An Empirical Evaluation of Chinese College Admissions Reforms Through A Natural Experiment

Yan Chen[a,b,1], **Ming Jiang**[c,1], **and Onur Kesten**[d]

[a]School of Information, University of Michigan, 105 South State Street, Ann Arbor, MI 48109-2112, USA; orcid.org/0000-0003-1287-4525; [b]Department of Economics, School of Economics and Management, Tsinghua University, Beijing, 100084, China; [c]Antai College of Economics and Management, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai 200030, China; orcid.org/0000-0001-7402-1072; [d]School of Economics, Level 5, Social Sciences Building, The University of Sydney, NSW 2006, Australia; orcid.org/0000-0002-9684-4093

This manuscript was compiled on October 1, 2020

College admissions policies affect the educational experiences and labor market outcomes for millions of students each year. In China alone, ten million high school seniors participate in the National College Entrance Exam to compete for seven million seats at various universities each year, making this system the largest centralized matching market in the world. The last twenty years have witnessed radical reforms in the Chinese college admissions system, with many provinces moving from a sequential (immediate acceptance) mechanism to some version of the parallel college admissions mechanism, a hybrid between the immediate and deferred acceptance mechanisms. In this study, we use a natural experiment to evaluate the effectiveness of the sequential and parallel mechanisms in motivating student college ranking strategies and providing stable matching outcomes. Using a unique data set from a province that implemented a partial reform between 2008 and 2009, we find that students list more colleges in their rank-ordered lists, and more prestigious colleges as their top choices, after the province adopts the parallel mechanism in its Tier 1 college admissions process. These listing strategies in turn lead to greater stability in matching outcomes, consistent with our theoretical prediction that the parallel mechanism is less manipulable and more stable than the sequential mechanism.

college admissions | market design | natural experiment | stability

Since the 1990s, economic research has played an increasingly important role in the practical design of market institutions, including auctions for spectrums, electricity, and other commodities (1, 2); tradable permit systems for pollution abatement and other environmental regulations (3); labor market clearinghouses (4–7); formal procedures for student assignments to public schools or colleges (8–10); centralized systems for the allocation of organs (11); and other related matching and trading processes (12). In many of these cases, the insights drawn from theoretical, experimental, and empirical research have complemented each other in influencing market design choices.

Our study provides additional insight for the design of markets, specifically college admissions processes, obtained from a natural experiment to evaluate centralized matching procedures for student assignments to colleges. The college assignment process has a significant impact on the student educational experiences as well as on broader labor market outcomes in countries that use a centralized college admissions system based on standardized test scores. These countries include Australia, Chile (13), China (14), Germany (15–18), Greece, Hungary, Ireland, Russia, Spain, Turkey (19), and the United Kingdom.

Our study focuses on China in particular, where standardized test scores have been used since 1952 to match students to colleges via a centralized system. The National College Entrance Examination, also known as *gaokao*, forms the foundation of the Chinese college admissions system. Each year, roughly ten million high school seniors compete for seven million seats at various universities in China, making this system the largest centralized matching market in the world (14). Given the extent and importance of the Chinese admissions process, it is important to understand how the choice of an admissions mechanism impacts assignment outcomes.

The centralized college admissions problem (19) has several unique properties compared to other matching problems such as school choice (8). One major differentiator is that students' priorities in college admissions are usually determined by their test scores on a standardized college entrance exam, rather than their place of residence, as in school choice problems. Therefore, college priorities are by and large identical across all colleges. Moreover, the prestige of a college is a major concern for virtually all students, leading to a near-universal preference for top universities with national prestige. This

Significance Statement

While college admission decisions impact the educational experiences and labor market outcomes for millions of students each year, the best method to determine admissions has been vigorously debated by academics and policymakers across the globe. Driving this debate are a number of major theoretical and practical innovations over the past two decades. Using a natural experiment from China, we evaluate the performance of the immediate acceptance mechanism, used for many years for college admissions in China and school choice in the US, and the newer parallel mechanism. We find that when provinces move to the parallel mechanism, students apply to more colleges and list more prestigious colleges first. We further find that the student-college matching outcome becomes more stable.

