Dynamic Matching Mechanism and Matching Stability in College Admissions: Evidence from Inner Mongolia

Le Kang* Wei Ha† Haoyuan (Marty) Chen‡ Yuhao Deng §

March 2024

Abstract

We present the first large-scale empirical evidence on the effects of adopting a dynamic matching mechanism, in replacement of the Immediate Acceptance (IA) mechanism, on matching stability in college admissions in China. In 2007, the Inner Mongolia Autonomous Region introduced the "Real-time Dynamic Mechanism", which allowed college applicants to change their college choices as many times as they want during a restricted time interval while seeing their tentative admission outcome when they made each choice. Using administrative data on test scores and admission outcomes of the universe of National College Entrance Exam (NCEE) takers from 2005 to 2011, we construct measures of justified envy, an indicator of matching stability. We use a generalized difference-in-differences framework and, in contradiction to the theoretical and experimental predictions from previous studies, find no evidence that the real-time dynamic mechanism improved matching stability in the first four years after its implementation. Our findings suggest that the real-time dynamic mechanism is much less effective in eliminating justified envy than the parallel mechanism, a hybrid of IA and the Deferred Acceptance (DA) mechanism, which is now widely adopted in other provinces in China.

Keywords: Inner Mongolia, Real-time Dynamic Mechanism, Immediate Acceptance (IA) Mechanism, Matching Stability, Justified Envy

JEL Codes: C78, D82, I23, I28

*Institute of Education, Nanjing University (lekang@nju.edu.cn).
†Institite of Economics of Education, Peking University (wha@pku.edu.cn).
‡Corresponding author. Department of Economics, University of Maryland (hchen915@umd.edu).
§Department of Biostatistics, School of Public Health, University of Michigan (yuhaoden@umich.edu).

1 Introduction

Every year, millions of students apply to colleges through a centralized application and admission system. At the center of such a system is the design of a matching mechanism that produces stable outcomes based on students' grades, preferences of students and colleges, and the capacity of each college. Three of the most well-known matching mechanisms are the immediate acceptance (IA) mechanism, the deferred acceptance (DA) mechanism, and the top trading cycles (TTC) mechanism. Among the three, the student-proposing deferred acceptance mechanism proposed by Gale and Shapley (1962) is often celebrated by the literature as the ideal mechanism since it is strategy-proof and ensures that truth-telling is a dominant strategy for students. However, recent experimental (Chen and Sönmez, 2006; Chen and He, 2021; Pais and Pintér, 2008; Pais, Pintér and Veszteg, 2011) and empirical evidence (Chen and Pereyra, 2019; Hassidim, Romm and Shorrer, 2016; Rees-Jones, 2018; Shorrer and Sóvágó, 2022) comparing DA and other mechanisms suggests that students often do not truthfully report their preferences. This leads to potentially unstable allocations. Moreover, these studies find that information has an important effect on truthful preference revelation.

For students, understanding the matching mechanism is key to a preferred matching. Take college admissions in China as an example, a good strategy in the ranking of schools is often as important, if not more important, than a good score in the college entrance exam. Anecdotal evidence reveals that college applicants often devote a tremendous amount of time to exam preparation but understand little of the college application and matching system by the time of application submission. To provide students with more information, education systems in many countries have introduced dynamic matching mechanisms. [1] One prominent example is the college admission system in the Chinese province of Inner Mongolia. Each year, almost 200,000 Inner Mongolia test takers of the

[1] For example, university admissions in Brazil, Germany, and Inner Mongolia (China) have adopted dynamic (or iterative) mechanisms to match students to colleges.

National College Entrance Exam (NCEE), also known as *Gaokao*, were matched to colleges across the country through a "Real-time Dynamic Mechanism". Under the real-time dynamic mechanism, students can submit and revise their tentative college-major choices as many times as they want during a restricted time interval on the day of application submission. Each time a student submits a college-major choice, her tentative admission outcome for that college-major is computed based on the pre-determined number of seats reserved for the Inner Mongolia province in that year and her relative *Gaokao* score to other students who submitted the same college-major choice. The tentative admission outcome is then made public to the student and is updated continuously accordingly for every other student. By providing real-time feedback to students about their current allocations, the mechanism intends to reduce information uncertainty and increase matching stability.

We present the first empirical evidence on the matching stability of the real-time dynamic matching mechanism. Using administrative data on individual-level *Gaokao* test scores and college admission outcomes for the universe of test takers from 2005 to 2011, we construct measures of justified envy (JE) and compare the JE outcomes under the real-time dynamic mechanism to those under IA. The concept of justified envy was introduced by Abdulkadiroğlu and Sönmez (2003), and since then stability and "no justified envy" were often used synonymously in the matching theory literature. Specifically, we say that *student s_1 justifiably envies student s_2 if student s_1 is matched to college c_1, student s_2 is matched to college c_2, and student s_1 scores higher than student s_2, but prefers c_2 to c_1*. Recent theoretical and experimental studies find that dynamic mechanisms, which allow for interactions and provide extra information, can outperform the static ones (Bó and Hakimov, 2020; Ding and Schotter, 2017; Dur, Hammond and Kesten, 2021; Gong and Liang, 2020; Grenet, He and Kübler, 2022; Klijn, Pais and Vorsatz, 2019; Stephenson, 2022). In particular, they have shown that dynamic mechanisms can lead to higher truth-telling rates, better matching efficiency, and more stable matching outcomes. Our paper fills the gap in empirical evidence by using administrative data to construct measures of matching stability and comparing them under different mechanisms.

Using a generalized difference-in-differences design, we find no evidence that the real-time dynamic mechanism is better than IA in eliminating justified envy. During the first four years after its implementation, the real-time dynamic mechanism did not decrease the proportion of applicants that justifiably envy at least one other applicant. It also had no impact on the average number of applicants justifiably envied by one. In addition to the traditional JE definition, we construct additional measures to capture the envy in both school quality and student test-score dimensions. These measures are similar to those in Kang et al. (2020) and we adjust them to the context of our setting. We find that the real-time dynamic mechanism led to just as much envy in these two dimensions as the IA, if not more. The results are robust to a variety of different specifications. Using event-study method, we find a significant increase in justified envy in the first year after the Inner Mongolia reform. It then dropped gradually to similar level as in the control provinces in the next few years. Synthetic control estimates yield consistent results. Then, we compare the real-time dynamic mechanism to the widely-adopted parallel mechanism and show that parallel mechanism leads to more stable matching than the real-time dynamic mechanism. Similar results are also found when we use rank range and match index as proxies of matching stability. [2] Finally, we discuss potential reasons for the unexpected results from the Inner Mongolia reform. The observed results could be explained by the immaturity of the mechanism design in its initial years and the "collusion" among applicants.

College admissions in China are unique in several ways. First, it is the largest centralized college admission system in the world, with millions of students being matched to hundreds of colleges every year. Second, the college entrance exams and matching allocations are centralized at the provincial level. In the last two decades, there have been major matching mechanism reforms in nearly all provinces. While IA used to be the most

[2] These alternative measures of matching stability are introduced by Ha, Kang and Song (2020).

prevalent admission mechanism, by 2018, all provinces except Inner Mongolia switched to the parallel mechanism, a hybrid of IA and DA. The timing and geographical variations in the reforms provide a unique opportunity for estimating the effects of matching mechanisms on matching stability outcomes. Third, to construct JE, we need to make assumptions about student preferences. The Chinese higher education system provides a context where we can reasonably assume students share homogeneous preferences over Tier-1 universities. [3] We present results where we assume students have homogeneous preferences over groups of universities categorized based on the average *Gaokao* scores of students admitted to the university.

Our results contribute to two strands of literature. First, we add to the literature on school matching mechanisms. Gale and Shapley in their seminal paper (Gale and Shapley, 1962) introduced the student-proposing deferred acceptance (DA) procedure. Over the years, the strategy-proof DA and its variations have been endorsed by the literature and have been adopted in many real-life student-school matching programs around the world (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu, Pathak and Roth, 2009; Balinski and Sönmez, 1999; Pathak and Sönmez, 2013). Other mechanisms, such as the immediate acceptance (IA) mechanism, the so-called "Boston mechanism", and the parallel mechanism are also used to match millions of students to schools and colleges every year (Chen and Kesten, 2017; Chen and Pereyra, 2019; Ergin and Sönmez, 2006; Ha, Kang and Song, 2020; Kang et al., 2020; Pathak and Sönmez, 2008).

On top of the extensive research on single-offer mechanisms, recent years have seen the emergence of dynamic matching mechanisms that match students to schools and colleges. Prominent examples include college admissions in Brazil, Germany, and the Chinese province of Inner Mongolia. We provide the first large-scale empirical evidence comparing matching stability of the real-time dynamic mechanism to its alternatives. Pre-

[3]We use university and college interchangeably in this paper. Chinese universities are categorized into different tiers. In our sample, there are Tier-1, Tier-2, Tier-3, and vocational universities, with Tier-1 universities better than Tier-2 universities, Tier-2 universities better than Tier-3 universities, and so on. In recent years, Tier-2 and Tier-3 universities are merged into one.

vious literature on the comparison of dynamic versus static mechanisms highlights the advantages of dynamic mechanisms. Klijn, Pais and Vorsatz (2019) compare dynamic versions of both the student-proposing DA and school-proposing DA to their static counterparts. They find that both dynamic mechanisms outperform the static versions in truth-telling preference revelation and matching stability. Bó and Hakimov (2020) and Bó and Hakimov (2022) find that a much higher proportion of stable outcomes is reached under iterative DA than under DA. They propose the "iterated deferred acceptance mechanism" to yield student-optimal stable matching. Grenet, He and Kübler (2022) study the mechanism in Germany's university admissions. Using both a theoretical model and survey evidence, they show that a dynamic multi-offer mechanism can improve matching efficiency. Our paper adds to this literature on the comparison of dynamic and static matching mechanisms.

Second, the results of our paper contribute to the literature on the college matching mechanisms in China. While IA used to be the dominant matching mechanism in college admissions, all provinces except Inner Mongolia have switched to the "Parallel mechanism" during the past twenty years. Chen and Kesten (2017) characterize the parallel mechanism within a system of matching mechanisms as in between IA and DA, and show that the parallel mechanism is more stable than IA. Further experimental and empirical research support this result (Chen and Kesten, 2019; Chen, Jiang and Kesten, 2020; Ha, Kang and Song, 2020; Kang et al., 2020). Other studies have explored the change in college preference submission timing from pre-exam to post-exam and from pre-score posting to post-score posting (Lien, Zheng and Zhong, 2016, 2017; Wu and Zhong, 2014). A few studies look at the matching mechanism in Inner Mongolia. Gong and Liang (2020) find, both theoretically and experimentally, that the Inner Mongolia dynamic mechanism is as stable as DA, is as efficient as IA, and induces a higher truth-telling rate than both DA and IA. Chen, Pereyra and Zhu (2022) in their paper referred to the Inner Mongolia mechanism as the "time-constrained dynamic mechanism", and show that the time-constrained feature of this mechanism could lead to worse outcomes than DA. They also use university-level application data published by the clearinghouse to show that dynamic mechanisms may achieve unstable outcomes and create winners and losers among students. However, their sample is small, preventing them from performing extensive empirical comparisons between different mechanisms. The results from our paper complement those presented in these studies by presenting large-scale empirical evidence on the matching results in Inner Mongolia. Our results also have important policy implications. A comparison of Inner Mongolia's real-time dynamic mechanism and the parallel mechanism adopted in the rest of the country shows that the parallel mechanism produces much more stable matching outcomes. An obvious solution to improve matching stability would be for Inner Mongolia to switch to the parallel mechanism.

This paper proceeds as follows. Section 2 describes college admissions in China and the real-time dynamic mechanism in Inner Mongolia. Sections 3 and 4 outline the data and the empirical strategy, respectively. Section 5 presents the effects of the Inner Mongolia Reform. Section 6 compares the real-time dynamic mechanism to the parallel mechanism. Section 7 provides a discussion of the real-time dynamic mechanism and potential reasons for why it fails to yield a more stable matching. Section 8 concludes.

2 Inner Mongolia's Real-time Dynamic Mechanism

2.1 College Admissions in China

College admissions in China are centralized, featuring a provincial-based admission quota system. The admission results are almost entirely based on the test scores from the National College Entrance Exam (NCEE), commonly known as the *Gaokao*. Students choose either STEM or non-STEM track at the end of 10th grade, study different subjects in the final two years of high school, and take the province-track-specific NCEE at the end of 12th grade. Each year, the Ministry of Education coordinates with the National Development and Reform Commission to determine how many students each college can admit from the next *Gaokao*. The only major decision each university needs to make is to allocate its capacities (i.e. admission quotas) among provinces. While the exam takes place on June 7th and 8th nationwide, the exam questions are not necessarily the same across different provinces. [4] Each student registers for *Gaokao* in the province where her household registration (*hukou*) is located, and competes with other students who take the exam in the same province.

