OSIPP Discussion Paper : DP-2025-E-002

The Effect of Coarse Score Labels on College Application Decisions*

June 6, 2025

Tomoharu Mori[1]

Associate Professor, College of Comprehensive Psychology, Ritsumeikan University

Katsuya Takii[2]

Professor, Osaka School of International Public Policy, Osaka University

【JEL Classification】C21, D91, I23

【Keywords】Inattention, college application, regression discontinuity, score labels

【Abstract】This study examines whether people are inattentive to once-in-a-lifetime events. Using data from Japanese entrance examinations, we show that, even in situations where there appears to be little difference in the actual admissions probability, students change their school of choice simply because the coarse label describing admissions probability has changed. To understand the mechanism behind this result, we model students' application decisions by incorporating inattention. The results suggest that, in once-in-a-lifetime situations, students do observe information more carefully than usual, but they still cannot be completely attentive. We also find that people with better information processing skills and who are accustomed to processing more information daily pay greater attention to information.

*The authors would like to thank the participants of the Applied Economics Workshop at Keio University, Contract Theory Workshop, Kansai Labor Seminar, OSIPP Economics Brownbag Seminar, RISS Workshop, Seminar at Osaka City University, Tokyo Labor Economics Workshop and Workshop on the Economics of Human Resource Allocation for useful comments. Financial support from the Japanese Ministry of Education, Culture, Sports, Science, and Technology (Grants-in-Aid for Scientific Research Nos 25245040, 18H03636 and 24K00265), and a Research Grant from Murata Science and Education Foundation are gratefully acknowledged.

[1] tmr15047@fc.ritsumei.ac.jp

[2] takii@osipp.osaka-u.ac.jp

1 Introduction

Growing evidence suggests that people are inattentive to opaque information (e.g., DellaVigna (2009), Maćkowiak et al. (2023)) in several economic situations, such as inattention to taxes (e.g., Chetty et al. (2009)), inattention in health plan choices (e.g., Heiss et al. (2021)) and inattention in the financial market (e.g., DellaVigna and Pollet (2009)). Gabaix (2019) attempts to unify literature and proposed a tractable measure of attention parameters, which is 0 if people do not pay attention to information at all and 1 if people perfectly pay attention to information. Reviewing the literature using this measure, Gabaix (2019) shows that there is substantial heterogeneity in the degree of attention parameters found in the previous literature. Building on this, Taubinsky and Rees-Jones (2018) and Morrison and Taubinsky (2023) examine the effect of incentive strength on the intensity of attention. This finding suggests that attention varies depending on the incentives triggered by a problem. The aforementioned raises a natural question: if the decision is a once-in-a-lifetime event, wouldn't individuals pay closer attention to the available information?

This study examines how coarse information regarding admission probabilities in entrance examinations for higher education influences students' application decisions when more precise information is available. In many countries, the market for higher education is decentralized without any flexible price mechanisms. Because the capacity of desirable colleges is typically smaller than the number of applicants in these countries, students must consider the probability of admission when deciding to which colleges to apply. Since the choice of which college to enroll in significantly impacts the rest of one's life, information regarding admission probability is important when students decide which school to apply to. This study examines whether students are still inattentive to such important information by constructing a model of inattention and estimating its parameters. We also assess how attention parameters differed among the different types of students.

For this purpose, we examine data that helped students infer the admission probability for the entrance exam and their choice of school in the Japanese university entrance examination. As we discuss in detail later, Japan's public university entrance examinations are divided into two stages: the first stage is the National Center Test for University Admission[1]—a test common to all public universities (held in mid-January). The second stage is each university's individual entrance examination (held in late February).

Depending on their performance in the first stage, students choose a one-degree program to apply to, which comprises a university and department pair. To make application decisions, it is important to know their relative academic standing and chances of success. In Japan, major university prep schools collect information on students' first-stage examination results and desired degree programs. Based on the collected data, these prep schools provide each student with information to infer the probability of admission to their desired degree programs. We investigate whether students changed their application behavior in their first-choice degree program after obtaining this new information.

A significant feature of the information returned by prep schools to students is that it contains coarse information on admission probability and precise information in the same place. One such piece of information is the score label (called "hantei") indicating the likelihood of success, represented on a scale of A, B, C, D, and E, where A indicates the highest and E the lowest admission probability area. The other piece of information provides more details showing the distribution of the applicants' scores (histogram), the location of each student, and the threshold of admission probability. The latter detailed piece of information encompasses the former coarse information and helps to more precisely understand the probability of admission to their desired degree programs.

Using these data, we investigate how students around a cutoff point change their application behavior from the degree program of first choice using a regression discontinuity design (RDD). Estimation results from the regression discontinuity design show that students' behavior "jumps" around the threshold. The jump is statistically significant at the cutoffs of B/C, C/D and D/E, but not that of A/B, where A/B means "between A and B." Around the cutoffs of B/C, C/D, and D/E, the probability of applying to the first-choice degree program decreases by 4.3%–6.7% due to the change in the score label from a higher probability label to a lower probability label. Simply, we find evidence that coarse information significantly

[1]The name was changed to "The Common Test for University Admissions" in 2020.

influences individual behavior even when precise information is available.

We also examine whether the probability of admission changed as a result of changes in application behavior. To this end, we implement a Fuzzy RDD design using changes in labels as instrumental variables. The results show that there are no statistically significant effects of changes in application behavior on admission probability around any cutoffs. If it is reasonable to assume that admission to a student's first-choice degree program yields higher utility, then this evidence suggests that changes in application behavior reduce the expected utility.

To understand the mechanism behind the evidence obtained, we model students' application decisions when they have a limited capacity for attention in the spirit of Gabaix (2014) and Gabaix (2019). We assume that, if students have more capacity for attention, their subjective probability of success is more influenced by precise information on the admission probability; if students have less capacity, their subjective probability is more influenced by coarse information. We parameterize the degree of attention and analyze how the attention parameter influences the students' application decisions. It is shown that the observed "jumps" can be considered as evidence that the students are subjected to the limited capacity of attention.

Furthermore, this model allows us to decompose the "jumps" obtained in the regression discontinuity design into two components: (1) An index measuring how much the subjective probability of admission to the first-choice degree program changes when the label shifts, relative to changes in the objective probability of admission to the first-choice degree program and (2) an index measuring how sensitive students are in changing their school choice when the objective admission probability changes. We refer to index (1) as the sensitivity of the subjective probability to label changes and index (2) as the sensitivity of the application behavior to changes in the objective probability. This theory also shows that the sensitivity of the subjective probability to label changes comprises two factors: the degree of inattention and the expected admission probability change from label changes. This decomposition helps us quantify the inattention's role in the obtained discrete "jumps."

Finally, we show that each component of the "jumps" can be identified using standard binary choice estimation, where the outcome takes the value of one when students change their application behavior from their first-choice degree program. In particular, the model shows that the sensitivity of subjective probability to label changes is identified by the coefficient of changes in the label relative to that of the percentage changes in the objective admission probability in the binary choice model, and that the sensitivity of application behavior to changes in objective probability is identified by the average marginal effects of the percentage changes in the objective admission probability in the binary choice model.

The results generated by our logit estimation show that we cannot reject the hypothesis that the theoretical prediction of the magnitude of the "jumps" is the same as that obtained by the RDD. While we also conduct several sub-sample analyses, the theoretical predictions of the magnitude of the "jumps" passed the validity test used by RDD for all sub-samples. This result greatly increases confidence in the validity of our model.

Using this model, we estimate the sensitivity of the subjective probability to label changes and the sensitivity of the application behavior to changes in the objective probability for several sub-samples. We find that the sensitivity of the subjective probability to label changes is significantly higher for the low-test-score group than for the high-test-score group and significantly greater for students living in non-metropolitan areas than for those living in metropolitan areas. The sensitivity of application behavior to changes in objective probability is significantly higher for the low-test-score group than for the high-test-score group, but significantly greater for students living in metropolitan areas than for those living in non-metropolitan areas.

Because the test score groups comprise students with high academic skills and metropolitan areas where many university prep schools are located, we interpret these results as follows: People with a greater ability and/or more experience in processing information tend to pay attention to more precise information. Conversely, while individuals with higher cognitive abilities do not respond as strongly to precise information, those with more experience in processing information do. These results suggest that it is important to educate students with lower cognitive skill levels and/or those living in areas with limited admissions information on how to interpret admissions information.

This study also demonstrates that, once the expected change in admission probability owing to label changes is known, the attention parameters can be identified. We calibrate the expected rate of change and estimated the magnitudes of the attention parameters.

We find that the average value of the attention parameter is 0.726 for the B/C cutoff and 0.667 for the C/D cutoff. Our sub-sample analysis shows a statistically significant difference in attention parameters between the low- and high-test-score groups, as well as between students living in non-metropolitan and metropolitan areas. The attention parameter ranged from 0.594 for students in non-metropolitan areas (for the C/D cutoff) to 0.854 for those in metropolitan areas (for the B/C cutoff).

Several studies provide attention parameters for the same spirit (e.g., Hossain and Morgan (2006), Chetty et al. (2009), DellaVigna and Pollet (2009), Brown et al. (2010), Lacetera et al. (2012), Allcott and Wozny (2014), and Taubinsky and Rees-Jones (2018)). However, no studies have investigated this issue in once-in-a-lifetime situations. According to Gabaix (2019), the mean and standard deviation of the attention parameters in these papers are 0.44 and 0.28, respectively. As the attention parameters in our results range from 0.594 to 0.854, the said results suggest that people seem to observe information more carefully than usual, but they cannot be completely attentive, even in once-in-a-lifetime situations.