Y.C., M.J. and O.K. designed research; M.J. analyzed data; and Y.C., M.J. and O.K. wrote the paper. Author names follow alphabetical order.

Y.C., M.J. and O.K. declare that they have no competing financial interests.

Data deposition: The data and analysis code reported in this paper are available at the open Inter-university Consortium for Political and Social Research (ICPSR) data repository (https://doi.org/10.3886/E121101V2).

[1]To whom correspondence should be addressed. E-mails: yanchen@umich.edu; mjiang@sjtu.edu.cn

universal criterion implies that student preferences are often highly correlated. As a result, college admissions are typically much more competitive than student allocations to schools within a district. These two factors raise the stakes in the college admissions process and potentially affect how students strategize under different mechanisms.

In the past two decades, the majority of Chinese provinces have moved from a "sequential" mechanism to various versions of a "parallel" mechanism in assigning students to universities. In applying these mechanisms, universities are divided into tiers according to their level of prestige. The sequential mechanism is a priority matching mechanism (20) executed sequentially across tiers of decreasing prestige. Within each tier, the Immediate Acceptance (IA) mechanism is applied, e.g., once the assignments in the first tier are finalized, the assignment process in the second tier starts, and so on. Despite its dominance in the admissions process until 2001, a pervasive criticism of the sequential mechanism* is that many high-scoring students often remain unassigned or end up being under-matched due to poor strategizing in providing their preferred college rankings (24).

To combat this issue, Chinese provinces more recently have moved to some version of a parallel mechanism (PA), where students are provided with choice-bands in which they can list several "parallel" colleges in decreasing desirability. Under PA, student applications are processed by these choice-bands, wherein each student is guaranteed to retain her score advantage for any college she lists within the same choice-band. This mechanism is perceived to alleviate the pressure experienced under IA by allowing students to aim for multiple colleges at the same time without the fear of losing their score advantage. For example, in Sichuan Province, where our dataset comes from, students can list up to five colleges within the same choice-band. Students can choose to allocate their choices across a mix of desirable-yet-risky and less-desirable-yet-safer options.

It is plausible to argue that the parallel mechanisms falls somewhere between the IA and the Deferred Acceptance (DA) mechanism (25). In a theoretical study of the Chinese college admissions reforms, Chen and Kesten (CK hereafter) (14) consider a parametric family of *application-rejection* mechanisms where each member is indexed by some positive number $e \in \{1, 2, \ldots, \infty\}$ of periodic choice-band sizes that allow the application and rejection process to continue before assignments are made permanent. In this family of mechanisms, as parameter e increases, one goes from IA ($e = 1$) to PA ($e \in [2, \infty)$), and then to DA ($e = \infty$). CK show that members of this family become "more manipulable" (26) and "less stable" (27) as one moves away from DA. While multiple equilibria may arise under any member of the family, their important insight is that it is students' first e choices that matter. We use these theoretical insights as a partial basis for our hypotheses in our natural experiment. Since the theoretical comparisons of IA and PA assume complete information and coordinated strategic play, it is important to test these predictions in the field to better gauge their policy implications.

Specifically, to complement and test the theoretical insights from CK, we study how different matching mechanisms in a centralized college admissions system affect students' preference ranking strategies and matching outcomes. Our study is based on a natural experiment that takes advantage of the move from IA to PA in Sichuan Province from 2008 to 2009. Since the year students participate in the college admissions process is mostly determined by their year of birth, well before the change in mechanism, students in our study do not self-select into the different mechanisms. This timing feature eliminates any concern about selection bias in our study. Moreover, since the mechanism change affects only a portion of the students in our experiment (Tier 1 students), we are able to use a difference-in-differences approach. Using this approach, we find that the results from our experiment confirm some of the theoretical predictions from CK. Lastly, our data set is truly unique, as we have the complete rank-ordered list (ROL) from each applicant in addition to each applicant's matching outcomes, whereas other empirical studies of the Chinese college admissions process do not have the ROL for each applicant (28–32). The ROL data enable us to make more precise inferences regarding students' strategic responses to the change in the matching mechanism.