Education authorities in each province are responsible for matching students from their own area to universities all over the country, given the provincial-level quota set by the universities, each student's *Gaokao* scores, and student preferences. Before 2003, all provinces used the immediate acceptance (IA) matching mechanism to allocate their students. Alternatives to IA were experimented with and implemented since 2003. To this date, almost all provinces switched to the parallel mechanism, an intermediate mechanism between IA and the deferred acceptance (DA) mechanism (Chen and Kesten, 2017). One notable exception is the province of Inner Mongolia. Starting in 2007, all *Gaokao* test takers in Inner Mongolia applying for Tier-1 universities in China were matched to these universities using the "Real-time Dynamic Mechanism". It was then extended to Inner Mongolia students applying to all Chinese universities in the following years.

2.2 Real-time Dynamic Mechanism

Under the real-time dynamic mechanism, students can submit and revise their tentative college and major choices as many times as they want during a restricted time interval

[4] Prior to 2003, *Gaokao* was held in July. Some provinces design their own exams, while others are divided into groups, such that students in provinces from the same group take the same exam. There are multiple reasons for not having a consistent nationwide exam, including the concerns of equity due to heterogeneous education resources and thus student ability across provinces, and differential preferences of regional education authorities. Since the competition in college admissions is entirely within-province, the differential difficulty of the exams does not raise equity issues.

on the day of college preference submission. [5] Each time a student submits a college-major choice, her tentative admission outcome for that college-major is computed based on the admission quota and her relative *Gaokao* score to other students who submitted the same college-major choice. The tentative admission outcome is then made public to the student and is updated continuously accordingly for every other student. This dynamic mechanism was designed to reduce information uncertainty and improve efficiency and matching stability.

Before the start of the college application (i.e., preference submission) process, each university publishes the admission quota for every major offered by the university. In comparison to IA, on the day of application submission, applicants receive extra information through two main channels. First, as an applicant selects a tentative college-major choice, she is provided with her *Gaokao* ranking among applicants from Inner Mongolia who made the same (tentative) college-major choice, the admission quota for this college-major in Inner Mongolia, and her admission result if nobody makes changes to the college-major choice before the final submission deadline. Second, at every full hour, the current admission *Gaokao* cutoff scores for all college-major pairs are made public to all applicants. The extra information provides applicants with more information about school vacancies and allows them to update their expectations of the admission probability at each school.

The real-time dynamic mechanism is implemented as the following:

- All applicants are partitioned into N groups, $g_1, g_2, ..., g_N$, based on their *Gaokao* test scores. g_1 includes applicants with scores c within $[c_1, \bar{c}]$, g_2 includes applicants with scores c within $[c_2, c_1)$, ..., and g_N includes applicants with scores c within $[\underline{c}, c_{N-1})$. [6]

[5] Some college-major pairs are track-specific, while others are open to students from both the STEM and non-STEM track.

[6] The number of groups and the group cutoffs changed from year to year. For example, in 2007, there was only one group, including all Tier-1 university applicants. In 2008, there were four groups. Since 2021, test takers were divided into 10 groups: applicants with test scores 670 to 750, 640 to 669, 610 to 639, ...,430 to 459, 429 or below.

- On the day of college preference submission, all applicants enter the online application system at the same time. They can submit one college-major choice at a time, but can change their choices as many times as they want before the submission deadline.
- When an applicant submits a choice, she can observe its admission quota and her real-time ranking among all applicants that submitted the same choice. All applicants can also observe the admission cutoff scores for all college-major choices at every full hour.
- Each group g_i has a submission deadline T_i, with $T_1 < T_2 < ... < T_N$, i.e. applicants in a group with higher scores would have an earlier deadline than applicants in groups with lower scores. [7] At each submission deadline, each college-major admits students with the highest scores who selected them as their final choice, up to its admission quota.
- All applicants who are left unmatched to a school will enter a second round of matching after the submission deadline for the last group. [8]

We name the Inner Mongolia mechanism as the "real-time dynamic mechanism" because of its dynamic aspect, i.e. students interact with others dynamically before forming matches, and its real-time feedback feature. As we have discussed in the introduction section, previous theoretical and experimental studies looking at other dynamic matching mechanisms that are similar to the real-time dynamic mechanism have shown that dynamic mechanisms would outperform their static versions in matching stability, matching efficiency, and truth-telling preference revelation. Since essentially the only difference between the real-time dynamic mechanism and IA, its predecessor in Inner Mongolia, is that college applicants are allowed to observe their tentative admission outcomes and freely

[7] In recent years, each group had one extra hour to adjust than the group with the deadline just before it. Students are not obliged to submit a college-major choice by any time before the submission deadline for their group, but recent statistics show that over 90 percent of students put in a college-major choice during the first hour after the portal opens. See https://www.nm.zsks.cn/kszs/ptgk/ggl/202207/t20220720_43209.html, retrieved on 6/2/2023.

[8] In recent years, over 95 percent of students were matched to their selected programs in the first round.

adjust their college application choices based on this extra information, IA can be considered as the real-time dynamic mechanism's static counterpart. Based on established evidence, one would predict that the real-time dynamic mechanism leads to more stable allocations than IA. This paper compares the real-time dynamic mechanism to IA using national-level data and provides evidence against the theoretical and lab predictions.

3 Data

We use a unique administrative data set of college admission records between 2005 and 2011. It contains individual-level information on the universe of *Gaokao* test takers who were admitted to any college, including the year of the test, province, exam track (STEM or non-STEM), test scores, and the admitted university and major. Since Inner Mongolia reformed its matching mechanism in 2007, the data allow us to make causal inferences on the effect of the real-time dynamic mechanism by comparing admission results in Inner Mongolia and those from other provinces which remained using IA during the entire time period.

3.1 Justified Envy

The main outcome variable is a measure of justified envy (JE), a key indicator of matching stability. Consider a college matching problem, with n students, denoted by $s_1, s_2, ..., s_n$, and m colleges, denoted by $c_1, c_2, ..., c_m$. Each student has preferences over the entire set of colleges plus remaining unmatched. Colleges rank all students based on their *Gaokao* scores and prefer students with higher scores, given the pre-determined admission quota. A matching $M : S \rightarrow C$ is a list of assignments that assign each student to a school, including the null assignment, with each school admitting no more than its admission quota. Now, given a matching M, we say that *student s_i justifiably envies student s_j if $M(s_i) = c_i$, $M(s_j) = c_j$, and student s_i scores higher in Gaokao than student s_j, but prefers c_j to c_i.*

Figure 1 illustrates the JE measures. Consider a matching problem with four colleges, denoted by c_1, c_2, c_3, c_4, and eight applicants, denoted by letters A to H. Assume that schools share homogeneous preferences and rank students by their *Gaokao* scores, and students also share homogeneous preferences and prefer college c_1 to c_2, c_2 to c_3, and c_3 to c_4. Figure 1 plots one possible matching outcome with student-rank on the horizontal axis (1 being the most preferred), and university-tier on the vertical axis (1 being the most preferred). For example, student A, represented by $(1, 2)$, has the highest score among the eight applicants and is matched to the second preferred school; student E, represented by $(5, 1)$, is ranked fifth in *Gaokao* scores and is matched to the most preferred school. Based on the definition of JE, we state that student A justifiably envies student E, since A scores higher than E, but is matched to a less preferred school than E. We assign the outcome variable *any justified envy* to 1 if this applicant justifiably envies any other applicant, and 0 otherwise. The shaded areas capture the areas with justified envy.

In the real world, it is possible that a student with a low score is matched to a much-preferred school. This would distort the JE measure as a dummy variable, and leave many high-scoring applicants with positive JE. Therefore, a simple comparison of two matching mechanisms with the standard JE measure might not reveal the actual degree of matching stability. We create seven additional measures of justified envy similar to those in Kang et al. (2020), with slight adjustments so that the measures are directly comparable across provinces of difference sizes. Table 1 summarizes all JE measures and provide an example using applicant D in Figure 1. The construction of these measures intends to capture JE from multiple dimensions, taking into account the envy from gaps in admitted universities (how much better of a university does the justifiably envied student get admitted to), from gaps in *Gaokao* performance (how much lower does the justifiably envied student rank in *Gaokao* scores), and from both of them combined. Together with the standard JE measure as a dummy variable, they are the main outcome variables of interest.

3.2 Homogeneous Preferences

The key assumption that we make in the construction of JE is homogeneous preference, from both the universities and the applicants. All universities rank students by their *Gaokao* scores, with two main exceptions. First, the *zizhu zhaosheng* policy provides an opportunity to earn college-specific bonus points for students who demonstrate exceptional ability in certain areas and pass further college-specific tests. The bonus points vary by university but normally range from 10 to 60 points out of 750, the perfect score in *Gaokao* in the majority of provinces. [9] However, the effect of *zizhu zhaosheng* on JE is very limited since the Ministry of Education restricted the number of such applicants to be less than five percent of total applicants and it only started to become widely adopted in most provinces after 2011, the end of our sample period. [10] Second, a few universities admitting students to special programs such as music, acting, or sports would put decent weight on additional interviews or program-related exams, apart from the performance in *Gaokao*. This is again a very small portion of higher education in China. Therefore, we can reasonably assume that universities have homogeneous preferences over applicants based on their test scores.

A common barrier in the empirical literature is the identification of student preferences. While we do not have individual-level data on student preferences, the Chinese higher education system provides a context where we can reasonably assume students share homogeneous preferences over Tier-1 universities. First, the higher education system in China is highly stratified, with a limited amount of resources being allocated to a small group of universities. Therefore, university prestige and selectivity are highly stable over time, sending consistent signals of school quality to students and parents all over the country. The signals are particularly similar to applicants in the same province.

[9] There are a few exceptions to the 750 perfect score. In 2021, for example, Shanghai had a perfect score of 660, Jiangsu 480, and Hainan 900. These areas are not included in our final analytic sample for a different reason (to be discussed in the next subsection).

[10] See http://www.moe.gov.cn/srcsite/A15/moe_776/s3110/201212/t20121212_150771.html, retrieved on 6/1/2023.

Second, unlike higher education systems in many Western countries, where universities sometimes differ greatly in tuition (public vs. private), size (liberal arts vs. large research universities), and cultures (sports), these other factors which potentially have large impacts on students' decision making in the U.S. play very little role in China. College tuition rates are regulated by the Ministry of Education and remain small over the years (at roughly 800 USD). College sports receive very little attention. Also, all Tier-1 universities, the focus of our study, are large, public research universities. Instead, students' preferences towards universities in China are highly influenced by their prestige and rankings. Moreover, since college admission is administrated at the provincial level and the main identification strategy relies on comparison of JE within-province, other factors such as distance to school, exposure to college advertisement and information, and local cultural norms are likely homogeneous among students from the same province.

Given these facts, however, it is still likely that students will have homogeneous preferences on universities of similar rankings and reputations. Thus, we partition all Tier-1 universities into groups based on the average *Gaokao* scores of students admitted to, and thus attending, the university. Therefore, universities in the same group will be of similar quality (Ha, Kang and Song, 2020). We assume that student preferences are homogeneous over universities of different groups while allowing for heterogeneous preferences on universities within-group. Specifically, we assign universities into rank groups (RGs) based on the average-admission-score-percentile of applicants from the same year-province-track cell, where the competition and the matching game take place. To construct the average-admission-score-percentile, we first compute the average *Gaokao* scores of the incoming freshman cohort for each university. We then locate each university's percentile in the distribution of all universities. In particular, RG-1 includes schools from the top 10%, RG-2 through RG-4 consists of schools from the top 10% to 30%, the top 30% to 60%, and the rest, respectively.

The selection of cutoffs generates discontinuity in group assignments for schools near the average admission scores cutoff. Therefore, we further restrict the homogeneity assumption to: *students have homogeneous preferences over schools in different RGs with at least 10 percentile distance apart.* In other words, taking universities from RG-1 and RG-2 as an example, we assume all applicants prefer any RG-1 university to any RG-2 university if there are at least 10 percent of schools between the RG-1 university and the RG-2 university when all schools are ordered by the average admission score. We validate our results under assumptions with different rules. We also present results where we construct RGs based on different cutoff selections.