Our study also provides evidence on how different observable characteristics influence attention parameters. Taubinsky and Rees-Jones (2018) theoretically and experimentally show that the heterogeneity of inattention has substantial welfare consequences in the context of non-salient tax. They also reported how individual characteristics influence attention parameters in the online appendix. Their study showed that individuals with higher attention levels exhibited stronger responses when tax rates increased. Compared with these results, we find that people who have high cognitive ability and live in metropolitan areas pay more attention to information. Since the metropolitan area dummy is a proxy for the amount of information people face, the results suggest that people who are better able to process information and are accustomed to processing more information on a daily basis pay more attention to it when faced with a once-in-a-lifetime event.

Our paper is also related to growing literature that investigates the role of information and beliefs in educational decisions (e.g., Hastings and Weinstein (2008), Jensen (2010), Zafar (2011), Stinebrickner and Stinebrickner (2012) and Stinebrickner and Stinebrickner (2014), Wiswall and Zafar (2015), Papay et al. (2016), Bleemer and Zafar (2018) and Tan (2023)). Similar to our work, Papay et al. (2016) examine the impact of labels on college-going decisions when a more precise score is available. Tan (2023) also examines the impact of college-score labels on labor market outcomes. However, because the purpose of the latter study was to investigate how information influences future educational investment decisions and/or labor market outcomes, the authors did not discuss the mechanism by which labels influence students' decisions. To explain the mechanism, this paper provides possible theories and evidence from the perspective of limited attention. This allows us to decompose the jump in the effects of labels and analyze the factors that influence the attention parameters. In this way, we can identify the types of students who require more guidance in interpreting admissions information.

The remainder of this paper is organized as follows: Section 2 provides an institutional background, focusing on the university entrance examination system and student behavior. Section 3 describes the data used in this analysis. Section 4 introduces the RDD and presents the main findings. Section 5 describes the development of the limited attention model, while Section 6 discusses the estimation of the key parameters. Section 7 presents the results of the structural estimation, following which Section 8 concludes. Additional materials and details are provided in the appendix.

2 Institutional Background

2.1 University Entrance Examination System in Japan

In Japan, university entrance examinations are held every year in the fall and winter, and admitted students enroll in schools when the school year starts in April. In the fall, entrance examinations called "recommendation examinations" or "AO examinations" are held. In these examinations, the evaluation is primarily based on interviews, essays, and high-school questionnaires, and there are few or no academic tests. In winter, an examination termed the general entrance examination, focused on academic tests, is held. This study focuses on general entrance exams.

There are two types of universities in Japan: public (national universities and public universities established by local governments) and private. Worldwide, famous research universities, such as the University of Tokyo and Kyoto University, are national institutions. In each region of Japan, public universities are often considered the most highly ranked universities. Additionally, the tuition fees of public universities are low, making them the most popular choice for talented students. Although there are no restrictions on taking entrance exams at private universities, there are strong restrictions on taking entrance exams at public universities. Because of the institutional constraints explained here, this study focuses on application behavior for public universities.

The entrance exams at public universities are divided into two stages. The first stage is the National Center Test for University Admissions[2] (henceforth the Center Test), which is the same for all students. The second stage comprises individual examinations at each university. The second stage is further divided into two rounds (early and later), but students are only allowed to apply to one school in each round. Furthermore, if a student completes the acceptance procedures for the early round (examinations are held at the end of February), he or she will not be eligible for the later round (examinations are held in mid-March), regardless of which public university he or she applies to. As a result, students must make careful and strategic choices regarding where and when to apply. Because a substantial majority of public

[2] The Center Test was renamed to the Common Test for University Admissions in 2011.

university seats are allocated to the early round, and because students' first-round choices are highly consequential, our analysis focuses on application behavior in the early round.

The Center Test is based on a mark-sheet system. The required subjects vary depending on the university or department to which the student applies. Typically, students must answer four to seven questions related to English, mathematics, Japanese, science, and social studies. Again, the required subjects vary depending on the university and faculty. English and Japanese are often required in humanities and social sciences departments, whereas English, mathematics, and science are often required in science departments. In some cases, there is a limit on the number of applicants who can take the second examination, in which case applicants are selected based on their performance in the first examination.

Students can only apply to universities after completing the first stage of the examination. The first-stage examinations are mark-based, and the answers are released the next day so that the students can keep track of their scores. Each student decides which university to apply for based on their score in the first-stage exam. If their performance in the first stage is worse (or better) than expected, there is a possibility that they may change the university they are applying to from their original assumption.

In Japan, students commonly decide which faculty they want to belong to at the application stage. Therefore, students must decide which faculty and university to apply for after the Center Test. In this paper, the university/faculty pairs that will be the options for future applications are called "degree programs."

Most students who take the exam are in their third (and final) year of high school, and account for approximately 90% of all students who take the Center Test. Many of the graduates are so-called "ronin" students who did not pass the university entrance exam of the previous year and decided to take the exam again. It is somewhat common for students who wish to enter high-level universities to take the entrance exam once as a ronin and then again; however, it is rare for students to take it more than once. This is probably because studying for university entrance exams involves substantial direct (generally studying at a preparatory school) and opportunity costs.

2.2 Information System by Major Prep School

An important factor in the development of a student's application strategy is his or her relative position among the students who plan to apply for his or her desired degree programs. Although the mean and variance of test scores in the first stage (the Center Test) are published before the application process, national statistics are not sufficiently informative for strategy development, because detailed relative information is limited.

To overcome this limitation, students commonly rely on information systems provided free of charge by major prep schools ("yobiko" in Japanese), which are specialized private educational institutions for university entrance exams. After the first stage of exams, these major prep schools collect information directly from students, including their self-assessed scores and a ranked list of the degree programs that they are considering applying to. The degree program ranked that is highest on the list is designated as the student's "first choice." In this paper, the primary outcome of interest is whether students actually apply to their declared "first choice" degree program.

Using the collected data, major prep schools tabulate detailed distributions of applicants' scores for each specific degree program. Prep schools set cutoff thresholds by combining historical score distributions with the first-stage examination results of the current year, dividing applicants into discrete admission-likelihood categories. Each student receives a personalized "student result sheet" from the prep school, which contains two types of information: (1) coarse, categorical labels indicating the admission likelihood (explained below) and (2) detailed numerical information, including the distribution of applicants' scores and the exact cutoff thresholds specific to the degree program of their choice. This exchange of information between students and prep schools is widespread and trusted by Japanese students.

The admission probabilities provided in the student result sheets are indicated by discrete categorical labels. The categorical label assigned to each student is determined solely by comparing the students' first-stage examination scores with the cutoff thresholds calculated by the prep schools. Importantly, prep schools do not disclose explicit numerical probabilities of admission. Instead, this system provides both coarse (categorical) information regarding admissions probabilities and detailed numeric information (thresholds and distributions) that enables students to infer more precise admissions probabilities through careful observation.

3 Data

We use the data provided by one of the major prep schools described earlier, covering the years 2011 and 2012. The dataset consists of three main components.

First, it includes the information that students submitted to the prep school and received from it immediately after taking the Center Test, as explained in the previous section. This dataset contains the following information: students' self-assessed scores on the Center Test, their first-choice university at the time, scores recalculated based on the subject weighting for each university, and score labels (from A to E).

Second, it contains the results of a follow-up survey conducted by the prep school after the university entrance examination period, which collected information on the universities that students applied to and their admission outcomes.

Third, it includes selected personal information maintained by the prep school that was made available to us for research purposes. To avoid identifying individuals, only limited personal information is included. The available data include gender (male or female), enrollment status (high school students or graduates), the type of high school attended (public or private), and the prefecture in which the high school is located.

In this study, we focus on students who, after taking the Center Test, selected a public university in the early round (zenki) admission process as their first-choice university and who actually applied to a public university in the first round. We exclude students for whom no such application was confirmed in the follow-up survey.

The summary statistics of the data used in this study are presented in Table 1 1. The sample comprised 375,626 individuals, with 58.34% being male and 41.66% being female. This gender imbalance reflects the fact that the university enrollment rate is higher among males than among females in Japan[3]. Regarding enrollment status, 89.31% are high-school

[3]According to the 2011 Basic School Survey by the Ministry of Education, Culture, Sports, Science and

students, who are generally around 18 years old, whereas 10.69% are high school graduates, typically around 19 or 20 years old.

As for regional distribution, 37.87% of the sample reside in metropolitan areas, defined as prefectures within Japan's three major metropolitan regions: Southern Kanto, Chubu, and Kinki [4]. The remaining 62.13% reside outside these areas. Notably, major preparatory schools are heavily concentrated in these metropolitan regions and provide students with greater access to educational resources.

The score label refers to the label assigned to the applicant's first-choice degree program. Label A accounts for 9.32% of the sample; label B, 11.33%; label C, 16.16%; label D, 17.16%; and label E, the largest group, represents 46.01% of the sample.

Concerning application behavior, 46.02% of individuals changed their application from their initial choice, whereas 53.98% did not. The admission outcomes indicated that 41.11% were admitted to national or public universities through a first-round (zenki) entrance examination. Note that some individuals who were not admitted in the first round may have gained admission through a second-round (kouki) entrance examination or from private universities.

In this study, we use the Center Test score as a proxy for cognitive ability. Specifically, we define the "Three-subject Center Test score" as the total of the scores in English, Mathematics, and Japanese—three core subjects—each scaled to 100 points. The average score is 208.33, with a standard deviation of 40.90, reflecting variation in academic performance within the sample.

The main explanatory variable (running variable in regression discontinuity design) in this study is the score on the Center Test. Each degree program differs in terms of which subjects on the Center Test are used and in what proportion. A typical example of a humanities and social sciences program is the use of a 900-point scale comprising 200 points each for English, 200 points for mathematics, 200 points for Japanese, 200 points for social studies, and 100 points for science.