In particular, we find that students list more colleges in their ROL under PA relative to IA. Of the top-listed colleges, we observe a 5% increase in the preference rankings for the most prestigious colleges. Overall, our results show that the added insurance of being able to designate some safe options increases the stability of our matching outcomes.

Related Literature

Our study makes several important contributions to the literature on matching markets. Within this literature, a common approach in testing matching mechanisms is to conduct a lab experiment. Doing so makes it possible to induce true preferences and thus accurately obtain various performance evaluations. Indeed, the school choice problem has been extensively studied using laboratory experiments that yield support for various mechanisms. For example, Chen and Sönmez (22) find that DA performs well in terms of truthful preference revelation, while Pais and Pintér (33) find that the Top Trading Cycles mechanism (TTC) is more efficient and less vulnerable to manipulation than either IA or DA in the school choice scenario. In experiments under the interim information condition, Featherstone and Niederle (34) find that incomplete information on the student side changes both mechanism efficiency and truthfulness, while Calsamiglia *et al.* (35) find that constraining students' ability to reveal their preferences leads to greater manipulation and lower efficiency. We refer the reader to a recent survey of the experimental literature on school choice and college admissions for further details (36).

Our paper contributes to the college admissions and broader matching literature by testing a common set of hypotheses using a natural experiment. The use of a field test provides higher external validity relative to laboratory experiments, since the latter is unable to capture the large scale and high stakes nature of the real-world college admissions process.

Empirical evaluations have been used to study the properties and performance of different matching mechanisms. For example, Mongell and Roth (37) study the "preferential bidding system" that matches students to sororities, and find that preference manipulation can prevent an unstable mechanism from unraveling. Braun *et al.* (15) study the centralized

*Such complaints are by now familiar from the school choice context where IA has come under extensive scrutiny due to its welfare and incentive shortcomings (8, 21–23).

college admissions in Germany, and find that high performing students who truth-tell due to a lack of understanding of the mechanism receive suboptimal placements. More recently, several empirically studies have taken a structural approach to examine the performance of matching mechanisms (38–41) and uncover true preferences from reported ROLs when the mechanism is not strategy-proof. In a related study using school choice data from Beijing, He (39) finds that teaching middle school parents to play the best response under IA may yield better outcomes than switching to DA. Another strand of empirical literature takes a more direct approach by using preference reports under strategy-proof mechanisms or surveys (42–45). In particular, Fack *et al.* (44) provide theoretical and empirical evidence showing that, assuming stability of the matching provides rich identifying information, while being a weaker assumption on student behavior, compared to assuming that students truthfully rank schools when applying for admission. The latter is corroborated by an online experiment using medical students immediately after their participation in the medical residence match which features a strategy-proof market design (46).

Finally, in the Chinese school choice and college admissions context, the college admissions mechanisms not only differ in their algorithm but also in the timing of students' preference submissions. Wu and Zhong (31) find that, under IA, better students are admitted to a top university when they submit their preferences before learning their test scores in the National College Entrance Exam, consistent with the theoretical prediction (47). Using lab experiments, Lien *et al.* (48) and Jiang (49) argue that requiring preference submissions before students take the exam can help correct the observed exam measurement error under IA. However, Pan (50) finds that pre-exam IA rewards overconfidence and creates more mismatches between students and schools. Comparing all three mechanisms in the Chinese school choice context in the laboratory, Chen and Kesten (51) find that PA is less manipulable and more stable than IA. Compared to Chen and Kesten (14, 51) who first characterize the Chinese college admissions mechanisms theoretically and then test them in the laboratory, we use a unique naturally occurring dataset to test their theoretical predictions surrounding the switch from IA to the new PA mechanism. In doing so, we are able to provide a clean body of support for their basic theoretical predictions.