3.3 Summary Statistics

We restrict the sample to students admitted to Tier-1 universities, with *Gaokao* scores above the province-track-year-specific Tier-1 admission cutoffs. These universities have higher reputations across provinces such that the homogeneous preference assumption is most likely to hold among students submitting applications to them. We also limit the sample to the province of Inner Mongolia and other municipalities which remained using IA from 2005 to 2011. These municipalities include Beijing, Shanxi, Heilongjiang, Shandong, Qinghai, and Gansu. Such selection allows for a comparison between matching outcomes in Inner Mongolia and in provinces adopting IA. Our final analytic sample includes 887,615 *Gaokao* test takers between 2005 and 2011.

Table 2 presents the summary statistics for individuals in our sample. The first two columns report the statistics of Inner Mongolia before and after the reform, respectively. The last two columns report the statistics of control provinces. Students in Inner Mongolia have always had their college preference submission day scheduled after their *Gaokao* scores were released. However, in some other provinces, the full preference submission period might take place before students take *Gaokao* (ex-ante submission), or after they take the exam but before the grades are posted (ex-interim submission). Some provinces also switched from ex-ante submission to ex-post submission during this time period. Table A1 shows the timing of the preference submission dates over the years for all provinces in our sample. The relative date of preference submission to the dates of *Gaokao* might have an impact on matching stability, so we control for them in the regression model. [11] In the control provinces, roughly half of all *Gaokao* test takers in our sample took the exam and submitted the college preferences with perfect information on their test scores. Less than 20 percent of college applicants were set to submit their preferences before they took the exam. Roughly 8 percent of students were matched to schools using the real-time dynamic mechanism, i.e. the Inner Mongolia post-2007 subsample, while the rest of them were matched to universities under IA. Over 80 percent of students selected the STEM track in 10th grade and took the corresponding college entrance exam two years later.

For each student, we construct eight JE measures as described in Table 1. According to our homogeneous preference assumption, a student s_1 justifiably envies another student s_2 if student s_1 is matched to college c_1, student s_2 is matched to college c_2, student s_1 scores higher than student s_2, but c_1 is in a lower rank group than c_2 and they are at least 10 percentile apart in the distribution of all schools. As shown in Table 2, the standard JE measure shows that over 80 percent of students justifiably envied at least one other student. However, as we have discussed, the high percentage could result from one student with low scores being admitted into a top university. From the additional JE measures, we see that for an average Inner Mongolia applicant before 2007, 2.2 percent of all applicants whom she was competing with had lower test scores but were admitted to a more-preferred school. This number decreased to 1.8 percent after the real-time dynamic mechanism went into effect. Applicants in the control provinces justifiably envied a higher proportion of all applicants. On average, they envied 3.2 percent of competing applicants pre-2007 and 2.7 percent of competing applicants post-2007.

For JE at the university rank group dimension, the average maximum degree of envy

[11] Wu and Zhong (2014) use data from an elite university and find that ex-ante submission under IA leads to better match fairness than ex-post submission under the parallel mechanism.

in Inner Mongolia is 1.407 before the reform and 1.627 after the reform, while the average maximum degree of envy in the control provinces is quite stable over time between 1.6 and 1.65. The maximum possible degree of envy is 3 since we constructed four rank groups. Different from the demonstrations in Figure 1 and Table 1, we standardize the degree of envy for the student-rank dimension such that it is also between 0 and 3. Thus, the magnitudes are comparable across measures from different dimensions. On average, the lowest-ranked individual an applicant justifiably envies is ranked equivalent to 0.85-1 RG below the student herself. The average degree of envy is around 0.9-1 in university rank group and 0.3-0.4 in student rank. From the diagonal distance perspective, the maximum degree of envy is around 1.9 and the average degree of envy is around 1 in both Inner Mongolia and the control provinces. In general, however, we observe that these JE measures increased in Inner Mongolia, but not in control provinces, after the reform.

4 Empirical Strategy

We use a generalized difference-in-differences design to identify the causal effect of Inner Mongolia's switch to real-time dynamic mechanism (from IA) on matching stability through various JE measures. The econometric model takes the following form:

$$Y_{itpk} = \alpha_t + \mu_p + \beta Dynamic_{tp} + \gamma Timing_{tp} + \lambda Rank_i + \delta N_{tpk} + \eta_k + \varepsilon_{itpk} \quad (1)$$

where Y_{itpk} is the JE measure for student i taking *Gaokao* in year t, province p, and track k. $Dynamic_{tp}$ is a dummy variable that is equal to 1 if province p in year t was implementing the real-time dynamic mechanism, and 0 otherwise. $Timing_{tp}$ indicates the timing of preference submission, $Rank_i$ denotes student i's percentile rank within the year-province-track cell, and N_{tpk} denotes the total number of students within that cell. α_t, μ_p, and η_k are year, province, and track fixed effects, respectively. The standard errors are clustered at the province-track level. The coefficient of interest is β, representing the difference-in-differences estimates of the effect of Inner Mongolia's dynamic mechanism.

We also present results using an event-study design. The regression model is given by:

$$Y_{itpk} = \alpha_t + \mu_p + \sum_{\substack{\tau=-2 \\ \tau\neq-1}}^{\tau=4} \beta_\tau Dynamic_{\tau p} + \gamma Timing_{tp} + \lambda Rank_i + \delta N_{tpk} + \eta_k + \varepsilon_{itpk} \tag{2}$$

where $Dynamic_{\tau p}$ is set to 1 if province p introduced the real-time dynamic mechanism τ year ago, and 0 otherwise. We omit 2006, the year before the introduction of the real-time dynamic mechanism. As we have discussed, during the time period of our sample, most other provinces in China also saw a transition from IA to the parallel mechanism, a hybrid of IA and DA. The event-study representation allows for a straightforward comparison with results from previous studies (Kang et al., 2020) on the effect of such a transition to the parallel mechanism.

Moreover, since there were only 6 provinces that remained using IA from 2005 to 2011, and that these provinces were of different sizes and had distinct characteristics, the parallel trend assumption might not necessarily hold. Figure 2 shows the trends of unadjusted justified envy in Inner Mongolia and other provinces in the control group. The first subfigure presents the average probability of justifiably envying any other applicant, and the second subfigure presents the average number of justifiably envied applicants, as a proportion of all competing applicants. In both subfigures, we do not see evidence against a parallel trend. In fact, the pre-reform trends in these two measures are very similar in Inner Mongolia and in other provinces. Nevertheless, to further validate our results, we use the synthetic control method (Abadie and Gardeazabal, 2003; Abadie, Diamond and Hainmueller, 2010; Abadie, 2021) to construct a weighted combination of these provinces. The predictors of the JE measures we use include the proportion of STEM (versus non-STEM) test takers, the number of applicants in a year-province-track cell, and pre-intervention JE measures. We choose the corresponding parameters such that the resulting synthetic Inner Mongolia best approximates the trajectory of the JE measures in pre-intervention Inner Mongolia. Table A2 provides the province weights in the synthetic Inner Mongolia.

5 Effects of Inner Mongolia Reform

In this section, we first present the effects of Inner Mongolia reform on justified envy (JE) measures from the generalized difference-in-differences regressions specified by Equation 1. We also present results from the same regression model using different sets of rank groups (RGs), and from relaxing the homogeneous preference assumption. Then, we show the event-study estimates from Equation 2, followed by the synthetic control illustrations. Finally, we present the robustness checks and heterogeneous effects.

5.1 Main Results

Table 3 presents the main results from the difference-in-differences design specified by Equation 1. The first column shows the coefficients on the likelihood of justifiably envying any other applicant. Changing from IA to the real-time dynamic mechanism increases the likelihood of justified envy by 1.8 percentage points, which is not statistically significantly different from 0. Column 2 indicates that the switch to real-time dynamic mechanism had no impact on the number of students one justifiably envies.

Columns 3 through 8 show results for degrees of justified envy in three dimensions: how much better of a university did the justifiably envied students get admitted to, how much lower did the justifiably envied students rank in terms of their *Gaokao* scores, and a combination of the above two. Consistent with the results in Columns 1 and 2, there is no evidence supporting that the real-time dynamic mechanism reduced justified envy when compared to IA. In the student rank dimension, the real-time dynamic mechanism led to a 0.155 standard deviation (sd) increase in the maximum degree of envy and a 0.174 sd increase in the average degree of envy. The coefficients in the university rank group dimension, while being less significant, are also positive and of decent size in magnitude (0.247 sd and 0.080 sd). When we combine the envy in the university quality and student test-score dimension, we find that the maximum degree of envy increased by 0.255 sd and the average degree of envy increased by 0.126 sd. These estimates suggest that the real-time dynamic mechanism was just as bad as IA in eliminating justified envy, if not worse.

Recall that we first rank schools based on the *Gaokao* scores of the freshman cohort from the current year, and assign the top 10% of schools to the first rank group (RG-1), then schools from the top 10% to 30%, top 30% to 60%, and the rest to RG-2 through RG-4, respectively. The RGs are constructed with these thresholds because students normally have clearer preferences among top-quality schools than among low-quality schools. Nevertheless, the determination of the cutoffs between these rank groups may seem arbitrary. We adjust the rank group selections and run regressions with the alternative selections. The results are shown in Table 4. Panel A shows estimates when we rank schools based on the *Gaokao* scores of admitted students from the previous year. The difference between estimates in Panel A and Table 3 would capture any drastic change in school quality or student preferences towards schools in two consecutive years. [12] In Panel B we assign 25% of schools to each rank group. In Panel C we assign universities into eight groups instead of four groups based on the following cutoff percentiles: 5%, 10%, 20%, 30%, 45%, 60%, and 80%. Panels D and E are analogous to panels A and B except that there are now eight groups. Of all panels, no evidence suggests that the real-time dynamic mechanism eliminated more justified envy than IA. The conclusions drawn from Table 3 are robust to different assignments of rank groups.

Since some students might be indifferent between a lower-ranked university in a higher RG and a top university in a lower RG, for our main regressions we stated the homogeneous preference assumption as: *students have homogeneous preferences over schools*

[12]Since we do not have 2004 data, we drop all 2005 observations.

in different RGs with at least 10 percentile distance apart. In other words, all applicants prefer a RG-1 university to a RG-2 university if and only if there are at least 10 percent of schools in between them in the ranking of schools. Here, we loosen the justified envy definition and test if the results would be any different. Table 5 presents the results from relaxing the percentile distance restriction. In Panel A we alter the restriction from a 10 percentile distance to 5 percentile distance, and in Panel B we remove the restriction such that students have homogeneous preferences over schools in different RGs, regardless of their specific location within the RGs. The results in Table 5 are highly consistent with those from the main specification. Using alternative homogeneous preference assumptions does not impact the regression results.

5.2 Event-Study Estimates

Figure 3 presents the event-study results from Equation 2 on all eight JE measures. Each subfigure plots the point estimates and the corresponding 95% confidence intervals. The horizontal axis represents the number of years from the Inner Mongolia reform, with 2007 being year 0 in all subfigures. The graphs show no clear pre-trend prior to the Inner Mongolia reform. The first column of Figure 3 shows that the real-time dynamic mechanism led to no improvement in the probability of justifiably envying any other applicant, and that it did not reduce the number of justifiably envied students per applicant in the first four years after the reform. The probability of justifiably envying at least one other applicant increased by around 10 percentage points (0.25 sd) in Inner Mongolia in the first two years after the reform, before dropping to a similar level as the control provinces.

For JEs in the university quality and student test-score dimensions, we also observe a surge in envy in the reform year. Using statistics from Table 3, it suggests a 0.80 sd increase in maximum degree of envy and a 0.60 sd increase in average degree of envy in university rank groups in 2007. When we combine both dimensions of envy, these numbers increased to 0.83 sd and 0.66 sd, respectively. The degree of envy in all dimensions dropped in subsequent years, but remained at a higher level than those in the control group until the final year, when they became statistically indistinguishable. Thus, at its initial stage, the transition from IA to the real-time dynamic mechanism clearly led to less stable matching. The instability is observed both in the probability of envying other applicants, and in the magnitudes of the envy. We do see that over the years since the reform, there was a decreasing trend in JE measures, indicating that applicants in later years potentially learned from the experience of previous applicants and adjusted their preference submissions strategy accordingly. Therefore, it is possible that the same trend continued after 2011 and eventually all JE measures reached an equilibrium where the Inner Mongolia's real-time dynamic mechanism produced a more stable matching than IA. Overall, these results are consistent with our findings above that the real-time dynamic mechanism is no better than IA in eliminating justified envy.