Technology, the university enrollment rate for four-year universities was 50.6% for males and 44.6% for females.

[4] Saitama, Chiba, Tokyo, and Kanagawa (Southern Kanto); Gifu, Shizuoka, Aichi, and Mie (Chubu); and Shiga, Kyoto, Osaka, Hyogo, Nara, and Wakayama (Kinki).

As an indicator of a student's behavior after receiving information such as score labels, we use the indicator variable of whether the student actually applied to the "first-choice" degree program he or she wanted at the time he or she submitted the information. If a student feels that his or her test score is sufficient to pass the exam, he or she will generally apply to the degree program that he or she had hoped for at the time of information submission (immediately after the center exam). However, if he or she feels that his or her grades are insufficient, he or she will consider changing his or her degree program. This can be achieved by changing to a lower-ranked university in the same department or by changing to a department in the same university that is easier to pass through. We test whether this behavior changed significantly with the threshold of the score label using methods such as RDD.

4 Regression Discontinuity and Finding

4.1 Regression Discontinuity Design

First, we use a sharp RDD to determine whether student behavior changed around the score-label threshold. We define $\boldsymbol{\phi}_i'\mathbf{x}_n$ as student n's score in their "first-choice'' degree program i, where the parameter $\mathbf{x}_n \in [\underline{x}, \overline{x}]^J$ is the vector of scores at center exams for each subject, J is the number of subjects, and $\boldsymbol{\phi}_i$ is the vector of weight for each subject in the ith degree program sets. Let x_i^L be the threshold of the score label $L \in \{A, B, C, D\}$, which varies with degree program i. To apply the regression discontinuity design, we use this threshold to standardize the students' scores.

$$\hat{x}_{i,n}^L = \boldsymbol{\phi}_i'\mathbf{x}_n - x_i^L. \tag{1}$$

$\hat{x}_{i,n}^L$ represents the deviation of the score from the score-label threshold. We use $\hat{x}_{i,n}^L$ as a running variable in our RDD design.

For our statistical tests, we utilize Calonico et al. (2014) to obtain bias-corrected point estimates by employing local linear functions, optimal bandwidths, and valid confidence intervals.

We estimate the discontinuity parameter $\hat{\beta}_{RDD}$ using the following model:

$$\beta_{RDD} = \mu_{+} - \mu_{-}, \tag{2}$$

where

$$\mu_{+} = \lim_{\hat{x} \to 0^{+}} \mu(\hat{x}), \mu_{-} = \lim_{\hat{x} \to 0^{-}} \mu(\hat{x}), \mu(\hat{x}) \equiv E[Y_i | X_i = \hat{x}]. \tag{3}$$

The running variable $\hat{x}$ means $\hat{x}_{i,n}^{L}$ as defined by Equation (1), and the cutoff threshold is normalized to zero. The outcome variable of interest is the indicator variable of whether students changed their application to a degree program that was not their first choice at the time they submitted the information.

We utilize the `rdrobust` package in R, as described in Calonico et al. (2015), for parameter estimation and graphical explanation. The optimal bandwidth for the estimation is selected in a data-driven manner following Calonico et al. (2014), which balanced the bias and variance to achieve a robust inference. We parameterize $\mu(x)$ by using a local linear function. Robustness checks using local quadratic functions yielded similar results, as presented in Table A1 in the Appendix.

4.2 Main Findings of RDD

Figure 1 shows the results of the regression discontinuity design analysis around the four cutoff points (panels A-D). Each panel corresponds to a specific cutoff, where the x-axis represents the score relative to the cutoff (with the cutoff set at 0) and the y-axis indicates the probability of change from the first-choice degree program. The solid red lines show the estimated local linear regression results fitted separately on either side of the cutoff. The dots represent the binned averages of the observed data at evenly spaced intervals.

These panels show that around the cutoff lines B, C, and D, the probability exhibits slight negative jumps, suggesting potential discontinuities in these regions. Contrastingly, around the A cutoff line, there is no evidence of a jump in probability.

Table 1 summarizes the estimates from the regression discontinuity design analysis for the

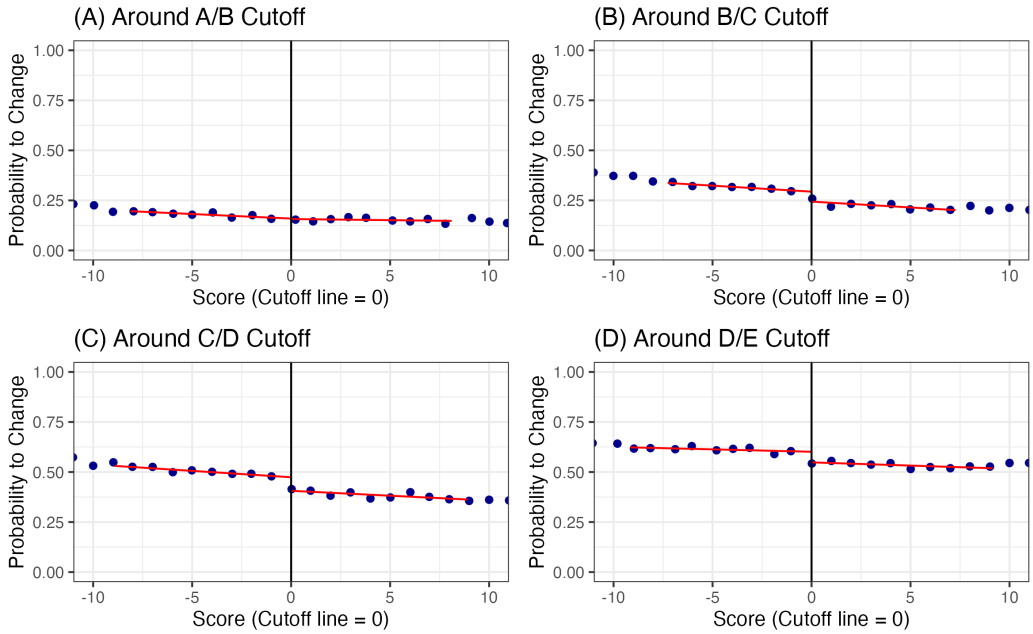

Figure 1: Probability of Changing from First-Choice University and Relative Score to Cutoff Line

Notes: The figures illustrate the results of a regression discontinuity design analysis. The outcome variable is the indicator variable of whether the student changed their application to a degree program that was not their first choice at the time they submitted the information. The running variable is the deviation of the score from the threshold, $\hat{x}_{i,n}^{L}$. The solid lines represent local linear regression estimates, and the dots represent binned averages of the outcome variable.

four cutoff points (A–D). The table presents the estimated discontinuity parameters, standard errors, 95% confidence intervals, sample sizes on the left and right sides of the cutoff, and the bandwidths used in the local linear regression.

The discontinuity parameters have a certain magnitude and are statistically significant for thresholds B, C, and D. Specifically, around the cutoffs, they indicate that the probability of applying to the first-choice university and program decreases by 4.3%–6.7% due to the change in score label. Contrastingly, for threshold A, the discontinuity parameters are small and statistically insignificant.

The results of the regression discontinuity design estimation indicate that student behavior changes around the score thresholds. This finding suggests that some students make decisions based on coarse, rather than detailed, information. The latter sections of this paper develop a decision-making model based on the theory of limited attention and analyze the mechanism of jumps around thresholds.

4.3 Tests for the RDD analyses

A key assumption in the regression discontinuity design is that individuals cannot manipulate their position around the threshold. Because applicants are unaware of the exact threshold at the time of submission, the risk of manipulation is minimized. Nevertheless, statistical tests are conducted to assess the validity of the RDD by detecting any potential manipulation.

First, we perform a density test to detect any discontinuity in the density of the running variables at the threshold. Second, we estimate the regression discontinuity specification for a set of observed baseline characteristics to examine whether any predetermined covariates exhibited discontinuous changes at the threshold.

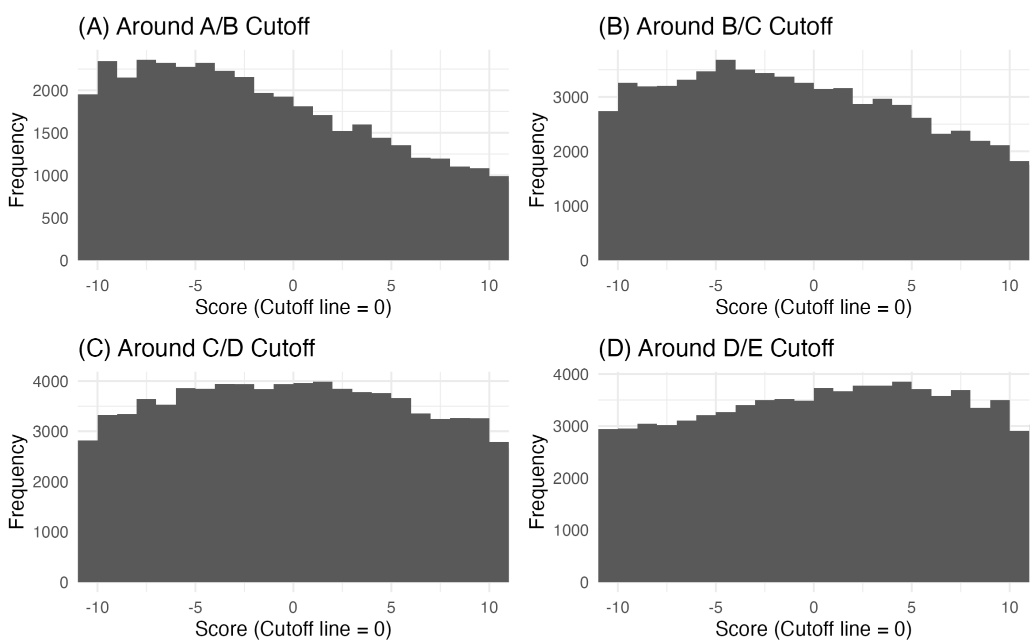

Figure 2: Histograms of Frequency around Cutoff Line
Notes: This figure presents histograms of score distributions around different cutoff points. The x-axis represents the score relative to each cutoff (normalized so that the cutoff line is at 0), while the y-axis shows the frequency of observations.