Theory and Hypotheses

In this section, we introduce the college admissions problem, describe a family of mechanisms, and summarize the main theoretical results pertaining to this family. These theoretical results form the basis of our empirical evaluation.

We begin by defining the college admissions problem. Specifically, a *college admissions problem* (19) is a tuple (S, C, P_s, P_C), consisting of: (1) a set of students $S = \{s_1, ..., s_n\}$; (2) a set of colleges $C = \{c_1, ..., c_m\}$; (3) a capacity vector $q = (q_{c_1}, ..., q_{c_m})$ where q_{c_i} is the capacity of college c_i; (4) a list of student preferences $P_S = (P_{s_1}, ..., P_{s_n})$ where P_{s_i} is the strict preference relation of student s_i over colleges including the no-college option (with an unlimited quota); and (5) a list of college preferences $P_C = (P_{c_1}, ..., P_{c_m})$ where P_{c_i} is the strict preference relation of college c_i over a set of students, determined by students' scores on the centralized college entrance exam. Therefore, $P_{c_i} = P_{c_j}, \forall i, j \in \{1, ..., m\}$. A matching μ is an allocation of college slots (and the no-college option) to students such that the number of students assigned to any college does not exceed its quota.

A matching μ is *nonwasteful* if no student prefers a college that has an unfilled quota. A matching μ is *envy-free* if there is no student-college pair (c, s) such that student s prefers college c to the college she is assigned to, and college c prefers student s to at least one student who is assigned to it. A matching is *stable* if it is nonwasteful and envy-free. A matching is *Pareto efficient* if there is no other matching that makes all students as well off and at least one student better off.

The recent literature focuses on analyzing weaker properties than stability, such as (justified) envy-freeness, i.e., fairness. Wu and Roth (52) consider envy-free matchings in a many-to-one matching environment and show that the set of such matchings forms a lattice. In a similar vein, Kamada and Kojima (53), motivated by various distributional constraints, focus on finding fair matchings that are student-optimal and apply their results to the Japanese daycare market.

A college admissions mechanism, or simply a mechanism, is an algorithm that selects a matching for each problem. A mechanism is Pareto efficient (stable) if it always selects Pareto efficient (stable) matchings. A mechanism is *strategy-proof* if no student ever gains by misrepresenting his preferences.

Prior to 2001, the sequential, mechanism (or IA) was the prevalent college admissions mechanism in China. However, after 2001, a number of provinces began to adopt various versions of the parallel mechanism. By 2018, variants of PA had been adopted in all provinces. We next discuss an algorithm that describes a general family of mechanisms that nest IA, PA, and DA.

In the parametric *application-rejection algorithm* family, a member is indexed by a periodic choice-band size e that represents the number of choices the algorithm goes through when allocations are tentative before they become final.[†] In this mechanism, students first submit their complete ROL before the allocation process starts. The algorithm is described as follows.

Round t $\geq$0:

- Each unassigned student from the previous round applies to his $te + 1$-st choice college. Each college c considers its applicants. Those students with the highest score are tentatively assigned to college c up to its quota. The rest of the applicants are rejected.

In general,

- Each rejected student, who is yet to apply to his $te + e$-th choice college, applies to his next choice. If a student has been rejected from all his first $te + e$ choices, then he remains unassigned in this round and does not make any applications until the next round. Each college c considers its applicants. Those students with the highest score are tentatively assigned to college c up to its quota. The rest of the applicants are rejected.

- The round terminates whenever each student is either assigned to a college (including the no-college option) or

[†] Several provinces use asymmetric versions of this algorithm where the size of the choice-band also varies across rounds. See CK for further explanation of these variations as well as a historical account of the Chinese college admissions process.

is unassigned in this round, i.e., he has been rejected from all his first $te+e$ choices. At this point, all tentative assignments become final and the quota of each college is reduced by the number of students permanently assigned to the college.