5.3 Synthetic Control Results

We present the synthetic control results in Figure 4. The black line in each subfigure represents the gap in justified envy measures between Inner Mongolia and the synthetic Inner Mongolia. To assess the significance of our estimates, we also conduct a series of placebo tests by applying the synthetic control methods to every other "untreated" province in our sample. In each test, we treat one of the other provinces as the reform province and compute the estimated effect of the hypothetical reform. The estimated effects are represented by the gray lines in each subfigure. They provide us with the distribution of the estimated gaps if no intervention took place.

As shown in Figure 4, the synthetic Inner Mongolia almost perfectly matches Inner Mongolia in justified envy outcomes prior to the Inner Mongolia reform in 2007. The trends in JE measures after 2007 were very similar to those in Figure 3, reaffirming our conclusions above.

5.4 Robustness Checks

We check the robustness of our results with a variety of different specifications. Table 6 shows the results of the robustness checks. Inner Mongolia implemented the real-time dynamic mechanism in 2007 for tier-1 universities, and then expanded to all universities starting in 2008. Since we limit the sample to students admitted to tier-1 universities, our main results treat 2007 as the first year of treatment (year 0). Panel A of Table 6 presents the difference-in-differences results from dropping the year of 2007. Results from 2007 would not reflect the equilibrium levels of justified envy under the real-time dynamic mechanism both because it is the first year of the reform and because that the mechanism was only applied to students matched to Tier-1 universities. The estimates in Panel A are consistent with the event-study results, indicating that most positive effects observed in Table 3 came from the initial treatment year. However, even after dropping 2007, there is no evidence for the real-time dynamic mechanism as a better mechanism at eliminating justified envy.

Our definition of homogeneous preference relies on the assumption that college applicants treat college quality as the dominant factor in their formation of college preferences. As we have discussed previously, China's unique centralized education system has laid a solid and supportive foundation for this assumption. However, it is possible that some applicants put heavy weight on the location of their schools and would prefer a college in their own province to an out-of-province college even if the own-province college was of much lower quality. Therefore, we drop student observations who were matched to a school from their own province and present the regression results in Panel B. The preferences of these students had the highest likelihood of violating our homogeneous preference assumption. The results from the truncated sample show highly consistent significance level with the main results.

In Panel C we drop students who were admitted to majors that have the highest or lowest admission cutoff scores within the university. Some schools might have a "star" major or a very poor major that distorts students' preferences towards schools. For instance, a student might prefer a computer science major in an engineering-oriented school in rank group 2 (RG-2) to the same major in a liberal arts-oriented school in rank group 1 (RG-1). By our assumption of homogeneous preferences, we assume that every student prefers a RG-1 university to a RG-2 university if they are at least 10 percentile apart in the distribution of all universities. However, this might not hold in our example, where the specific major within the school plays an important role in student preferences. Therefore, we drop students who were admitted to such majors since they had a higher likelihood of violating the homogeneous preference assumption. The results are shown in Panel C. While the significance levels of some estimates differ from previous specifications, most coefficients remain positive and insignificant.

Finally, since each student submitted school-major choices instead of school choices alone. We build the same number of rank groups and with the same cutoff criteria as in the main specification, but use a school-major pair as an data point rather than the school itself. In other words, we rank all school-major pairs within a year-province-track cell based on their admission cutoff scores. We then create four rank groups, with 10%, 20%, 30%, and 40% of school-major pairs in each rank group. Thus, student preferences are at the school-major level rather than at the school level. With the same homogeneous preference assumption, the estimates in Panel D again show that the real-time dynamic mechanism did not lead to a more stable matching than IA. Overall, our main results are robust to different specifications.

5.5 Heterogeneous Effects

We discuss heterogeneous effects of the Inner Mongolia reform in two dimensions: student test scores and county-level economic status. Figure 5 shows the effects of the switch to real-time dynamic mechanism on applicants at different test score distributions. It shows that most of the increase in JEs that we observe in Table 3 came from applicants with test scores in the range of 50th-75th percentile. Moreover, the (undesired) effect on JE was most pronounced in the student rank dimension. In other words, after the Inner Mongolia reform, applicants with test scores within this range experienced much larger gap in test scores between themselves and other applicants that they envied. Applicants that obtained better matching outcomes were with even lower priority than before.

College applicants with either low test scores, in particular those below the 40th percentile, or very high test scores, in particular those above the 85th percentile, experienced similar justified envy under the real-time dynamic mechanism and IA. One explanation is that the baseline mean of JE was low for these applicants. The amount of justified envy by test score distribution under the real-time dynamic mechanism and under IA is presented in Figure A.3. In all subfigures we see hump-shaped curves: justified envy is the lowest at both ends of the test score distribution. For JE in the student rank dimension, it is highest for applicants between 50th and 75th score percentile, consistent with our findings in Figure 5 that the effects were largest for applicants in this range.

The observed shape in JE is expected given the definitions of JE measures and the homogeneous preference assumption that we make. For low-performing applicants, they were less likely to envy others for two main reasons: first, less applicants had lower scores than they did; and second, most schools that they were matched to were in the last rank group and by definition they would not envy each other within a rank group. For students with very high test scores, they had more precise expectation about the probability of admission since there were less applicants with higher priority than they did. Therefore, they were also less likely to envy others, although most applicants had scores below them. For those in the middle of the test score distribution, however, they faced both the uncertainty about choices from higher-ranked applicants and the "threats" from lower-ranked applicants who had the potential to achieve better matching outcome. These evidence suggests that policy adjustments should gear towards medium-ranked applicants because they have the highest justified envy and justified envy measures are most sensitive to policy changes among these applicants.

Moreover, applicants from counties with different income levels might have different experience from the Inner Mongolia reform. Using the *Gaokao* id , we are able to match 77.6% of applicants to the counties that took the exam in. We then identify applicants that took the exam in one of the 832 officially-designated severely impoverished counties. The list of impoverished counties was determined by the State Council of the People's Republic of China in 2012 based on county-level average income. [13] Using their *Gaokao* id, we are able to match 77.6% of applicants to the counties that they took the exam in. Table 7 reports the difference-in-differences estimates from the main specification for the impoverished counties and all other counties that were successfully matched. In severely impoverished counties, college applicants experienced larger increase in justified envy after the real-time dynamic mechanism was implemented. The proportion of applicants that justifiably envied at least one other applicant increased by 3 percentage points, almost twice as much as the effect observed in Table 3, and tripled the coefficients for non-impoverished counties. Significant increases in justified envy were also found in the university RG and student rank dimension. By contrast, for non-impoverished counties in Panel B, all eight coefficients were smaller in magnitudes than those in Panel A. Many of them were also indistinguishable from zero. This suggests that applicants from more disadvantaged backgrounds had less desirable matching outcomes than others in the same province-year-track cell after the reform. However, this is unsurprising as applicants from poor counties are expected to have less information gains from the reform for at least two reasons: first, they might not fully understand the dynamic mechanism due to the lack of information and instructions; second, they had less internet and technology resources to be able to fully take advantage of the dynamic mechanism.

[13] A full list of these counties are available on the National Rural Revitalization Administration website.

6 Comparison with Parallel Mechanism

Within the last two decades, all provinces except Inner Mongolia have switched from the immediate acceptance (IA) mechanism to a "parallel mechanism" to match their students to colleges in China. Chen and Kesten (2017) characterize the parallel mechanism as a hybrid of the deferred acceptance (DA) mechanism and IA and show theoretically that the transition from IA to the parallel mechanism would produce more stable matching. Their findings are supported by other experimental and empirical studies (Chen and Kesten, 2019; Chen, Jiang and Kesten, 2020; Ha, Kang and Song, 2020; Kang et al., 2020). Among these studies, Kang et al. (2020) are the first to use large-scale empirical data to compare parallel mechanism to IA in multiple fairness measures. We replicate their results using our adjusted justified envy measures and compare the real-time dynamic mechanism to the parallel mechanism through their individual comparisons with IA.

Table 8 shows the effects of switching from IA to parallel mechanism on matching stability. As our paper uses the same data set, in Panel A we replicate the results in Kang et al. (2020) with the adjusted JE measures in this paper. We use the regression model specified by Equation 1. Similar to our main design, we compare JE outcomes of provinces that switched to the parallel mechanism to those of provinces that kept using IA between 2005 and 2011, exploiting the staggered implementation of the parallel mechanism. [14] We exclude Inner Mongolia from the sample since, unlike all other provinces in China, Inner Mongolia neither switched to the parallel mechanism nor remained using IA. The estimates suggest that the parallel mechanism is better at eliminating justified envy across all measures. The likelihood of justifiably envying any other applicant dropped by 5.1 percentage points (0.13 sd); the number of justifiably envied students as a proportion of all competing applicants also decreased by 1.3 percentage points (0.26 sd). In other words, when parallel mechanism is adopted, it is significantly less likely that a student

[14] Table A3, extracted from Kang et al. (2020), shows the detailed timing of the policy reform in each province.

with lower scores than the other student would be matched to a more-desired school. Estimates of other dimensions of JE support this conclusion. When we combine envy in school quality and student test-score, there was a 0.50 sd decrease in maximum degree of envy and a 0.22 sd decrease in average degree of envy. The estimates are in sharp contrast to those in Table 3, where no significant negative effects are found. Therefore, while switching from IA to the real-time dynamic mechanism did not impact JE, switching from IA to the parallel mechanism significantly improved JE elimination in various dimensions.

In Panel B of Table 8 we keep the full 2005-2011 sample, including provinces that kept using IA, provinces that switched to the parallel mechanism, and Inner Mongolia. Then, we add to our original model an additional binary variable, $Parallel_{tp}$, which indicates if province p in year t adopts the parallel mechanism. [15] We then report the estimates of coefficients for $Dynamic_{tp}$ and $Parallel_{tp}$. They represent the effects of switching from IA to real-time dynamic mechanism and parallel mechanism, respectively. Consistent with previous results, Panel B of Table 8 shows that parallel mechanism outperformed IA in matching stability, while the real-time dynamic mechanism did not. The coefficients for $Dynamic_{tp}$ confirm our main results with higher power due to the increase in sample size. Moreover, Figure A.1 presents the event-study estimates of the effects of switching from IA to parallel mechanism. Figure A.2 shows similar results using the synthetic control method and displaying the effects on matching stability in Zhejiang province, one of the provinces that switched to IA, as an example. A comparison of these two figures and Figures A.1 and A.2 further supports our conclusions from above.

Finally, Ha, Kang and Song (2020) introduce two other proxies of matching stability: the "rank range" and the "match index". Assuming that universities have homogeneous preferences on students, strictly preferring students with higher *Gaokao* scores, and students have homogeneous preferences on universities, strictly preferring schools with

[15] The corresponding regression model is $Y_{itpk} = \alpha_t + \mu_p + \beta_1 Dynamic_{tp} + \beta_2 Parallel_{tp} + \gamma Timing_{tp} + \lambda Rank_i + \delta N_{tpk} + \eta_k + \varepsilon_{itpk}$.

higher ranks based on revealed preferences. Then, a unique stable matching would be one that perfectly stratifies students. Given an ordering of all applicants from highest-scoring to lowest-scoring, in a stable matching it is divided into $N+1$ groups (N schools plus $\emptyset$). Student-groups are then matched to universities, with higher-scoring student-groups being matched to higher ranked schools. "Rank range" and "match index" represent the level of stratification precision. Rank range is constructed as the difference between the highest and lowest student-rank within each year-university-track-province admission cell. In a perfectly stratified matching, the rank ranges are the smallest possible for every admission cell, being equal to the number of students admitted (quota) minus one. Match index is a standardized version of the rank range and is defined as (rank range + 1)/quota. It is equal to one in the perfectly stratified matching. As a result, all else equal, rank range and match index would be smaller when matching stability improves. We construct the rank range and match index measures for each year-university-track-province admission cell and compare these measures under the real-time dynamic mechanism, parallel mechanism, and IA. Table 9 reports the difference-in-differences estimates under the same specifications used in the main results. The first two columns show that the rank range and match index did not change since Inner Mongolia implemented the real-time dynamic mechanism. From columns 3 and 4, we see evidence of significantly reduced rank range and match index after provinces switched from IA to the parallel mechanism. The last two columns show consistent results that the parallel mechanism yielded more stable matching than both IA and the real-time dynamic mechanism.