Figure 2 shows histograms of the frequency around the four different cutoff points (panels A-D). Across all cutoff points, the frequency remained nearly unchanged before and after the cutoff. The results of the manipulation test proposed by McCrary (2008) and Cattaneo et al. (2020) indicate that all p-values are large, suggesting no significant difference in frequency across the cutoff. Specifically, the p-values for cutoffs A/B, B/C, C/D, and D/E are $p = 0.766$, $p = 0.260$, $p = 0.796$, and $p = 0.419$, respectively.

To assess whether predetermined covariates exhibited discontinuous changes at the threshold, we estimate the discontinuity parameters for key baseline characteristics: male dummy, graduate dummy, and metropolitan dummy. The results presented in Figure A1 in the Appendix indicate almost no significant discontinuities at the cutoff. The exception is the metropolitan dummy at the C cutoff, but the estimated discontinuity parameter is small. These findings suggest that individuals just above and below the threshold are comparable in terms of observable characteristics, thus supporting the validity of the regression discontinuity design.

4.4 Sub-sample analyses

In this subsection, we detail a sub-sample analysis which we conduct to gain a more detailed understanding of the mechanism driving discontinuity around the threshold. Specifically, we divide the data by gender (male or female), Three-subject Center Test scores (above or below the in-sample median), and region (metropolitan or non-metropolitan).

At the A/B, C/D, and D/E cutoffs, no substantial differences are observed across sub-samples (Table A3 in the Appendix). Therefore, this discussion focuses on the B/C cutoff. Table 3 presents the results of the sub-sample analysis for the B/C cutoff. The estimates indicate no substantial gender differences, as the discontinuities for male and female applicants are similar in magnitude, with overlapping confidence intervals.

However, differences emerged when examining sub-samples based on the Three-subject Center Test score and regional background. Applicants with low Center Test scores—defined as scores below the in-sample median of the three-subject Center Test—exhibited a more pronounced threshold effect than those with higher scores, defined as scores above the median. Similarly, regional differences are evident, with nonmetropolitan applicants experiencing larger discontinuities than metropolitan applicants. These findings suggest that the threshold effect is more pronounced among lower-scoring or non-metropolitan applicants.

4.5 RDD for admission probability

The results of the regression discontinuity design analyses indicated that the application behavior changed around the cutoff. This raises the question of whether such changes affect the probability of admission to applied universities. In this subsection, we examine whether the probability of admission changes discontinuously at the threshold.

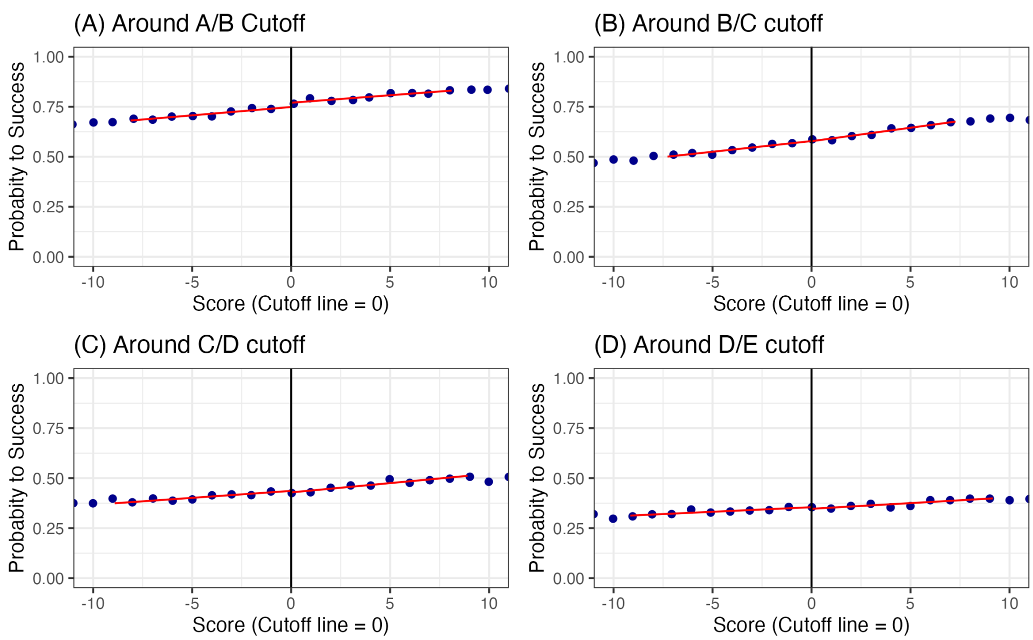

Figure 3: Probability of Success and Relative Score to Cutoff Line
Notes: The figures illustrate the results of a RDD analysis. The outcome variable is an indicator of the student's success in gaining admission to the public university they applied to. The running variable is the deviation of the sore from the threshold. The solid lines represent local linear regression estimates, and the dots represent binned averages of the outcome variable.

Figure 3 applies the same regression discontinuity design as in the previous analyses, but with the outcome variable defined as whether the student was admitted. As illustrated in the figure, no discontinuity is observed at any of the cutoffs. Table 4 presents the statistical estimates for the regression discontinuity design. For all cutoffs, the null hypothesis that the discontinuity is zero cannot be rejected. These results suggest that, although some individuals changed their applied university, there was no difference in the probability of admission around the cutoff.

The regression discontinuity design described above examines whether there is a change in the probability of admission around the cutoff. However, this approach does not directly capture the causal effect of changes in application behavior from the first choice on the probability of admission. To estimate this causal effect, a fuzzy RDD is implemented using the changes in the label as an instrument for the changes in application behavior from the first choice.

Table 5 lists the results of the fuzzy regression discontinuity design. For the A/B cutoff, the standard errors are large, resulting in imprecise estimates. This is because of the absence of a discontinuity at the A/B cutoff. For the other cutoffs, although the coefficients are positive, the standard errors are very large, and the estimates are not statistically significant. These results suggest it is not possible to reject the hypothesis that a change in application behavior induced by discontinuity does not affect the probability of admission. If it is reasonable to assume that admission to a student's first-choice degree program yields higher utility, then these results suggest that changes in application behavior reduce the expected utility.

5 The Model of Limited Attention

The previous section indicated that the probability of a student changing their degree program of choice varies greatly, depending on the student's label, which is received from the prep school. In the current section, we detail the construction of a model that accounts for this fact, based on the assumption that students have a limited attention capacity. The model serves to decompose β into a measure of the sensitivity of the subjective probability to label changes, and a measure of the sensitivity of behavior to changes in the objective probability.

Assume that i is a student n's degree program prior to receiving information from a preparatory degree program. We model the subjective probability of student success in enrollment i using

$$P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\omega\left(\mathbf{B}_{i,n}\right),$$

where the parameter $\mathbf{x}_n \in [\underline{x}, \overline{x}]^J$ is the vector of scores at center exams for each subject and $\boldsymbol{\phi}_i$ is the vector of weight for each subject that the ith degree program sets. The probability $P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right) \in [0,1]$ is the subjective average probability of a student with a center exam score of $\mathbf{x}_n$ being accepted to the degree program i after receiving information from the prep school. The parameter, $\omega\left(\mathbf{B}_{i,n}\right) \in \left[0, \frac{1}{P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)}\right]$ is a subjective weight attached by the student n who selects the ith degree program as their first choice. We assume the weight is determined by the characteristics of the students n and the ith degree program, $\mathbf{B}_{i,n}$. Note that i is a function of n, as it represents the first-choice degree program selected by the nth student. However, for ease of reading, the notation is omitted when there is no risk of confusion.

After receiving the information, we assume that the student can recognize the objective probability of success, $P\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)$ if they pay thorough attention to the information. Because the attention of the student is limited, their understanding of $P\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)$ is $P^a\left(\boldsymbol{\phi}_i\mathbf{x}_n\right)$ where

$$P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right) = G\left(P\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\right).$$

We call G an attention function, which is specified below.