The algorithm terminates when each student has been assigned to a college or has received the no-college option. At this point, all the tentative assignments become final. This family of mechanisms nests IA and DA as extreme cases, and PA as an intermediate case (14). Specifically, IA is obtained when $e = 1$, PA when $2 \leq e < \infty$, and DA when $e = \infty$. In this family, IA is the only Pareto efficient mechanism, whereas DA is the only stable or strategy-proof mechanism. In our study of college admissions in Sichuan Province, $e = 5$.

In their theoretical study, CK find that a move from one extreme mechanism to the other yields a trade-off in terms of strategic immunity and stability. At the individual strategy level, they show that, whenever any given member can be manipulated by a student, any member with a smaller e number can also be manipulated but not vice versa (Theorems 1 & 3). This implies that the PA mechanism used in Sichuan Province (where $e = 5$) is less manipulable than its predecessor, the IA mechanism. This leads to our first hypothesis:

Hypothesis 1 (Manipulability). Students will manipulate their preferences less under PA compared to IA.

In our field setting, although true preferences are not directly observable, we can infer preference manipulation through a number of patterns, such as listing a safe college as one's top choice, where a safe option may be a less prestigious college, or through the length of the submitted rank-ordered list. The theory in CK suggests that under IA, in equilibrium, the choices other than the top choice do not matter, whereas the first five choices matter under PA (for Sichuan). If students understand this observation, we expect to see a longer rank-ordered list under PA.

Continuing with the theoretical predictions of CK, they suggest that students under PA are able to list their equilibrium assignments under IA as a safety option while also listing their more desirable options higher up in their preference list, which yields an outcome at least as good as that under IA (Proposition 5). While the top choice is the most critical decision under IA, the first e choices are of utmost importance when determining final assignments under PA. This leads to our next hypothesis:

Hypothesis 2 (Insurance). Students will list more prestigious/more preferred colleges as their first choices under PA, compared to the IA mechanism.

In terms of choice accommodation, CK show that the IA mechanism is more generous in allocating students to their first choice than PA. This leads to the following hypothesis:

Hypothesis 3 (Choice Accommodation). IA will assign a higher number of students to their reported first choices than will PA.

In terms of stability, CK show theoretically that the PAs are more stable than the IAs they replace (Theorems 2 and 4). This leads to our final hypothesis:

Hypothesis 4 (Stability). PA will be more stable than IA.

Finally, we note that CK find no clear dominance of DA over PA, or PA over IA, due to the multiplicity of equilibria, even though the dominant strategy equilibrium outcome of DA Pareto dominates any equilibrium outcome of IA (21). Based on the predictions of CK, we are agnostic with regard to the efficiency comparison of the two mechanisms in our study.

Data and Empirical Methods

Our dataset consists of the college admissions data of a County in the Sichuan Province in southwestern China for the years 2008 and 2009. The County had a population of 1.47 million with 87% rural in 2008 and 2009, with a per capita GDP of USD 994 in 2008 and 1117 in 2009, below the national average of USD 3524 and 3828, respectively.[‡] For our study, we obtain the following student data: test score on the National College Entrance Exam, rank-ordered list of colleges, college admission outcome, and demographics. Compared to prior empirical studies of Chinese college admissions, our data set is unique in that we have each student's rank-ordered list.

Chinese colleges are categorized into tiers of decreasing prestige and quality. For example, Tier 1 colleges are generally considered better than Tier 2 colleges, etc. To determine college placement assignments, admissions mechanisms are executed sequentially across tiers. When assignments in the first tier are finalized, the assignment process in the second tier starts, and so on. Our dataset contains all students who participated in the Tier 1, Tier 2, and Tier 3 admissions process in 2008 and 2009.

For the period of our dataset, students first received their test scores and relative standings among all the students in the province and then completed their rank-ordered lists of colleges. The Provincial College Admissions Office determined whether a student was eligible to participate in the admissions of each tier by setting up an endogenously determined cutoff score, such that the number of students above the Tier 1 cutoff was approximately 120% of the total quota of all Tier 1 colleges; the number of students above the Tier 2 cutoff was approximately 120% of the total quota of all Tier 1 and Tier 2 colleges; etc.[§]

Additionally, there were two separate matching markets each year for the two academic tracks: humanities and social sciences (shortened as humanities henceforth), and science and engineering (shortened as STEM henceforth). Students self-select into one of the two tracks in their second year of high school, and subsequently prepare for and then take the corresponding set of exams. Likewise, each college has a separate quota for each of the two tracks.