7 Discussion

The results of our paper suggest that the real-time dynamic mechanism, implemented by Inner Mongolia starting in 2007, did not improve college matching stability upon the immediate acceptance (IA) mechanism, which was adopted by Inner Mongolia and most other provinces before 2007. Our findings are quite surprising given the previous theoretical literature on dynamic matching mechanisms (Bó and Hakimov, 2020, 2022; Grenet, He and Kübler, 2022; Klijn, Pais and Vorsatz, 2019), which show that dynamic mechanisms generally outperform their static versions in truth-telling preference revelation, matching efficiency and stability. Gong and Liang (2020) also study Inner Mongolia's college matching mechanism. They show both theoretically and experimentally that the real-time dynamic mechanism is as stable and efficient as IA, and induces a higher truth-telling rate than both the deferred acceptance (DA) mechanism and IA. Therefore, it is important to understand why the real-time dynamic mechanism did not empirically yield an ideal matching stability outcome as expected. Government statistics show that on average an Inner Mongolia applicant changed her college-major choices 130 times in 2018 during the day of college preference submission, so it is clear that students are well-aware of the policy and are adjusting their application decisions based on the tentative admission outcomes. [16] Although we cannot empirically test them, we provide a few potential reasons for the unexpected results from the Inner Mongolia reform.

7.1 Immaturity of Design of the Mechanism

While the "real-time dynamic mechanism" was first implemented in 2007, it has evolved year after year. As discussed in Section 2.2, under the real-time dynamic mechanism, all participants were divided N groups. Different groups had different preference submission deadlines such that only applicants with similar scores will be confirming their choices at the same time. The number of groups and their selection criteria have changed since 2007. Table A4 shows the selection of these applicant groups over years. In 2007, all applicants applying to Tier-1 universities formed only one group. This number increased

[16] The public report is released by the Inner Mongolia Education Examination Authority. See https://www.nm.zsks.cn/zxyw/201808/t20180801_1177.html, retrieved on 8/17/2023. The average number of changes made by Inner Mongolia applicants were over 90 in all years between 2014-2018, when data were made publicly available.

to four in 2008, and seven in 2010.

When more groups are formed, applicants in each group would be provided with more information about school vacancies and thus would be more confident about their probability of admission to any school they select (Luflade, 2018). Take the extreme example when each group consists of only one applicant, this applicant would be certain about which school she would get in and which school she would not because all applicants (groups) who have scores higher than she does have already submitted their choices. The uncertainty about the probability of admission comes from the uncertainty about the choices that applicants with higher priority (i.e., higher test scores) would be making, and this uncertainty decreases in the number of groups because applicants in higher-ranked groups have an earlier submission deadline.

The patterns that we observe in Figure 3 support this argument: matching stability was the worst in 2007 and has since improved gradually as applicants were divided into more and smaller groups. While we do not have admission data after 2011, we could anticipate more stable matching in more recent years as the number of groups kept increasing to nine in 2016 and ten in 2021. For college applicants, learning the dynamic mechanism and the formation of a best application strategy could also take a longer time than what our data covers because college application takes place only once a year. Therefore, it is possible that the dynamic mechanism would eventually lead to a steady state matching equilibrium that is more stable than what IA or the parallel mechanism would generate.

7.2 (Un)strategic Undermatch and "Collusion"

Since applicants can observe their real-time admission outcomes given the tentatively submitted college-major preferences, their final preference submissions are largely affected by the observed outcomes of their previous choices. If a student realizes that her desired college has more tentative applicants with higher scores than she expected, then she might update her expectation and turn to a less-desired school for final submission.

Given the nature of the dynamic mechanism, it is common that students are wavering among plausible options. For high scorers, rather than directly choosing their top choice, they could be testing out different schools or temporarily settling down at schools for which they over-qualify. Both of these behaviors are particularly prevalent among risk-averse applicants. However, in both cases, these high scorers drive up the admission score cutoffs of schools that would have been perfect matches for applicants with relatively lower *Gaokao* scores. These lower scorers are then mistakenly scared away from submitting applications to these schools, missing the opportunity to be admitted to schools that they are qualified for. This opens spaces for some risk-loving applicants with even lower scores to get into better schools than other applicants with higher scores.

Given the descriptions above, it has also become clear that if student A, whose preferred school is S, can persuade (as many as possible) applicants with higher scores to strategically submit to an under-matched school, S, until the final submission deadline, it would discourage other applicants who have similar scores and are competing with A to apply to school S. Such "collusion" behavior weakly dominates the status quo for A, conditional on A being able to find such high achievers. Anecdotal evidence reveals that this is a common phenomenon in Inner Mongolia as many applicants successfully persuade their friends to collude with them.

Any decent size of such "collusion" would disrupt a stable matching and increases our measures of justified envy. This type of behavior could become less popular when applicants are partitioned into smaller groups, but remains plausible due to the heterogeneity in test scores within each group.

8 Conclusion

Our paper is the first to use large-scale empirical data to examine the effects of switching from immediate acceptance (IA) mechanism to the "real-time dynamic mechanism" on college matching stability in Inner Mongolia. Although previous theoretical and experimental studies have predicted that dynamic matching mechanisms would outperform static mechanisms in matching stability, we find no evidence that matching stability improved in Inner Mongolia in the first four years after the implementation of the real-time dynamic mechanism.

Stability and "no justified envy" were often used synonymously in the matching theory literature. We construct measures of justified envy (JE) and compare the JE outcomes under the real-time dynamic mechanism to those under IA. Using a generalized difference-in-differences design, we find no evidence that the real-time dynamic mechanism is better than IA in eliminating justified envy. It did not decrease the likelihood of an applicant justifiably envying at least one other applicant nor the number of justifiably envied applicants. This is also evident in six other JE measures that we constructed to capture the envy in school quality, in student test-score, and in a combination of the two. Event-study estimates show that JE increased significantly in the first year of the reform, and gradually dropped to levels similar to those in provinces adopting IA. Then, we compare our results to those from Kang et al. (2020), who study the impacts of switching from IA to the "parallel mechanism" on matching stability. We show that the parallel mechanism was significantly better at eliminating JE than the real-time dynamic mechanism, indicating a more stable matching under the parallel mechanism. Similar results are found when we use rank range and match index as proxies of matching stability. Finally, we discuss a few potential reasons for why the real-time dynamic mechanism did not yield expected results. The evolution of the mechanism over years could explain the pattern we observe in the event-study figures. The submission decisions of high-scorers could also have large subsequent impacts on the decisions and matching outcomes of low-scorers.

The results of our paper has important policy implications. For policy makers aiming to create stable matching between *Gaokao* test takers and Chinese universities, the parallel mechanism has shown to be more effective than the real-time dynamic mechanism. An obvious path for education authorities in Inner Mongolia would be to switch their college matching mechanism to the parallel mechanism. One limitation of our results is we only observe matching outcomes in the first four years after the Inner Mongolia reform. Future research on the real-time dynamic mechanism can look at its long-term effects on the matching stability. As more recent cohorts of applicants demonstrate improved learning and take advantage of better-developed mechanism, we might see different matching outcomes that favor the real-time dynamic mechanism.

References

Abadie, Alberto. 2021. "Using synthetic controls: Feasibility, data requirements, and methodological aspects." Journal of Economic Literature, 59(2): 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." Journal of the American Statistical Association, 105(490): 493–505.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The economic costs of conflict: A case study of the Basque Country." American Economic Review, 93(1): 113–132.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School choice: A mechanism design approach." American Economic Review, 93(3): 729–747.

Abdulkadiroğlu, Atila, Parag A Pathak, and Alvin E Roth. 2009. "Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match." American Economic Review, 99(5): 1954–1978.

Balinski, Michel, and Tayfun Sönmez. 1999. "A tale of two mechanisms: student placement." Journal of Economic Theory, 84(1): 73–94.

Bó, Inácio, and Rustamdjan Hakimov. 2020. "Iterative versus standard deferred acceptance: Experimental evidence." The Economic Journal, 130(626): 356–392.

Bó, Inácio, and Rustamdjan Hakimov. 2022. "The iterative deferred acceptance mechanism." Games and Economic Behavior, 135: 411–433.

Chen, Li, and Juan Sebastián Pereyra. 2019. "Self-selection in school choice." Games and Economic Behavior, 117: 59–81.

Chen, Li, Juan S Pereyra, and Min Zhu. 2022. "Time-constrained Dynamic Mechanisms for College Admissions." arXiv preprint arXiv:2207.12179.

Chen, Yan, and Onur Kesten. 2017. "Chinese college admissions and school choice reforms: A theoretical analysis." Journal of Political Economy, 125(1): 99–139.

Chen, Yan, and Onur Kesten. 2019. "Chinese college admissions and school choice reforms: An experimental study." Games and Economic Behavior, 115: 83–100.

Chen, Yan, and Tayfun Sönmez. 2006. "School choice: an experimental study." Journal of Economic theory, 127(1): 202–231.

Chen, Yan, and Yinghua He. 2021. "Information acquisition and provision in school choice: an experimental study." Journal of Economic Theory, 197: 105345.

Chen, Yan, Ming Jiang, and Onur Kesten. 2020. "An empirical evaluation of Chinese college admissions reforms through a natural experiment." Proceedings of the National Academy of Sciences, 117(50): 31696–31705.

Ding, Tingting, and Andrew Schotter. 2017. "Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms." Games and Economic Behavior, 103: 94–115.

Dur, Umut, Robert G Hammond, and Onur Kesten. 2021. "Sequential school choice: Theory and evidence from the field and lab." Journal of Economic Theory, 198: 105344.

Ergin, Haluk, and Tayfun Sönmez. 2006. "Games of school choice under the Boston mechanism." Journal of Public Economics, 90(1-2): 215–237.

Gale, David, and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." The American Mathematical Monthly, 69(1): 9–15.

Gong, Binglin, and Yingzhi Liang. 2020. "A dynamic matching mechanism for college admissions: Theory and experiment." Working paper.

Grenet, Julien, YingHua He, and Dorothea Kübler. 2022. "Preference discovery in university admissions: The case for dynamic multioffer mechanisms." Journal of Political Economy, 130(6): 1427–1476.

Hassidim, Avinatan, Assaf Romm, and Ran I Shorrer. 2016. ""Strategic" Behavior in a Strategy-Proof Environment." 763–764.

Ha, Wei, Le Kang, and Yang Song. 2020. "College matching mechanisms and matching stability: Evidence from a natural experiment in China." Journal of Economic Behavior & Organization, 175: 206–226.

Kang, Le, Wei Ha, Yang Song, and Sen Zhou. 2020. "Matching Mechanisms and Justified Envy: Stability on the Intensive Margin." Available at SSRN: https://ssrn.com/abstract=4548508.

Klijn, Flip, Joana Pais, and Marc Vorsatz. 2019. "Static versus dynamic deferred acceptance in school choice: Theory and experiment." Games and Economic Behavior, 113: 147–163.

Lien, Jaimie W, Jie Zheng, and Xiaohan Zhong. 2016. "Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory." Experimental Economics, 19: 116–150.

Lien, Jaimie W, Jie Zheng, and Xiaohan Zhong. 2017. "Ex-ante fairness in the Boston and serial dictatorship mechanisms under pre-exam and post-exam preference submission." Games and Economic Behavior, 101: 98–120.

Luflade, Margaux. 2018. "The value of information in centralized school choice systems." Duke University, 4(6): 7.

Pais, Joana, Agnes Pintér, and Róbert F Veszteg. 2011. "College admissions and the role of information: An experimental study." International Economic Review, 52(3): 713–737.

Pais, Joana, and Ágnes Pintér. 2008. "School choice and information: An experimental study on matching mechanisms." Games and Economic Behavior, 64(1): 303–328.

Pathak, Parag A, and Tayfun Sönmez. 2008. "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism." American Economic Review, 98(4): 1636–1652.

Pathak, Parag A, and Tayfun Sönmez. 2013. "School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation." American Economic Review, 103(1): 80–106.

Rees-Jones, Alex. 2018. "Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match." Games and Economic Behavior, 108: 317–330.

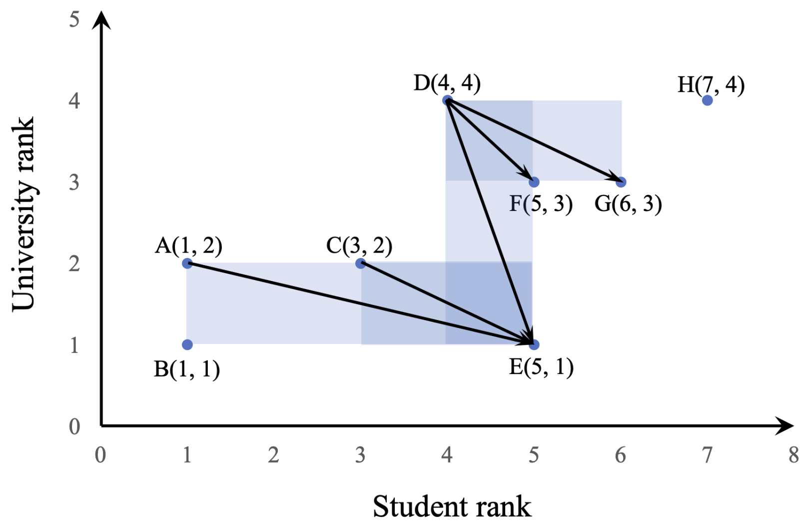
Shorrer, Ran I, and Sándor Sóvágó. 2022. "Dominated choices in a strategically simple college admissions environment: The effect of admission selectivity." Working paper.