Using this G function, the probability of the student n not applying to the ith degree program after receiving information from the prep school given that they initially choose the ith degree program is

$$\begin{aligned}
& 1 - q\left(\boldsymbol{\phi}_i'\mathbf{x}_n, \mathbf{B}_{i,n}\right) \\
= \; & \Pr\left[\begin{array}{c} P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\omega\left(\mathbf{B}_{i,n}\right)\left[U\left(\mathbf{B}_{i,n}\right)v_{i,n} + F(n)\right] + \left(1 - P^a\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\omega\left(\mathbf{B}_{i,n}\right)\right)F(n) \\ \leq V\left(\mathbf{B}_{i,n}\right)|\boldsymbol{\phi}_i'\mathbf{x}_n, \mathbf{B}_{i,n} \end{array}\right], \\
= \; & \Pr\left[G\left(P\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\right)\omega\left(\mathbf{B}_{i,n}\right)U\left(\mathbf{B}_{i,n}\right)v_{i,n} \leq V\left(\mathbf{B}_{i,n}\right)|\boldsymbol{\phi}_i'\mathbf{x}_n, \mathbf{B}_{i,n}\right],
\end{aligned}$$

where $F(n)$ is the value of failure, $U(\mathbf{B}_{i,n})$ is the additional value by attending the degree program i for student n, $V(\mathbf{B}_{i,n})$ is the opportunity costs of student n's choosing the ith degree program and $v_{i,n}$ is the taste shocks. Assume that $\ln v_{i,n}$ has a continuous distribution function Γ

$$\begin{aligned} 1-q\left(\boldsymbol{\phi}_i'\mathbf{x}_n,\mathbf{B}_{i,n}\right) &= \Pr\left[v_{i,n} \leq \frac{V(\mathbf{B}_{i,n})}{G\left(P\left(\boldsymbol{\phi}_i'\mathbf{x}_n\right)\right)\omega(\mathbf{B}_{i,n})U(\mathbf{B}_{i,n})} | \boldsymbol{\phi}_i'\mathbf{x}_n,\mathbf{B}_{i,n}\right], \\ &= \Gamma\left(\Delta(\mathbf{B}_{i,n}) - \ln G\left(P\left(\boldsymbol{\phi}_i\mathbf{x}_n\right)\right)\right), \end{aligned}$$

where $\Delta(\mathbf{B}_{i,n}) \equiv \ln V(\mathbf{B}_{i,n}) - \ln U(\mathbf{B}_{i,n}) - \ln\omega(\mathbf{B}_{i,n})$.

Consider a cutoff point of "B" and "C" as an example[5] and assume that the admission probability of the cutoff is θ_B. Define $\hat{x}_{n,i}$ by

$$\begin{aligned} \hat{x}_{i,n} &= \boldsymbol{\phi}_i\mathbf{x}_n - x_i^B, \ \textit{where} \\ \theta_B &= P\left(x_i^B\right). \end{aligned}$$

In this specification, if $\hat{x}_{i,n} \geq 0$, the grade is labeled as "B" and if $\hat{x}_{i,n} < 0$, the grade is labeled as "C". Using x_i^B and $\hat{x}_{i,n}$, we can rewrite the application probability by $1-q\left(\boldsymbol{\phi}_i'\mathbf{x}_n,\mathbf{B}_{i,n}\right) = 1-q\left(x_i^B+\hat{x}_{i,n},\mathbf{B}_{i,n}\right)$. Hence,

$$1-q\left(x_i^B+\hat{x}_{i,n},\mathbf{B}_{i,n}\right) = \Gamma\left(\Delta(\mathbf{B}_{i,n}) - \ln G\left(P\left(x_i^B+\hat{x}_{i,n}\right)\right)\right). \tag{4}$$

Define $1-q^e(\hat{x})$ as the expected probability of applying to a degree program other than the degree program originally planned, given $\hat{x}$:

$$1-q^e(\hat{x}) \equiv \int \Gamma\left[\Delta(\mathbf{B}_{i,n}) - \ln G\left(P\left(x_i^B+\hat{x}\right)\right)\right] dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n}=\hat{x}\right), \tag{5}$$

where $Q\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n}=\hat{x}\right)$ is the conditional probability of $\mathbf{B}_{i,n}$ and x_i^B when $\hat{x}_{i,n}=\hat{x}$.

[5] As an example, we model decision-making between cutoffs B and C, although the same logic can be applied to other cutoffs. In the estimation, parameters are estimated for the intervals between cutoffs B/C and C/D. We do not estimate parameters for the cutoffs A/B and D/E because admission probabilities for students in categories A and E cannot be reliably extrapolated from cutoff probabilities.

Using the function $1-q^e(\hat{x})$, we can define the parameter β using

$$\beta \equiv 1-q^e(0)-\lim_{\varepsilon\uparrow 0}\left[1-q^e(-\varepsilon)\right]. \tag{6}$$

We assume that $Q\left(\mathbf{B}_{i,n}, x_i^B|\hat{x}_{i,n}=\hat{x}\right)$ is continuous in $\hat{x}$. Using equations (4), (5) and (6), we can derive the theoretical counterpart of the parameter β as follows.

$$\beta=\int\begin{bmatrix}\Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right)-\ln G\left(P\left(x_i^B\right)\right)\right]\\ -\Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right)-\ln\lim_{\varepsilon\uparrow 0}G\left(P\left(x_i^B-\varepsilon\right)\right)\right]\end{bmatrix}dQ\left(\mathbf{B}_{i,n}, x_i^B|\hat{x}_{i,n}=0\right). \tag{7}$$

Equation (7) implies that $\beta=0$, if $\lim_{\varepsilon\uparrow 0}G\left(P\left(x_i^B-\varepsilon\right)\right)=G\left(P\left(x_i^B\right)\right)$ and $\beta<0$, if $\lim_{\varepsilon\uparrow 0}G\left(P\left(x_i^B-\varepsilon\right)\right)<G\left(P\left(x_i^B\right)\right)$. Specifically, our empirical finding of negative β can be explained if the attention function G has a positive discontinuous jump at x_i^B.

To justify the discontinuity of function G, we specify the attention function as follows:

$$\begin{aligned}G\left(P\left(x_i^B+\hat{x}\right)\right) &= \left[\theta_C+\left(\theta_B-\theta_C\right)I\left(P\left(x_i^B+\hat{x}\right)\geq\theta_B\right)\right]^{1-m}P\left(x_i^B+\hat{x}\right)^m,\\ &= \left[\theta_C+\left(\theta_B-\theta_C\right)I\left(\hat{x}\geq 0\right)\right]^{1-m}P\left(x_i^B+\hat{x}\right)^m,\end{aligned}$$

where θ_C and θ_B are cutoff threshold values calculated by the prep schools for the grades "C" and "B", where $\theta_B>\theta_C$, while $m\in[0,1]$ is an attention parameter, which is introduced by Gabaix (2014, 2019). When $m=1$, $G\left(P\left(x_i^B+\hat{x}\right)\right)=P\left(x_i^B+\hat{x}\right)$. Students can perfectly recognize the probability reported by the prep school. When $m=0$, $G\left(P\left(x_i^B+\hat{x}\right)\right)=\theta_C+\left(\theta_B-\theta_C\right)I\left(\hat{x}\geq 0\right)$. Students do not process information well and only care about labels "B" and "C." Hence, if their score from the center exam exceeds the cutoff value of x_i^B, and, therefore, $\hat{x}\geq 0$, they consider their probability of success to be θ_B. If their score does not exceed the cutoff value, and, therefore, $\hat{x}<0$, they consider their probability of success to be θ_C.

Applying this attention function to equation (7), we obtain

$$\beta = \int \begin{bmatrix} \Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right) - \ln\theta_B^{1-m} P\left(x_i^B\right)^m\right] \\ -\Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right) - \ln\theta_C^{1-m} P\left(x_i^B\right)^m\right] \end{bmatrix} dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n} = 0\right).$$

Taking the first-order approximation with respect to $\ln\theta$ and evaluating $\ln\theta_B$, we derive the following proposition:

Proposition 1. *The theoretical counterpart of the parameter β is approximately decomposed by (1) the measure of the sensitivity of subjective probability to label changes, I, which is an index showing how much the subjective admission probability changes when the label changes from C to B relative to when the objective admission probability changes, and (2) the measure of the sensitivity of behavior to changes in objective probability, K, which is an index showing how sensitive students change their decision on schools of interest when the objective admission probability changes.*

$$\beta \approx IK, \tag{8}$$

where

$$I \equiv g_\theta \frac{1-m}{m}, \; K \equiv \int \frac{\partial 1 - q\left(x_i^B, \mathbf{B}_{i,n}\right)}{\partial \ln P\left(x_i^B\right)} dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n} = 0\right),$$

$g_\theta \equiv \frac{\theta_B - \theta_C}{\theta_C}$.

The mechanism underlying Proposition 1 becomes more intuitive when recognizing that the indicator $mI = g_\theta(1-m)$ represents an index that captures the extent to which the subjective admission probability changes when the label shifts from C to B, while the indicator $K/m = (1/m)\int \frac{\partial 1 - q\left(x_i^B, \mathbf{B}_{i,n}\right)}{\partial \ln P\left(x_i^B\right)} dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n} = 0\right)$ represents an index that measures the sensitivity of students' application decisions to changes in their subjective admission probability. Accordingly, Proposition 1 can be reinterpreted as decomposing (1) the effect of changing labels on subjective probability and (2) the influence of subjective probability on application behavior.

One difficulty is that we cannot directly observe subjective probabilities in the data, which makes it challenging to identify both mI and K/m. To address this issue, we normalize mI by m—which captures the effect of objective probability on subjective probability—in order to interpret the magnitude in terms of changes in the objective measure. This normalization facilitates the decomposition presented in Proposition 1.

The proposition also shows that parameter I comprises the degree of inattention, $\frac{1-m}{m}$, and the expected admission probability change from label changes, g_θ. Note that, $\theta_B > \theta_C$ and $g_\theta > 0$. Therefore, if $m = 1$, $\beta = 0$ and if $m \in (0,1)$, the sign of β is equal to the sign of $\int \frac{\partial 1-q\left(x_i^B, \mathbf{B}_{i,n}\right)}{\partial \ln P\left(x_i^B\right)} dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n} = 0\right)$. Because the theory predicts $\frac{\partial 1-q\left(x_i^B, \mathbf{B}_{i,n}\right)}{\partial \ln P\left(x_i^B\right)} < 0$, this theory suggests that a negative β can be considered as evidence that students do not thoroughly pay attention to the probability $m \in (0,1)$.

