, m^H, \overline{m}, \overline{q}, M, c_{-1}) > P(d_2 = A \mid x^L, z^L, m^L, \overline{m}, \overline{q}, M, c_{-1})$. Intuitively, if the committee rejects either applicant, their m does not matter. But since their success probabilities are equal, accepting applicant H yields a higher expected payoff than accepting applicant L.

Theorem 5. *Consider any two applicants, one with $a = 1$ and one with $a = 0$, but with equal success probabilities and therefore different values of x or z. Suppose $P(a' = 1) = p > 0$. Then for all sufficiently small p, the $a = 1$ applicant has a strictly greater Round 2 acceptance probability.*

Proof. Suppose there are two arbitrary applicants in year t, one with $a = 1$ and the other with $a = 0$, but with $P(y = S \mid x^1, z^1, 1 - \overline{a}_{-1}) = P(y = S \mid x^0, z^0, \overline{a}_{-1})$. Unlike in the previous two proofs, these equal success probabilities imply that unless $\overline{a}_{-1} = \frac{1}{2}$, it must be that $(x^1, z^1) \neq (x^0, z^0)$. I will show that $P(d_2 = A \mid x^1, z^1, 1, \overline{a}_{-1}, t) > P(d_2 = A \mid x^0, z^0, 0, \overline{a}_{-1}, t)$.

Define for each round the choice-specific payoffs for acceptance (A) and rejection (R):

$$\pi_2^A \equiv P(y = S \mid x, z, |a - \overline{a}_{-1}|) + \beta \mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1') \mid a], \quad \pi_2^R \equiv L(y_2^* = S \mid t) + \beta \mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1')],$$

$$\pi_1^A \equiv \mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t'], \quad \pi_1^R \equiv L(y_1^{*\prime} = S \mid t') + \beta \mathbb{E}[v_1(x'', \overline{a}', t'', \varepsilon_1'')]. \tag{55}$$

It is the case that:

$$P(d_2 = A \mid x, z, a, \overline{a}_{-1}, t) = \frac{\exp(\pi_2^A/\sigma)}{\exp(\pi_2^A/\sigma) + \exp(\pi_2^R/\sigma)}, \text{ where:}$$

$$\mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1') \mid a] = \int_{\widetilde{\overline{a}}} \int_{\widetilde{t'}} \int_{\widetilde{x'}} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(\widetilde{x'})\, dF_t(\widetilde{t'})\, dF_{\overline{a}|a}(\widetilde{\overline{a}} \mid a), \text{ where:}$$

$$\begin{aligned}
\mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t'] &= P(a' = 1) \int_{\widetilde{z'}} \sigma \log\left(e^{\pi_2^{A\prime}(1)/\sigma} + e^{\pi_2^{R\prime}/\sigma}\right) dF_{z|x}(\widetilde{z'} \mid x') \\
&\quad + [1 - P(a' = 1)] \int_{\widetilde{z'}} \sigma \log\left(e^{\pi_2^{A\prime}(0)/\sigma} + e^{\pi_2^{R\prime}/\sigma}\right) dF_{z|x}(\widetilde{z'} \mid x'),
\end{aligned} \tag{56}$$

where $\pi_2^{A\prime}(a')$ denotes next year's Round 2 acceptance payoff evaluated at $a' \in \{0, 1\}$.

Per-period utilities are bounded and $\beta \in (0, 1)$, so the value functions exist and are unique. The only two places a appears in $P(d_2 = A \mid x, z, a, \overline{a}_{-1}, t)$ are in the success probabilities, which are assumed to be equal for the $a=1$ and $a=0$ applicants, and in the distribution of $\overline{a}$ conditional on a. I parameterize $\overline{a} \mid a \sim \mathcal{N}(a\beta_a, \sigma_a^2)$, where by assumption $\beta_a > 0$. β_a is positive because $a=1$ applicants are more likely than $a=0$ applicants to be in a year with a high percentage of accepted $a=1$ applicants. As such, multiple years of data are necessary to identify β_a. Then:

$$F_{\overline{a}|a}(\overline{a} \mid 1; \beta_a, \sigma_a^2) = \Phi\left(\frac{\overline{a}-\beta_a}{\sigma_a}\right) < \Phi\left(\frac{\overline{a}}{\sigma_a}\right) = F_{\overline{a}|a}(\overline{a} \mid 0; \beta_a, \sigma_a^2). \tag{57}$$

Therefore, $\overline{a} \mid 1$ first-order stochastically dominates $\overline{a} \mid 0$. To apply this stochastic dominance, I must show that $\mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1') \mid a]$'s integrand consisting of all but the outermost integral is strictly increasing in $\overline{a}$. But the only place $\overline{a}$ appears there is in $\mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t']$, and the only place $\overline{a}$ appears in that expectation is in $P(y' = S \mid x', z', \overline{a}; \theta_S) = \frac{\exp(f_S(x', z', |a'-\overline{a}|)\theta_S)}{1+\exp(f_S(x', z', |a'-\overline{a}|)\theta_S)}$.