Stephenson, Daniel. 2022. "Assignment feedback in school choice mechanisms." Experimental Economics, 1–25.

Wu, Binzhen, and Xiaohan Zhong. 2014. "Matching mechanisms and matching quality: Evidence from a top university in China." Games and Economic Behavior, 84: 196–215.

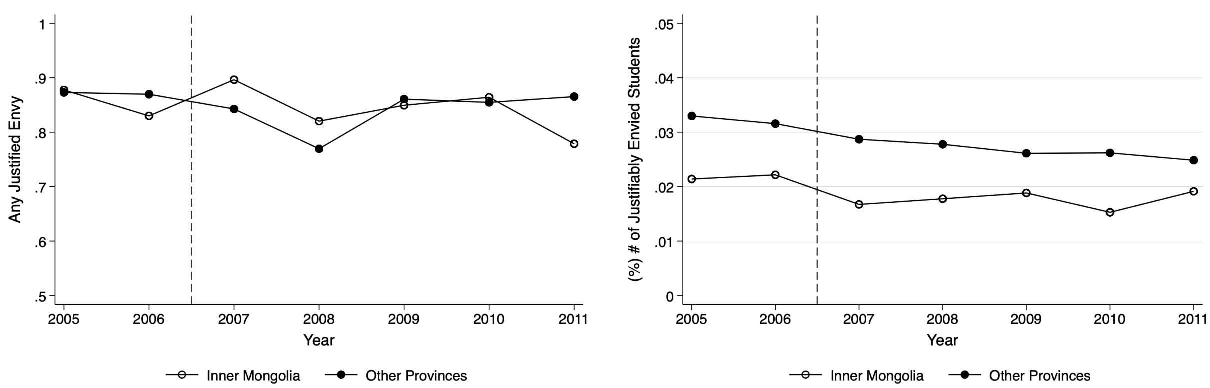
Figures and Tables

Figure 1: Illustration of Justified Envy Measures



Notes: This figure plots a possible matching outcome and illustrates the justified envy measures. The horizontal axis represents student ranking and the vertical axis represents college ranking. A student justifiably envies another student if the other student ranks lower but is matched to a higher-ranked school. On the figure, each student justifiably envies all other students to her bottom-right. The shaded areas capture the areas with justified envy.

Figure 2: Unadjusted Justified Envy Trends

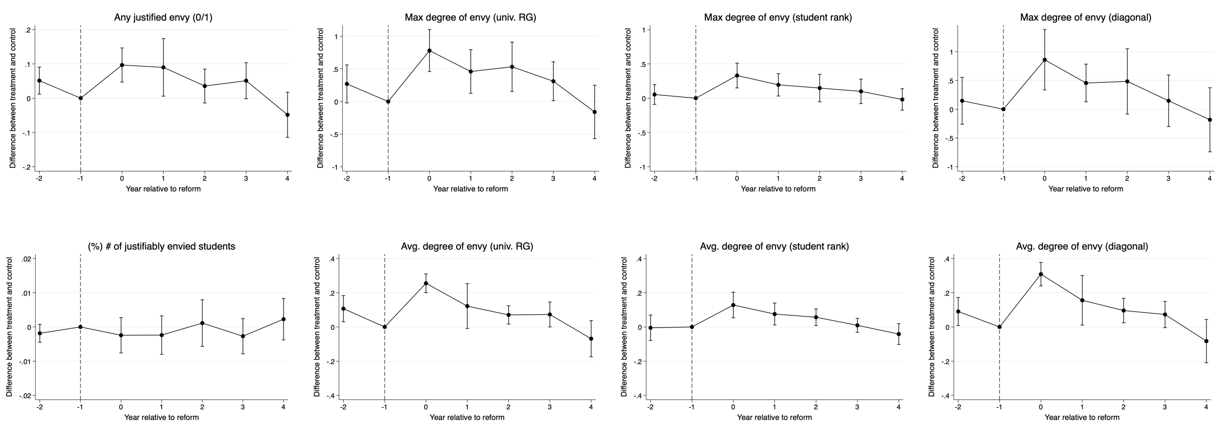


(a) Any Justified Envy (0/1)

(b) Number of Envied Students (%)

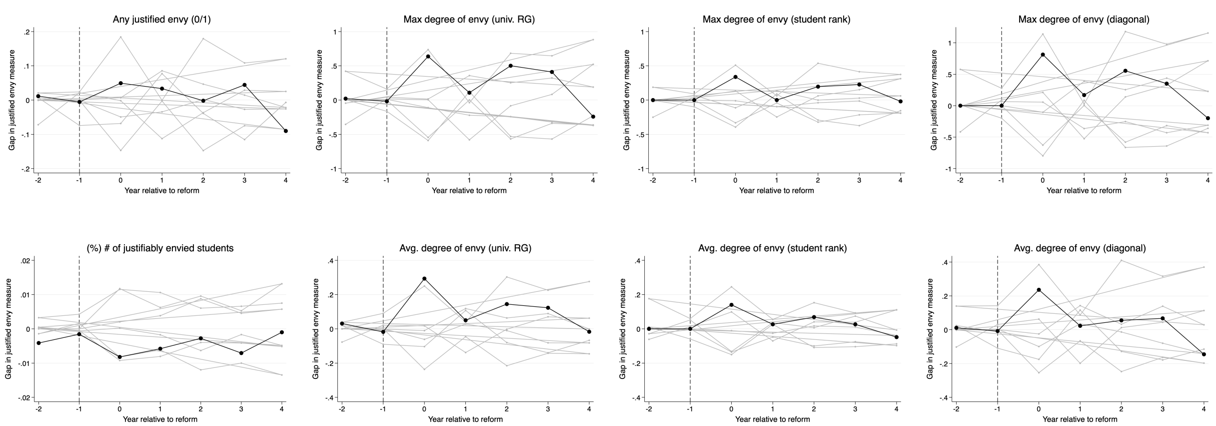
Notes: This figure shows the unadjusted justified envy outcomes in Inner Mongolia and other provinces in the control group beween 2005 and 2011. Subfigure (a) presents the unadjusted average probability of justifiably envying any other applicant. Subfigure (b) presents the unadjusted average number of justifiably envied applicants, as a proportion of all competing applicants.

Figure 3: Event-Study Estimates of the Effects of Inner Mongolia Reform



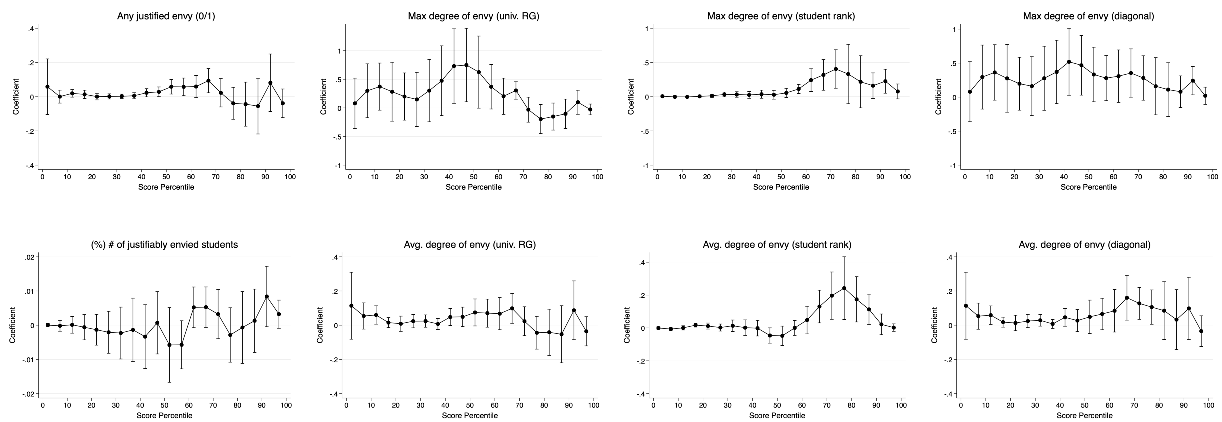
Notes: This figure presents the event-study results from Equation 2 on all eight JE measures. Each subfigure plots the point estimates of the difference between Inner Mongolia and the control provinces and their corresponding 95% confidence intervals. We set 2007 as first treatment year, i.e. year 0.

Figure 4: Synthetic Control Estimates of the Effects of Inner Mongolia Reform



Notes: This figure presents the effects of switching from IA to the real-time dynamic mechanism from the synthetic control design. The black line in each subfigure represents the gap in the corresponding JE measure between Inner Mongolia and the synthetic Inner Mongolia. The gray lines show the placebo tests, where we treat each of the control provinces as the reform province and compute the estimated effect of the hypothetical reform.

Figure 5: Effects on JE Measures by Test Score Distribution



Notes: This figure presents the effects of Inner Mongolia reform on JE measures at different test score percentiles. Applicants in each year-province-track cell are partitioned into 20 groups based on their test score percentiles. Each subfigure plots the estimated coefficients for the corresponding group and their 95% confidence intervals.

Table 1: Justified Envy Measures

JE Measures	Description	Graphical Representation (An Example with Applicant D in Figure 1)	Value
Any Justified Envy (0/1)	An indicator variable for whether an applicant justifiably envies at least one other applicant.	Whether there is a point to the bottom-right of D.	1
(%) Number of Justifiably Envied Students	The number of justifiably envied applicants as a proportion of the total number of competing applicants.	The proportion of points to the bottom-right of D.	$\frac{3}{7}$
Max degree of JE (university rank group)	The maximum envy in university tiers between an applicant and all competing applicants.	$\max\left\{\lvert DE\rvert_y, \lvert DF\rvert_y, \lvert DG\rvert_y, 0\right\}$	3
Average degree of JE (university rank group)	The average envy in university tiers between an applicant and all competing applicants.	$\left(\lvert DE\rvert_y + \lvert DF\rvert_y + \lvert DG\rvert_y\right)/7$	$\frac{5}{7}$
Max degree of JE (student rank)	The maximum envy in student ranking between an applicant and all competing applicants.	$\max\left\{\lvert DE\rvert_x, \lvert DF\rvert_x, \lvert DG\rvert_x, 0\right\}$	2
Average degree of JE (student rank)	The average envy in student ranking between an applicant and all competing applicants.	$\left(\lvert DE\rvert_x + \lvert DF\rvert_x + \lvert DG\rvert_x\right)/7$	$\frac{4}{7}$
Max degree of JE (diagonal)	The maximum envy, combining both the university tier and the student ranking dimensions, between an applicant and all competing applicants.	$\max\left\{\lvert DE\rvert, \lvert DF\rvert, \lvert DG\rvert, 0\right\}$	$\sqrt{10}$
Average degree of JE (diagonal)	The average envy, combining both the university tier and the student ranking dimensions, between an applicant and all competing applicants.	$\left(\lvert DE\rvert + \lvert DF\rvert + \lvert DG\rvert\right)/7$	$\frac{\sqrt{10}+\sqrt{5}+\sqrt{2}}{7}$

[*] *Notes:* This table reports all justified envy measures used in our paper. For each measure, one example using an applicant in Figure 1 is provided. In the table, $|\cdot|$ denotes length, $|\cdot|_x$ denotes distance in the horizontal axis, $|\cdot|_y$ denotes distance in the vertical axis. These measures are developed by making slight adjustments to those in Kang et al. (2020). We replace the level value of the number of justifiably envied students with a fraction. We also use the average degree of justified envy in all three dimensions (university, student, and these two combined) instead of the summation of degree of justified envy. These adjustments allow us to compare all justified envy measures across regions with different sizes of the pool of applicants.