6 Identification and Estimation of I, K and m

Proposition 1 shows that our model allows us to decompose β into a measure of the sensitivity of the subjective probability to label changes, I and a measure of the sensitivity of behavior to changes in the objective probability, K. Additionally, it shows that we can obtain information on the attention parameter m from I. In this section, we discuss the identification and estimation of I, K and m from the data.

Note that the probability of changing the schools of interest is

$$\begin{aligned}
&1 - q\left(x_i^B + \hat{x}_{i,n}, \mathbf{B}_{i,n}\right) \\
&= \Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right) - (1-m)\ln\left[\theta_C + \left(\theta_B - \theta_C\right) I\left(\hat{x}_{i,n} \geq 0\right)\right] - m \ln P\left(x_i^B + \hat{x}_{i,n}\right)\right], \\
&\approx \Gamma\left[\Delta\left(\mathbf{B}_{i,n}\right) - (1-m)\ln\theta_C - \gamma I\left(\hat{x}_{i,n} \geq 0\right) - m \ln P\left(x_i^B + \hat{x}_{i,n}\right)\right], \qquad (9)
\end{aligned}$$

where $\gamma = (1-m) g_\theta$. Note that, if we divide the coefficient of $I\left(\hat{x}_{i,n} \geq 0\right)$, γ, by that of $\ln P\left(x_i^B + \hat{x}_{i,n}\right)$, m, we can identify the measure of the sensitivity of subjective probability to label changes, $I = g_\theta \frac{1-m}{m}$ in Proposition 1. Therefore, using the results of Proposition 1, we obtain the following proposition.

Proposition 2. *Considering any binary estimation of Equation (9), I and K are identified by.*

$$I = \frac{\gamma}{m}, \tag{10}$$

$$K = \int \frac{\partial 1 - q\left(x_i^B, \mathbf{B}_{i,n}\right)}{\partial \ln P\left(x_i^B\right)} dQ\left(\mathbf{B}_{i,n}, x_i^B | \hat{x}_{i,n} = 0\right). \tag{11}$$

In addition, if g_θ is known, we can also identify m from I using

$$m = \frac{1}{1 + \frac{I}{g_\theta}}. \tag{12}$$

Proposition 1 shows that $\beta \approx IE$. Hence, once we identify I and K, we can test the validity of this prediction. Because β can be estimated from the regression discontinuity design, the magnitude of discrete jumps can be credible. If we cannot reject the hypothesis that $\hat{\beta}$ constructed by the estimates of structural estimation is equivalent to $\hat{\beta}$ estimated from the regression discontinuity design, we can be more confident in the predictions of our model.

Assume that Γ has a logistic distribution with the mean 0 and the variance of $\frac{\pi^2\sigma^2}{3}$ and that $\Delta\left(\mathbf{B}_{i,n}\right) = \upsilon'\mathbf{M}_i + \lambda'\mathbf{L}_n$, where $\mathbf{B}_{i,n} = \{\mathbf{M}_i, \mathbf{L}_n\}$, $\mathbf{M}_i$ is the vector of the characteristics of ith degree program, and $\mathbf{L}_n$ is the vector of the individual characteristics of the nth students. Then the probability of changing the first-choice degree program is estimated by a logit model, such that

$$1 - q\left(x_i^B + \hat{x}_{i,n}, \mathbf{B}_{i,n}\right) = \frac{\exp y}{1 + \exp y}, \tag{13}$$

where

$$y \approx \tilde{\mu} + \tilde{\upsilon}'\mathbf{M}_i + \tilde{\lambda}'\mathbf{L}_n + \tilde{\gamma} I\left(\hat{x}_{i,n} \geq 0\right) + \tilde{m} \ln P\left(i | x_i^B + \hat{x}_{i,n}\right),$$

$\tilde{\mu} = -\frac{(1-m)}{\sigma} \ln \theta_C$, and $\tilde{\upsilon} = \frac{\upsilon}{\sigma}$, $\tilde{\lambda} = \frac{\lambda}{\sigma}$, $\tilde{\gamma} = -\frac{(1-m)}{\sigma} g_\theta$ where $\tilde{m} = -\frac{m}{\sigma}$. Using the equation

(13) and Proposition 2, we estimate I and K in the next section by

$$I = \frac{\tilde{\gamma}}{\tilde{m}}, \tag{14}$$

$$K = \frac{\sum_n^N \frac{\partial 1-q\left(x^B_{i(n)}, \mathbf{B}_{i(n),n}\right)}{\partial \ln P\left(x^B_{i(n)}\right)}}{N}. \tag{15}$$

where $i(n)$ is the nth student's initial choice of a degree program.

Proposition 2 also shows that, given the knowledge of g_θ, equation (12) identifies m. An obvious question concerns a reasonable value of g_θ. Proposition 1 shows that g_θ can be identified by $g_\theta = \frac{\theta_B - \theta_C}{\theta_C}$. Using the knowledge of θ_B and θ_C[6] set $g_\theta = 1/2$. We also estimate our model using the C/D interval. For a C/D interval, $g_\theta = \frac{\theta_C - \theta_D}{\theta_D}$. In this case, $g_\theta = 1$. With these assumptions regarding the value of g_θ, we also estimate the attention parameter m from the data in the next section.

The specific variables and numerical values used to estimate Equation (13) are as follows: For $\mathbf{M}_i$, we use dummy variables to indicate each university (not a degree program). For $\mathbf{L}_n$, we use a male dummy, a high school graduate dummy, prefecture dummies indicating the location of the student's high school, a dummy for whether the student's high school is public, and the interaction terms between prefecture dummies and the student's Three-subject Center Test score (=Japanese score + English score + mathematics score). The prep school provides admission probabilities only at the cutoff points and does not indicate the probabilities for other scores. As the objective admission probabilities P used in the estimation, we employ probabilities calculated through linear interpolation[7].

7 Results of Structural Estimation

Tables 6 and 7 present the results of the RDD and structural estimation for various subgroups. This table includes RDD estimates with confidence intervals and the corresponding structural

[6] We cannot disclose this knowledge to maintain the confidentiality of the data source.

[7] For example, if a student's score label is C, using our undisclosed knowledge of θ_B and θ_C, the objective probability is calculated as $P = (\theta_B - \theta_C)(\phi_i \mathbf{x}_n - x_i^C)/(x_i^B - x_i^C) + \theta_C$

estimation results. The analysis is performed with different subgroups, categorized by gender (male or female), the Three-subject Center Test score (sum of English, Mathematics, and Japanese, below or above the median), and the region of high school attended (metropolitan or non-metropolitan).

The results show that all structural estimation results lie within the confidence intervals of the RDD estimates, and most coefficients closely approximate their corresponding RDD estimates. This alignment supports the validity and robustness of the structural estimation results.

Tables 8 and 9 present the estimates of β from the structural estimation along with the corresponding parameters m, I, and K, with their 95% confidence intervals calculated using the delta method. Parameter m represents the attention parameter, I is a measure of the sensitivity of the subjective probability to label changes, and K is a measure of the sensitivity of behavior to changes in the objective probability.

The results for m, I, and K collectively reveal important patterns in students' decision-making behavior. The attention parameter m ranges between 0.594 and 0.854, indicating the presence of inattention in the students' behavior. This suggests that many students rely on coarse information rather than fully processing detailed data. The estimates of I are positive and significantly different from zero, indicating that the subjective probability of success changes when the label changes. Similarly, the estimates of K are negative and significantly different from zero, suggesting that students are more likely to apply to their first-choice universities or programs as the objective probability of success increases.

To further investigate the heterogeneity in the attention parameter, we conduct subgroup analyses based on observable characteristics. Regarding gender, there are no significant differences in the overall parameters. The exception is the parameter I and, therefore, m at the C/D cutoff. In this case, females had significantly larger values than males, although this did not hold in other cases. As such, we conclude that there is little difference between males and females in terms of both inattention and application behavior.

Regarding the three-subject Center Test scores, a significant difference is observed between the high- and low-score groups—that is, those above and below the median—at the B/C cutoff. The absolute value of I is larger for the low Center Test score group, indicating that their behavioral changes around the cutoff threshold are greater. Similarly, the absolute value of K is also larger for the low Center Test score group, suggesting that their behavioral response to changes in the probability of admission is stronger. The attention parameter m is smaller for the low Center Test score group, implying that the degree of attention is lower in this group. Although the differences in parameters at the C/D cutoff are smaller and sometimes not statistically significant, the overall trends are similar to those observed at the B/C cutoff.

One possible interpretation of these results is related to students' cognitive abilities. Three-subject Center Test scores were considered to be associated with cognitive abilities. These results suggest that, the higher the cognitive ability, the more likely the students are to analyze detailed information when making decisions. Conversely, poorer cognitive abilities may lead students to make decisions based solely on coarse information.

Regional differences were evident for both cutoffs. For I, the value is larger for the non-metropolitan group, indicating that differences in the cutoff label lead to greater changes in subjective probability in this group. Consistent with these results, the value of m is considerably larger for the metropolitan group, indicating that the overall attention is greater among metropolitan students. For K, the absolute value is higher for the metropolitan group for both cutoffs. Moreover, the difference is statistically significant at the C/D cutoff, suggesting that they responded more strongly to differences in the probability of admission.

These results indicate that metropolitan students tend to focus more on detailed information, whereas non-metropolitan students tend to rely more on coarse information. One possible interpretation of this finding is the difference in the amount of information available to the students. Generally, metropolitan areas have a higher concentration of universities, high schools, and preparatory schools, making it easier for students to access examination-related information. Consequently, metropolitan students are more likely to understand and utilize detailed information.