By assumption, $f_S(\cdot)$ is strictly increasing and $\theta_{S,|a-\overline{a}_{-1}|} > 0$. $\theta_{S,|a-\overline{a}_{-1}|}$ is positive because a large value of $|a-\overline{a}_{-1}|$ means that applicant a is diverse relative to the previous year's accepted applicants and is therefore more likely to succeed. Recall $P(a'=1)=p$. Taking the limit $p \to 0$:

$$\begin{aligned}
&\mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t'] \\
&\quad \to \int_{\widetilde{z'}} \sigma \log(e^{\{P(y'=S|x',z',\overline{a})+\beta\mathbb{E}[v_1(x'',\overline{a}',t'',\varepsilon_1'')|0]\}/\sigma} + e^{\{L(y_2^{*\prime}=S|t')+\beta\mathbb{E}[v_1(x'',\overline{a}',t'',\varepsilon_1'')]\}/\sigma}) dF_{z|x}(\widetilde{z'} \mid x').
\end{aligned} \tag{58}$$

$P(y'=S \mid x', z', \overline{a})$ is clearly strictly increasing in $\overline{a}$, and by the continuity of $E[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t']$ in p, I have that it, and in turn $E[v_1(x', \overline{a}, t', \varepsilon_1') \mid a]$'s integrand, are also strictly increasing in $\overline{a}$ for sufficiently small p. Combining this result with $\overline{a} \mid 1 \succ_{FOSD} \overline{a} \mid 0$, I obtain the inequality:

$$\begin{aligned}
\mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1') \mid 1] &= \int_{\widetilde{\overline{a}}} \int_{\widetilde{t'}} \int_{\widetilde{x'}} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(\widetilde{x'})\, dF_t(\widetilde{t'})\, dF_{\overline{a}|a}(\widetilde{\overline{a}} \mid 1) \\
&> \int_{\widetilde{\overline{a}}} \int_{\widetilde{t'}} \int_{\widetilde{x'}} \sigma \log\left(e^{\pi_1^A/\sigma} + e^{\pi_1^R/\sigma}\right) dF_x(\widetilde{x'})\, dF_t(\widetilde{t'})\, dF_{\overline{a}|a}(\widetilde{\overline{a}} \mid 0) \\
&= \mathbb{E}[v_1(x', \overline{a}, t', \varepsilon_1') \mid 0].
\end{aligned} \tag{59}$$

Therefore, by Equation (56), Inequality (59), and $P(y=S \mid x^1, z^1, 1-\overline{a}_{-1}) = P(y=S \mid x^0, z^0, \overline{a}_{-1})$, I have that $P(d_2 = A \mid x^1, z^1, 1, \overline{a}_{-1}, t) > P(d_2 = A \mid x^0, z^0, 0, \overline{a}_{-1}, t)$. Intuitively, if the committee rejects either applicant, their a does not matter. But since their success probabilities are equal, accepting applicant $a=1$ yields a higher expected payoff than accepting $a=0$.

In this proof, I must assume that $P(a'=1)$ is sufficiently small, and not just $P(a'=1) < \frac{1}{2}$. The reason is that if $P(a'=1)$ is too large, then $\mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t']$ may not be strictly increasing in $\overline{a}$, considering $P(y'=S \mid x', z', 1-\overline{a})$ is strictly decreasing in $\overline{a}$. This indeterminancy can be explored by calculating $\frac{\partial}{\partial \overline{a}}\mathbb{E}[v_2(x', z', a', \overline{a}, t', \varepsilon_2') \mid x', \overline{a}, t']$. As such, to be sure the committee is more likely to accept the $a=1$ applicant, the committee must know that there will be sufficiently few $a'=1$ applicants. This assumption is strong, but possible if $a'=1$ is sufficiently rare for the committee to know this will happen. However, even without forward-looking dynamics, if $\overline{a}_{-1} < \frac{1}{2}$, then the committee will be more likely to accept an $a=1$ applicant than an $a=0$ applicant with identical x and z values. The reason is that $P(y=S \mid x, z, 1-\overline{a}_{-1}) > P(y=S \mid x, z, \overline{a}_{-1})$.