Table 2: Summary Statistics

	Inner Mongolia		Control Provinces	
Sample size: 887,615	Pre-2007 (1)	Post-2007 (2)	Pre-2007 (3)	Post-2007 (4)
Individual Characteristics				
Percentile rank	49.4	49.4	49.3	49.3
Ex-ante submission	0.000	0.000	0.183	0.152
Ex-interim submission	0.000	0.000	0.362	0.300
Ex-post submission	1.000	1.000	0.456	0.548
Stem track	0.811	0.811	0.817	0.825
Admitted to RG1 university	0.067	0.080	0.064	0.076
Admitted to RG2 university	0.208	0.221	0.222	0.227
Admitted to RG3 university	0.286	0.275	0.284	0.305
Admitted to RG4 university	0.439	0.424	0.431	0.392
JE Outcomes				
Any justified envy (0/1)	0.853	0.839	0.871	0.838
	(0.354)	(0.367)	(0.335)	(0.369)
(%) # of justifiably envied students	0.022	0.018	0.032	0.027
	(0.042)	(0.034)	(0.063)	(0.053)
Max degree of envy (univ. RG)	1.407	1.627	1.647	1.623
	(0.776)	(0.968)	(0.925)	(0.993)
Avg. degree of envy (univ. RG)	0.922	0.919	0.957	0.924
	(0.391)	(0.415)	(0.389)	(0.423)
Max degree of envy (student rank)	0.856	0.951	0.986	0.963
	(0.686)	(0.751)	(0.744)	(0.779)
Avg. degree of envy (student rank)	0.310	0.322	0.377	0.349
	(0.205)	(0.225)	(0.287)	(0.254)
Max degree of envy (diagonal)	1.649	1.945	1.905	1.939
	(0.830)	(1.014)	(0.949)	(1.058)
Avg. degree of envy (diagonal)	1.009	1.016	1.082	1.035
	(0.428)	(0.459)	(0.447)	(0.476)
Observations	19,393	70,497	186,612	611,113

* *Notes:* This table reports the summary statistics of individual-level *Gaokao* and college admission outcomes from 2005 to 2011. Post-2007 includes years 2007 through 2011. The sample includes Inner Mongolia and six other municipalities and provinces that remained as using IA from 2005 to 2011 (Beijing, Shanxi, Heilongjiang, Shandong, Qinghai, Gansu). The rank and justified envy outcomes are computed at the year-province-track level. Four university rank groups (RGs) are constructed, and universities are assigned to the RGs based on the average admission cutoff score percentile for the university within each year-province-track cell. Details of the construction of rank groups and justified envy measures are discussed in the data section.

Table 3: Effects of Adopting Real-time Dynamic Mechanism (versus IA)

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dynamic	0.018	0.000	0.239*	0.033	0.119**	0.045***	0.262*	0.059**
	(0.015)	(0.003)	(0.118)	(0.020)	(0.045)	(0.014)	(0.147)	(0.022)
in SD	[0.051]	[0.002]	[0.246]	[0.079]	[0.154]	[0.175]	[0.255]	[0.126]
Dep. Mean	0.845	0.027	1.624	0.931	0.965	0.352	1.926	1.043
Dep. SD	(0.362)	(0.054)	(0.973)	(0.415)	(0.768)	(0.259)	(1.029)	(0.468)
Observations	887,615	887,615	887,615	887,615	887,615	887,615	887,615	887,615
R-squared	0.34	0.09	0.48	0.41	0.45	0.50	0.43	0.42

* *Notes:* This table reports the main difference-in-differences regression results. The regression model includes year, province, track, student percentile rank fixed effects, application pool size, and preference submission timing controls. Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Results from Using Other University Rank Group Assignments

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 4 RGs, t-1 (Obs.=750,835)								
Dynamic	-0.007	-0.004	-0.094	-0.032	0.038	-0.025	-0.093	-0.044
	(0.021)	(0.003)	(0.075)	(0.039)	(0.055)	(0.019)	(0.097)	(0.042)
R-squared	0.35	0.12	0.51	0.44	0.54	0.51	0.41	0.42
Panel B: 4 RGs, width=25% (Obs.=887,615)								
Dynamic	-0.026*	-0.001	0.026	-0.004	0.010	-0.025*	0.067	-0.012
	(0.014)	(0.002)	(0.039)	(0.023)	(0.032)	(0.014)	(0.058)	(0.028)
R-squared	0.52	0.08	0.57	0.55	0.40	0.51	0.45	0.53
Panel C: 8 RGs (Obs.=887,615)								
Dynamic	0.015	-0.004	0.339	0.058	0.263***	0.049	0.525*	0.089
	(0.013)	(0.003)	(0.198)	(0.041)	(0.083)	(0.028)	(0.290)	(0.051)
R-squared	0.38	0.09	0.50	0.33	0.49	0.51	0.42	0.38
Panel D: 8 RGs, t-1 (Obs.=750,835)								
Dynamic	0.002	-0.007*	0.136	-0.078	0.191*	-0.041	0.134	-0.094
	(0.017)	(0.004)	(0.149)	(0.059)	(0.090)	(0.036)	(0.154)	(0.072)
R-squared	0.36	0.12	0.52	0.33	0.55	0.51	0.40	0.38
Panel E: 8 RGs, width=12.5% (Obs.=887,615)								
Dynamic	0.012	-0.002	0.239*	-0.065	0.257***	0.044	0.476**	-0.018
	(0.012)	(0.002)	(0.117)	(0.054)	(0.081)	(0.027)	(0.188)	(0.052)
R-squared	0.40	0.09	0.56	0.48	0.49	0.52	0.41	0.46

* *Notes:* This table reports the difference-in-differences results with alternative university rank group assignments. In the main regressions, we assign the top 10%, top 10% to 30%, top 30% to 60%, and the rest of schools to RG-1 through RG-4, respectively. Panel A shows estimates where we rank schools based on the *Gaokao* scores of admitted students from the previous year. Panel B assigns 25% of schools to each rank group based on the ordered performance of admitted students from the current year. Panel C assigns universities into eight groups instead of four groups based on the following cutoff percentiles: 5%, 10%, 20%, 30%, 45%, 60%, 80%. Panels D and E are similar to panels A and B except that there are now eight groups. Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Results from Using A Looser Justified Envy Definition

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 5 Percentile (Obs.=887,615)								
Dynamic	0.015	-0.001	0.235*	0.028*	0.127**	0.046***	0.268	0.055***
	(0.009)	(0.003)	(0.116)	(0.014)	(0.055)	(0.014)	(0.156)	(0.016)
R-squared	0.34	0.09	0.47	0.41	0.49	0.53	0.41	0.42
Panel B: No Restriction (Obs.=887,615)								
Dynamic	0.015*	-0.002	0.236*	0.028**	0.138**	0.044***	0.275	0.054***
	(0.007)	(0.003)	(0.116)	(0.011)	(0.063)	(0.014)	(0.162)	(0.015)
R-squared	0.35	0.10	0.47	0.41	0.50	0.54	0.40	0.43

* *Notes:* This table reports the difference-in-differences results with alternative justified envy definitions. In the main regressions, we define that students have homogeneous preferences over schools in different RGs with at least 10 percentile distance apart. Here, we loosen the percentile distance restriction to 5 percentile, or no restriction. Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Robustness Checks

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Drop year 2007 (Obs.=773,651)								
Dynamic	0.008	0.001	0.149	-0.002	0.076*	0.025**	0.142	0.016
	(0.016)	(0.003)	(0.111)	(0.021)	(0.041)	(0.012)	(0.134)	(0.022)
R-squared	0.35	0.09	0.48	0.42	0.45	0.49	0.42	0.42
Panel B: Drop students matched to local schools (Obs.=752,443)								
Dynamic	0.016	0.004	0.268**	0.037	0.138***	0.033*	0.294*	0.055*
	(0.017)	(0.003)	(0.111)	(0.022)	(0.043)	(0.019)	(0.143)	(0.027)
R-squared	0.38	0.12	0.49	0.44	0.50	0.53	0.45	0.44
Panel C: Drop the best and worst major(s) in each school (Obs.=764,999)								
Dynamic	0.054**	-0.002	0.145	0.044	0.045	0.005	0.056	0.040
	(0.020)	(0.002)	(0.155)	(0.025)	(0.097)	(0.032)	(0.181)	(0.037)
R-squared	0.28	0.09	0.44	0.37	0.40	0.46	0.40	0.38
Panel D: Build RGs based on school-major combinations (Obs.=887,615)								
Dynamic	0.002	-0.001	0.099	0.019	0.094*	0.051***	0.239**	0.047
	(0.036)	(0.002)	(0.082)	(0.040)	(0.050)	(0.010)	(0.094)	(0.039)
R-squared	0.42	0.06	0.46	0.46	0.49	0.60	0.49	0.49

* *Notes:* This table reports the results from the robustness checks. Panel A shows results from dropping observations from 2007. In Panel B we drop all student observations who were matched to a school in their own province. Panel C presents results from dropping students admitted to the best and worst majors in each school. In Panel D we build rank groups with the same criteria as in the main results, but assign school-major pairs rather than schools to the rank groups. Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Heterogeneous Effects

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Severely Impoverished Counties								
Dynamic	0.030***	0.000	0.357**	0.045*	0.147**	0.045	0.439**	0.067*
	(0.004)	(0.003)	(0.132)	(0.021)	(0.065)	(0.030)	(0.171)	(0.037)
Observations	96,184	96,184	96,184	96,184	96,184	96,184	96,184	96,184
R-squared	0.36	0.10	0.43	0.38	0.50	0.46	0.39	0.37
Panel B: Other Counties								
Dynamic	0.011	0.000	0.207	0.024	0.091*	0.037**	0.200	0.048*
	(0.010)	(0.002)	(0.135)	(0.018)	(0.048)	(0.016)	(0.163)	(0.023)
Observations	592,784	592,784	592,784	592,784	592,784	592,784	592,784	592,784
R-squared	0.34	0.09	0.50	0.43	0.48	0.52	0.45	0.43

* *Notes:* This table reports the heterogeneous effects of the switch to real-time dynamic mechanism. Panel A presents the results for severely impoverished counties, and Panel B presents the results for all other counties. Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effects of Adopting Parallel Mechanism (versus IA)

	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Excluding Inner Mongolia (Obs.=3,574,195)								
Parallel	-0.051**	-0.013***	-0.438***	-0.084***	-0.191***	-0.067***	-0.514***	-0.111***
	(0.020)	(0.001)	(0.086)	(0.024)	(0.045)	(0.020)	(0.102)	(0.030)
in SD	[-0.129]	[-0.257]	[-0.456]	[-0.185]	[-0.258]	[-0.240]	[-0.497]	[-0.217]
Dep. Mean	0.804	0.022	1.464	0.884	0.869	0.325	1.739	0.986
Dep. SD	(0.397)	(0.050)	(0.960)	(0.454)	(0.741)	(0.278)	(1.034)	(0.513)
R-squared	0.39	0.10	0.47	0.42	0.42	0.40	0.41	0.40
Panel B: Full Sample (Obs.=3,664,085)								
Dynamic	0.010	-0.003**	0.265**	0.035**	0.124***	0.027**	0.309**	0.051***
	(0.010)	(0.001)	(0.103)	(0.013)	(0.036)	(0.013)	(0.122)	(0.014)
in SD	[0.026]	[-0.065]	[0.276]	[0.077]	[0.168]	[0.097]	[0.300]	[0.100]
Parallel	-0.049**	-0.013***	-0.426***	-0.079***	-0.186***	-0.064***	-0.499***	-0.106***
	(0.019)	(0.001)	(0.086)	(0.024)	(0.045)	(0.020)	(0.103)	(0.031)
in SD	[-0.124]	[-0.261]	[-0.444]	[-0.175]	[-0.251]	[-0.231]	[-0.483]	[-0.207]
Dep. Mean	0.805	0.022	1.467	0.885	0.871	0.325	1.743	0.987
Dep. SD	(0.396)	(0.050)	(0.960)	(0.453)	(0.741)	(0.277)	(1.033)	(0.512)
R-squared	0.39	0.10	0.47	0.42	0.42	0.40	0.41	0.40

* *Notes:* Panel A of this table replicates the difference-in-differences estimates in Kang et al. (2020) (with the adjusted JE measures) using the model specified by Equation 1. Panel B of this table extends the results by including Inner Mongolia in the sample. The full data sample includes all Chinese *Gaokao* test-takers who were admitted to any college between 2005 and 2011 (and exclude Hunan province, which switched to the parallel mechanism before 2005). In Panel A we exclude the province of Inner Mongolia and compare provinces that switched to the parallel mechanism to provinces that remained using IA between 2005 and 2011. In Panel B we keep the full sample and estimate the effects of switching from IA to the real-time dynamic mechanism or the parallel mechanism on justified envy using a regression model of the following form: $Y_{itpk} = \alpha_t + \beta_1 Dynamic_{tp} + \beta_2 Parallel_{tp} + \gamma Timing_{tp} + \lambda Rank_i + \delta N_{tpk} + \mu_p + \eta_k + \varepsilon_{itpk}$. Standard errors are clustered at the province-track level. We report the estimates of coefficients β_1 and β_2 in this table. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