8 Conclusion

This study investigates how students interpret and respond to coarse information about their admission probabilities in the context of Japanese university entrance examinations. Using a RDD, we find that students exhibited significant behavioral responses to changes in score labels even when more precise information about their relative academic standing is available. Specifically, we observe a sharp decline in the probability of applying for the first-choice program when the score label shifted from a higher to a lower category.

To better understand the mechanisms underlying this sharp decline, we develop and estimated a model of limited attention-to-application decisions. The model allows us to decompose the observed discontinuous jumps into two distinct components: (1) the sensitivity of subjective probability to label changes, and (2) the sensitivity of application behavior to changes in objective probability.

The estimation results indicate that both components are important. Specifically, we observe changes in subjective probability driven by label shifts, as well as changes in application behavior in response to changes in objective probabilities. We also estimate attention parameters. The estimated attention parameter is approximately 0.7. The results confirm that students do not fully incorporate detailed information when making application choices because of limited attention, although they are relatively more cautious than indicated by previous literature.

Our analysis reveals significant heterogeneity in attention parameters across student groups. The subjective probabilities of students with lower Three-subject Center Test scores and those from non-metropolitan areas relies more on categorical labels than on precise probabilities, likely because of differences in cognitive ability and access to educational resources. Contrastingly, high-scoring and metropolitan students process detailed information more effectively. These findings suggest that policies aimed at improving college application decisions should consider these disparities and provide targeted support to students facing disadvantages, such as limited access to admission information and lower academic achievement.

Although this study demonstrates that coarse information significantly influences students' application behavior, the reasons behind the use of such information remain unclear. Future research should explore why admission information is often presented as categorical labels rather than as continuous probability distributions. Understanding the potential benefits of coarse information—such as ease of interpretation and reduced cognitive burden—could provide insights into its role in decision-making and whether alternative formats might improve outcomes.

Additionally, admissions information is not only processed by students themselves but also by their families and high school advisors, who may influence application decisions. Future studies ought to examine how different stakeholders interpret and utilize this information to identify whether discrepancies exist in their understanding of score labels versus precise probabilities. Investigating who pays attention to which aspects of the information and how this affects final application choices could further enhance our understanding of decision-making in college applications.

Variable	Frequency	Percentage
Total Sample	375626	
Gender		
Male	219124	58.34
Female	156502	41.66
Enrollment status		
HS students	335465	89.31
HS graduates	40161	10.69
Region		
Metropolitan	142261	37.87
Non-metropolitan	233365	62.13
Score label		
A	35025	9.32
B	42573	11.33
C	60718	16.16
D	64472	17.16
E	172838	46.01
Application behavior		
Change from the first choice	172872	46.02
No change	202754	53.98
Admission		
Admitted	154423	41.11
Not admitted	221203	58.89
Varible	Mean	S.D.
Three-subject Center Test score	208.33	40.90

Table 1: Summary Statistics
Notes: This table presents the summary statistics of the data. "HS" represents high school. Metropolitan areas are defined as prefectures within Japan's three major metropolitan regions: Southern Kanto, Chubu, and Kinki. Score label refers to the label assigned to the applicant's first-choice degree program. The admission outcome relates to the students' applied-for public degree program. "Three-subject Center Test score'' is the total of the scores in English, mathematics, and Japanese, each converted to a scale of 100 points.

Cutoff	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	-0.008	0.011	[-0.030, 0.014]	42573	35025	8.109
B/C	-0.051	0.011	[-0.073, -0.028]	60718	42573	7.275
C/D	-0.067	0.010	[-0.087, -0.047]	64472	60718	8.912
D/E	-0.043	0.010	[-0.063, -0.022]	172838	64472	9.157

Table 2: Estimates of Regression Discontinuity Design
Notes: This table presents estimates from a regression discontinuity design analysis based on Calonico et al. (2014). The outcome variable is the indicator variable of whether the student changed their application to a degree program that was not their first choice. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014).

Cutoff	Subsample	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
B/C	Male	-0.045	0.015	[-0.073, -0.016]	37748	26579	7.349
B/C	Female	-0.055	0.017	[-0.089, -0.021]	22970	15994	8.297
B/C	High Center Test score	-0.030	0.016	[-0.062, 0.002]	27826	23513	5.815
B/C	Low Center Test score	-0.075	0.017	[-0.109, -0.042]	32892	19060	8.544
B/C	Metropolitan	-0.027	0.017	[-0.061, 0.007]	24405	16521	7.197
B/C	Non-Metropolitan	-0.072	0.014	[-0.100, -0.044]	36313	26052	8.085

Table 3: Estimates of Regression Discontinuity Design: Sub-sample Analysis
Notes: This table presents estimates from a regression discontinuity design analysis based on Calonico et al. (2014). The outcome variable is the indicator variable of whether the student changed their application to a degree program that was not their first choice. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014). "High Center Test score" and "Low Center Test score" indicate whether a student's three-subject Center Test score is above or below the sample median, respectively.

Cutoff	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	0.016	0.013	[-0.009, 0.041]	42573	35025	8.933
B/C	-0.007	0.012	[-0.030, 0.016]	60718	42573	7.735
C/D	-0.020	0.011	[-0.041, 0.002]	64472	60718	6.764
D/E	-0.007	0.008	[-0.024, 0.009]	172838	64472	14.053

Table 4: Estimates of Regression Discontinuity Design: Probability of Success
Notes: This table presents estimates from a RDD analysis based on Calonico et al. (2014). The outcome variable is an indicator of the student's success in gaining admission to the public university they applied to. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014).

Cutoff	Coef.	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	-0.868	9.423	[-19.337, 17.601]	42573	35025	9.165
B/C	0.167	0.270	[-0.363, 0.697]	60718	42573	6.966
C/D	0.274	0.151	[-0.022, 0.570]	64472	60718	7.700
D/E	0.163	0.169	[-0.168, 0.493]	172838	64472	11.243

Table 5: Fuzzy RD Estimates: Impact of the Discontinuity on Success
Notes: This table reports the estimates from a fuzzy RDD, where the outcome variable is defined as whether the student was admitted, the treatment variable is the changes in application behavior from the first choice and the instrument is the changes in the label. The column labeled Coef. represent the local average treatment effects (LATE) for individuals around the cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014).

Cutoff	Sample	β_{RDD}	95%CI	β_{STR}
B/C	All	-0.051	[-0.073, -0.028]	-0.057
B/C	Male	-0.045	[-0.073, -0.016]	-0.056
B/C	Female	-0.055	[-0.089, -0.021]	-0.058
B/C	High Center Test score	-0.030	[-0.062, 0.002]	-0.041
B/C	Low Center Test score	-0.075	[-0.109, -0.042]	-0.074
B/C	Metropolitan	-0.027	[-0.061, 0.007]	-0.027
B/C	Non-metropolitan	-0.072	[-0.100, -0.044]	-0.076

Table 6: Results of Regression Discontinuity Design and Structural Estimation (B/C cutoff)
Notes: β_{RDD} represents the discontinuity shown in Table 2, and the 95%CI indicates the 95% confidence interval. β_{STR} represents the discontinuity obtained from the estimation of our inattention model. "High Center Test score" and "Low Center Test score" indicate whether a student's three-subject Center Test score is above or below the sample median, respectively.

Cutoff	Sample	β_{RDD}	95%CI	β_{STR}
C/D	All	-0.067	[-0.087, -0.047]	-0.079
C/D	Male	-0.056	[-0.081, -0.031]	-0.073
C/D	Female	-0.086	[-0.118, -0.053]	-0.089
C/D	High Center Test score	-0.069	[-0.104, -0.033]	-0.074
C/D	Low Center Test score	-0.063	[-0.092, -0.034]	-0.080
C/D	Metropolitan	-0.056	[-0.086, -0.027]	-0.057
C/D	Non-metropolitan	-0.077	[-0.103, -0.051]	-0.094

Table 7: Results of Regression Discontinuity Design and Structural Estimation (C/D cutoff)
Notes: β_{RDD} represents the discontinuity shown in Tables 2 and A3, while the 95%CI indicates the 95% confidence interval. β_{STR} represents the discontinuity obtained from the estimation of our inattention model. "High Center Test score" and "Low Center Test score" indicate whether a student's three-subject Center Test score is above or below the sample median, respectively.

Cutoff	Sample	β_{STR}	m	95%CI	I	95%CI	K	95%CI
B/C	All	-0.057	0.726	[0.677, 0.775]	0.189	[0.142, 0.236]	-0.302	[-0.278, -0.326]
B/C	Male	-0.056	0.725	[0.660, 0.789]	0.190	[0.129, 0.251]	-0.295	[-0.264, -0.326]
B/C	Female	-0.058	0.732	[0.656, 0.807]	0.183	[0.113, 0.254]	-0.317	[-0.280, -0.355]
B/C	High Center Test score	-0.041	0.770	[0.694, 0.845]	0.150	[0.086, 0.213]	-0.276	[-0.244, -0.309]
B/C	Low Center Test score	-0.074	0.694	[0.630, 0.757]	0.221	[0.155, 0.287]	-0.334	[-0.298, -0.369]
B/C	Metropolitan	-0.027	0.854	[0.771, 0.936]	0.086	[0.029, 0.143]	-0.315	[-0.280, -0.351]
B/C	Non-metropolitan	-0.076	0.655	[0.594, 0.717]	0.263	[0.191, 0.335]	-0.290	[-0.258, -0.322]

Table 8: Results of Structural Estimation (B/C cutoff)
Notes: β_{STR} (discontinuity), m (attention parameter), I, K are the estimated parameters of our model. The 95% CI represents the 95% confidence interval calculated based on the delta method. "High Center Test score" and "Low Center Test score" indicate whether a student's three-subject Center Test score is above or below the sample median, respectively.

Cutoff	Sample	β_{STR}	m	95%CI	I	95%CI	K	95%CI
C/D	All	-0.079	0.667	[0.619, 0.715]	0.500	[0.392, 0.608]	-0.158	[-0.142, -0.174]
C/D	Male	-0.073	0.694	[0.634, 0.755]	0.440	[0.315, 0.564]	-0.165	[-0.145, -0.186]
C/D	Female	-0.089	0.622	[0.543, 0.701]	0.607	[0.403, 0.811]	-0.146	[-0.120, -0.172]
C/D	High Center Test score	-0.074	0.698	[0.633, 0.762]	0.434	[0.300, 0.567]	-0.172	[-0.148, -0.195]
C/D	Low Center Test score	-0.080	0.644	[0.574, 0.715]	0.552	[0.381, 0.722]	-0.145	[-0.123, -0.168]
C/D	Metropolitan	-0.057	0.766	[0.696, 0.835]	0.306	[0.188, 0.424]	-0.185	[-0.160, -0.210]
C/D	Non-metropolitan	-0.094	0.594	[0.527, 0.660]	0.685	[0.496, 0.873]	-0.137	[-0.116, -0.158]

Table 9: Results of Structural Estimation (C/D cutoff)
Notes: β_{STR} (discontinuity), m (attention parameter), I, K are the estimated parameters of our model. The 95% CI represents the 95% confidence interval calculated based on the delta method. "High Center Test score" and "Low Center Test score" indicate whether a student's three-subject Center Test score is above or below the sample median, respectively.

References

Allcott, Hunt and Nathan Wozny, "Gasoline Prices, Fuel Economy, and the Energy Paradox," *Review of Economics and Statistics*, 2014, *96* (5), 779–795.

Bleemer, Zachary and Basit Zafar, "Intended college attendance: Evidence from an experiment on college returns and costs," *Journal of Public Economics*, 2018, *157*, 184–211.

Brown, Jennifer, Tanjim Hossain, and John Morgan, "Shrouded Attributes and Information Suppression: Evidence From the Field," *Quarterly Journal of Economics*, 2010, *125* (2), 859–876.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs: Robust nonparametric confidence intervals," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , _ , and _ , "rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs," *The R Journal*, 2015, *7* (1).

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

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

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

— **and Joshua M. Pollet**, "Investor Inattention and Friday Earnings Announcements," *Journal of Finance*, 2009, *64* (2), 709–749.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

— , "Behavioral Inattention," *Handbook of Behavioral Economics*, 2019, *2*, 261–343.

Hastings, Justine S. and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence From Two Experiments," *Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Heiss, Florian, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou, "Inattention and Switching Costs as Sources of Inertia in Medicare Part D," *American Economic Review*, 2021, *111* (9), 2737–2781.

Hossain, Tanjim and John Morgan, "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *Advances in Economic Analysis Policy*, 2006, *6* (2), Article 3, 1–27.

Jensen, Robert, "The (Perceived) Returns to Education and the Demand for Schooling Get," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor, "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 2012, *102* (5), 2206–2236.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt, "Rational Inattention: A Review," *Journal of Economic literature*, 2023, *61* (1), 226–273.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Morrison, William and Dmitry Taubinsky, "Rules of Thumb and Attention Elasticities: Evidence From Under- and Overreaction to Taxes," *Review of Economics and Statistics*, 2023, *105* (5), 1110–1127.

Papay, P. John, Richard J. Murnane, and John B. Willett, "The Impact of Test Score Labels on Human-Capital Investment Decisions," *Journal of Human Resources*, 2016, *51* (2), 357–388.

Stinebrickner, Ralph and Todd Stinebrickner, "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *Review of Economic Studies*, 2014, *81*, 426–472.

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

Tan, Brandon Joel, "The Consequences of Letter Grades for Labor Market Outcomes and Student Behavior," *Journal of Labor Economics*, 2023, *41* (3), 565–588.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies*, 2018, *85*, 2462–2496.

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

Zafar, Basit, "How Do College Students Form Expectations?," *Journal of Labor Economics*, 2011, *29* (2), 301–348.

Appendix

Cutoff	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	-0.012	0.013	[-0.037, 0.014]	42573	35025	12.219
B/C	-0.049	0.014	[-0.076, -0.022]	60718	42573	10.775
C/D	-0.066	0.013	[-0.092, -0.039]	64472	60718	11.313
D/E	-0.041	0.011	[-0.062, -0.020]	172838	64472	17.500

Table A1: Estimates of Regression Discontinuity Design Using Local Quadratic Function
Notes: This table presents estimates from a regression discontinuity design analysis based on Calonico et al. (2014). In Table 2, we use a local linear function, whereas in this table we use a local quadratic function. The outcome variable is the indicator variable of whether the student changed their application to a degree program that was not their first choice. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014).

Cutoff	Covariates	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	Male dummy	0.010	0.015	[-0.018, 0.039]	42573	35025	8.365
A/B	HS graduates dummy	0.002	0.011	[-0.019, 0.024]	42573	35025	7.802
A/B	Metropolitan dummy	-0.003	0.013	[-0.029, 0.022]	42573	35025	10.866
B/C	Male dummy	-0.010	0.013	[-0.034, 0.015]	60718	42573	6.703
B/C	HS graduates dummy	-0.015	0.008	[-0.031, 0.002]	60718	42573	9.206
B/C	Metropolitan dummy	-0.013	0.010	[-0.033, 0.007]	60718	42573	10.383
C/D	Male dummy	-0.020	0.011	[-0.041, 0.001]	64472	60718	7.894
C/D	HS graduates dummy	-0.010	0.007	[-0.025, 0.004]	64472	60718	9.255
C/D	Metropolitan dummy	-0.023	0.011	[-0.045, -0.002]	64472	60718	7.406
D/E	Male dummy	-0.008	0.010	[-0.027, 0.012]	172838	64472	10.716
D/E	HS Graduates dummy	-0.009	0.007	[-0.023, 0.004]	172838	64472	9.571
D/E	Metropolitan dummy	-0.002	0.010	[-0.021, 0.017]	172838	64472	11.797

Table A2: Estimates of Regression Discontinuity Design for Covariates
Notes: This table presents estimates from a regression discontinuity design analysis based on Calonico et al. (2014). The outcome variable is noted in the Sub-sample column. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014).

Cutoff	Subsample	$\hat{\beta}_{RDD}$	S.E.	95% CI	N_{left}	N_{right}	Bandwidth
A/B	Male	-0.014	0.016	[-0.045, 0.018]	26579	21098	6.824
A/B	Female	-0.002	0.013	[-0.028, 0.025]	15994	13927	12.703
A/B	High Center Test score	-0.009	0.017	[-0.042, 0.024]	20633	17804	5.888
A/B	Low Center Test score	-0.022	0.017	[-0.054, 0.011]	21940	17221	10.730
A/B	Metropolitan	-0.016	0.018	[-0.050, 0.019]	16521	11704	7.201
A/B	Non-Metropolitan	-0.005	0.013	[-0.031, 0.021]	26052	23321	10.423
B/C	Male	-0.045	0.015	[-0.073, -0.016]	37748	26579	7.349
B/C	Female	-0.055	0.017	[-0.089, -0.021]	22970	15994	8.297
B/C	High Center Test score	-0.030	0.016	[-0.062, 0.002]	27826	23513	5.815
B/C	Low Center Test score	-0.075	0.017	[-0.109, -0.042]	32892	19060	8.544
B/C	Metropolitan	-0.027	0.017	[-0.061, 0.007]	24405	16521	7.197
B/C	Non-Metropolitan	-0.072	0.014	[-0.100, -0.044]	36313	26052	8.085
C/D	Male	-0.056	0.013	[-0.081, -0.031]	39007	37748	9.350
C/D	Female	-0.086	0.017	[-0.118, -0.053]	25465	22970	9.236
C/D	High Center Test score	-0.069	0.018	[-0.104, -0.033]	27946	33955	5.596
C/D	Low Center Test score	-0.063	0.015	[-0.092, -0.034]	36526	26763	10.066
C/D	Metropolitan	-0.056	0.015	[-0.086, -0.027]	25887	24405	9.404
C/D	Non-Metropolitan	-0.077	0.013	[-0.103, -0.051]	38585	36313	9.437
D/E	Male	-0.043	0.013	[-0.069, -0.018]	94692	39007	10.181
D/E	Female	-0.044	0.015	[-0.073, -0.015]	78146	25465	11.375
D/E	High Center Test score	-0.045	0.012	[-0.069, -0.021]	71788	46553	8.868
D/E	Low Center Test score	-0.038	0.017	[-0.072, -0.004]	101050	17919	11.769
D/E	Metropolitan	-0.033	0.015	[-0.062, -0.003]	63744	25887	10.037
D/E	Non-Metropolitan	-0.051	0.012	[-0.076, -0.027]	109094	38585	11.322

Table A3: Estimates of Regression Discontinuity Design: Subsample Analysis
Notes: This table presents estimates from a regression discontinuity design analysis based on Calonico et al. (2014). The outcome variable is the indicator variable of whether the student changed their application to a degree program that was not their first choice. The column labeled $\hat{\beta}_{RDD}$ shows the estimated discontinuity parameters at each cutoff with standard errors (S.E.). The 95% confidence intervals (95% CI) are displayed in brackets. N_{left} and N_{right} denote the number of observations to the left and right of each cutoff, respectively. The Bandwidth column indicates the optimal bandwidth selected using the algorithm of Calonico et al. (2014). “High Center Test score” and “Low Center Test score” indicate whether a student’s three-subject Center Test score is above or below the sample median, respectively.