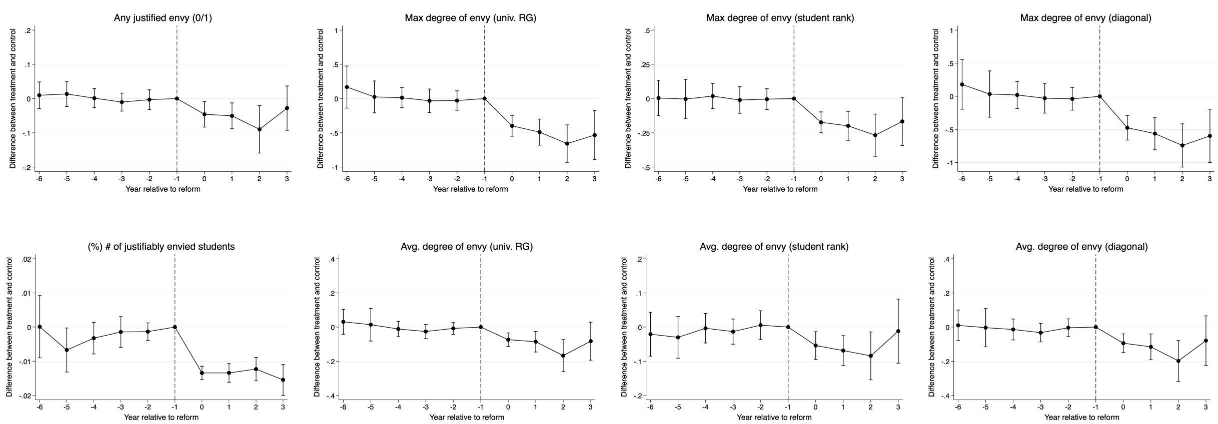
Table 9: Effects on Rank Range and Match Index

	Real-time Dynamic vs. IA		Parallel vs. IA		Both vs. IA	
	Rank Range (1)	Match Index (2)	Rank Range (3)	Match Index (4)	Rank Range (5)	Match Index (6)
Dynamic	42.39	0.28			380.93	0.37
	(1,183.73)	(0.67)			(853.98)	(0.57)
Parallel			-1,153.00***	-1.00***	-1,140.13***	-0.99***
			(313.58)	(0.20)	(308.66)	(0.20)
Observations	392	392	1,536	1,536	1,592	1,592
R-squared	0.83	0.41	0.83	0.35	0.83	0.35

* *Notes:* This table reports the difference-in-differences estimates of the effects of switching from IA to the real-time dynamic mechanism or the parallel mechanism on rank range and match index. The data are at the year-university-track-province level from 2005 to 2011. Columns 1 and 2 include Inner Mongolia and provinces that remained using IA. Columns 3 and 4 include provinces that switched to parallel mechanism and provinces that remained using IA. Columns 5 and 6 include all provinces (excluding Hunan). Standard errors are clustered at the province-track level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

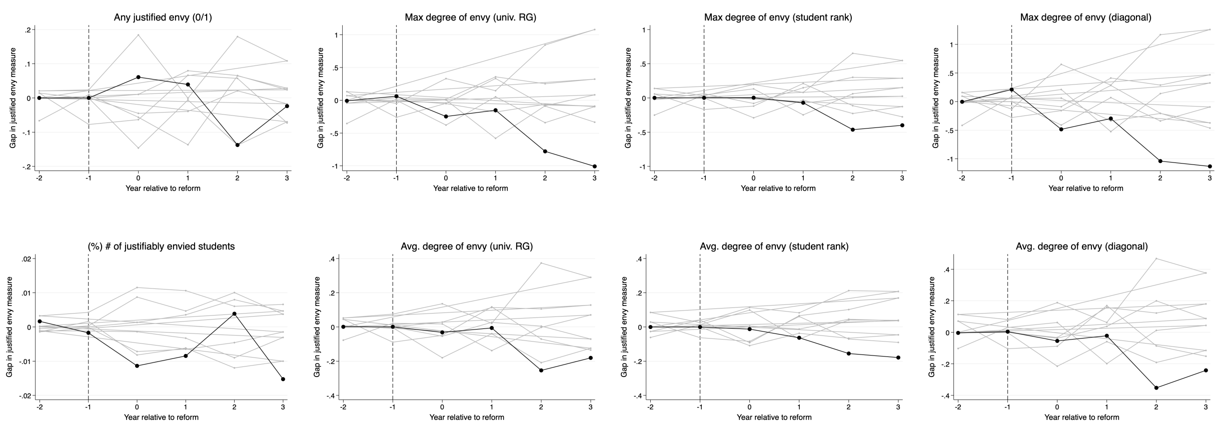
A Appendix: Supplementary Figures and Tables

Figure A.1: Event-Study Estimates of the Effects of Implementing Parallel Mechanism



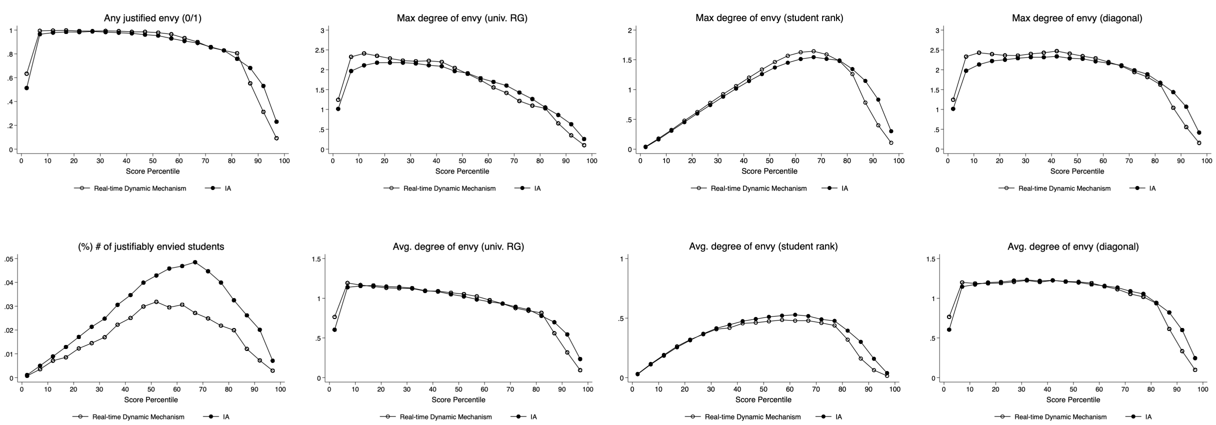
Notes: This figure reports the event-study estimates in Kang et al. (2020) using the model specified by Equation 2. Each subfigure plots the point estimates of the difference between provinces that switched from IA to the parallel mechanism and provinces that kept using IA, and their corresponding 95% confidence intervals. Difference provinces could have different treatment time, as specified by Table A3.

Figure A.2: Synthetic Control Estimates of the Effects of Implementing Parallel Mechanism



Notes: This figure presents the effects of switching from IA to the parallel mechanism from the synthetic control design. The black line in each subfigure represents the gap in the corresponding JE measure between Zhejiang (one of the provinces that switched from IA to the parallel mechanism) and the synthetic Zhejiang. Zhejiang reformed its matching mechanism in 2007 and data for Zhejiang province is available from 2005 to 2010. The gray lines show the placebo tests, where we treat each of the control provinces as the reform province and compute the estimated effect of the hypothetical reform.

Figure A.3: Justified Envy by Test Score Distribution



Notes: This figure presents the average JE measures along the test score distribution for the real-time dynamic mechanism and the immediate acceptance (IA) mechanism. Note that IA consists of the control provinces in all years and Inner Mongolia before 2007. Applicants in each year-province-track cell are partitioned into 20 groups based on their test score percentiles.

Table A1: Preference Submission Timing

Province	2005	2006	2007	2008	2009	2010	2011
Beijing	Ex-ante	Ex-ante	Ex-ante	Ex-ante	Ex-ante	Ex-ante	Ex-ante
Shanxi	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim
Inner Mongolia	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post
Heilongjiang	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim	Ex-interim
Shandong	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post
Gansu	Ex-interim	Ex-interim	Ex-interim	Ex-post	Ex-post	Ex-post	Ex-post
Qinghai	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post

* *Notes:* This table reports the timing of college preference submission period, i.e. the day(s) at which all applicants submitted their college applications. Ex-ante indicates that the preference submission date(s) take place before the first day of *Gaokao*. Ex-interim indicates that the preference submission date(s) take place after *Gaokao*, but before the scores are released. Ex-post indicates that the preference submission date(s) take place after the *Gaokao* scores are released for everyone.

Table A2: Province Weights in the Synthetic Inner Mongolia

Province/Municipality	Weights							
	Any Justified Envy	(%) Number of Envied Students	Degree of Justified Envy					
			University RG		Student Rank		Diagonal	
			Max	Avg.	Max	Avg.	Max	Avg.
Beijing	0.488	0	0.742	0.387	0.696	0.459	0.682	0.622
Shanxi	0	0	0	0.613	0.036	0.12	0.029	0
Heilongjiang	0	0	0	0	0.024	0.116	0.013	0
Shandong	0.512	1	0.258	0	0.205	0.124	0.248	0.378
Gansu	0	0	0	0	0.021	0.08	0.014	0
Qinghai	0	0	0	0	0.018	0.101	0.014	0

* *Notes:* This table reports the weights of each control province in the synthetic Inner Mongolia. The weights are chosen such that the resulting synthetic Inner Mongolia best reproduces the path of the JE measures in Inner Mongolia before the implementation of the real-time dynamic mechanism. Each column reports the weights used to reproduce the corresponding JE outcome.

Table A3: Timing of Policy Reform in Each Province

Province/Municipality	Preference Submission		Matching Mechanism
	Ex-interim	Ex-post	Parallel
Beijing	-	2015	2014
Tianjin	before 2005	2011	2010
Hebei	-	1999	2009
Shanxi	before 2005	2012	2012
Inner Mongolia	-	2002	-
Liaoning	before 2005	2014	2008
Jilin	-	2008	2009
Heilongjiang	before 2005	2013	2013
Shanghai	-	2017	2008
Jiangsu	-	2003	2005
Zhejiang	-	1999	2007
Anhui	before 2005	2007	2008
Fujian	-	2005	2009
Jiangxi	before 2005	2007	2009
Shandong	-	1998	2013
Henan	before 2005	2010	2010
Hubei	-	2004	2011
Hunan	-	2001	2003
Guangdong	-	2008	2010
Guangxi	-	2004	2009
Hainan	-	2002	2009
Chongqing	-	2006	2010
Sichuan	-	2005	2009
Guizhou	before 2005	2008	2009
Yunnan	-	2004	2009
Tibet	-	1996	2010
Shaanxi	before 2005	2010	2010
Gansu	before 2005	2008	2015
Qinghai	-	1996	-
Ningxia	-	1999	2009
Xinjiang	before 2005	2015	2011

* *Notes:* Source: Ha, Kang and Song (2020). This table shows the timing of policy reform in each province. All provinces started with ex-ante preference submission and IA mechanism. Some provinces phased into the ex-post submission by adopting the ex-interim submission first (before 2005). "-" indicates that this province never adopted this type of preference submission timing or matching mechanism. Hunan province is excluded from the sample since it switched to the parallel mechanism before 2005.

Table A4: Selection of Applicant Groups

Group	2007	2008	2009	2010	2011	2016	2021
Group 1	l-750	640-750	610-750	660-750	640-750	670-750	670-750
Group 2		580-639	550-609	630-659	610-639	640-669	640-669
Group 3		560-579	500-549	600-629	580-609	610-639	610-639
Group 4		l-559	l-499	570-599	550-579	580-609	580-609
Group 5				540-569	520-549	550-579	550-579
Group 6				510-539	490-519	520-549	520-549
Group 7				l-509	l-489	490-519	490-519
Group 8						460-489	460-489
Group 9						l-459	430-459
Group 10							l-429

* *Notes:* This table reports the selection of applicant groups in Inner Mongolia in each year. College applicants with scores that fall into a specific interval were assigned to the corresponding group. As we discussed in Section 2.2, each group had a different preference submission deadline. Groups with high-scorers would have an earlier deadline than groups with low-scorers. "l" represents the lowest possible score admitted to any Tier-1 university in each year.