Between the college entrance exams of 2008 and 2009, the government of Sichuan Province announced that it would change the college admissions mechanism from IA to PA for only its Tier 1 selection process. Since students participate in the college admissions process during their last year of high school, and the policy change was announced after the previous year's admission was complete, students were essentially selected into different treatment groups by birth. Thus, this context allows us to use the policy change as a natural experiment to study the effects of different matching mecha-

[‡] The national rural population was 53% and 52% for 2008 and 2009, respectively. Sources: National Bureau of Statistics; County Bureau of Statistics.

[§] See the online appendix of CK (14) for a detailed discussion of the Chinese college admissions process.

Table 1. Summary statistics

	2008				2009			
	Total	Female	Rural	STEM	Total	Female	Rural	STEM
Participated in Tier 1 admission	620	32.3%	80.2%	81.8%	768	36.3%	79.7%	80.7%
Participated in Tier 2 admission	2443	40.4%	80.8%	70.9%	2735	42.7%	83.1%	73.8%
Participated in Tier 3 admission	688	40.4%	75.0%	50.7%	605	48.3%	73.2%	57.4%
Participated in Tiers 1 and 2	122	43.4%	81.1%	77.9%	135	46.7%	81.5%	77.0%
Participated in Tiers 1, 2 and 3	2	100.0%	50.0%	100.0%	3	100.0%	100.0%	33.3%
Submitted Tier 1 ROL	717	30.54%	79.9%	82.0%	849	35.7%	79.5%	80.1%
Submitted Tier 2 ROL	2967	38.5%	80.7%	72.8%	3343	41.3%	82.7%	74.6%
Submitted Tier 3 ROL	876	49.2%	72.3%	52.6%	787	48.9%	72.7%	57.8%
Submitted Tiers 1 and 2 ROL	628	33.0%	80.6%	80.4%	723	37.2%	80.8%	78.2%
Submitted Tiers 1, 2 and 3 ROL	4	100.0%	25.0%	25.0%	7	71.4%	57.1%	28.6%

nisms on students' behavioral responses and college admissions outcomes.

Even though students are randomly selected into the different years by birth, we consider the possibility that there may be other differences across the two years, such as students' overall preferences for humanities versus STEM programs, that may impact our results. To address this possibility, we exploit the fact that only the Tier 1 mechanism changed from 2008 to 2009 in Sichuan Province, whereas the Tier 2 admissions mechanism remained the same. Therefore, we estimate the following difference-in-differences model:

$$y_i = \beta_0 + \beta_1 \cdot Y2009_i + \beta_2 \cdot Tier1_i + \beta_3 \cdot (Y2009_i \cdot Tier1_i) + \gamma \cdot \boldsymbol{X}_i + \epsilon_i,$$

where y_i is the outcome variable, measuring strategies or matching outcomes for each student. $Y2009_i$ and $Tier1_i$ are dummy variables that equal one for Year 2009 and Tier 1, respectively, and zero otherwise. The vector, $\boldsymbol{X}_i$, contains students' individual characteristics, including gender, residential status (rural or urban), academic track (humanities or STEM), and rank by test scores.

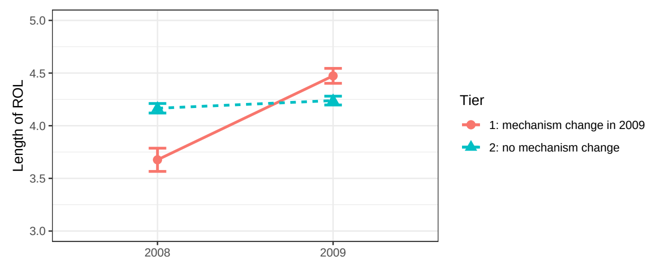
Table 1 presents the summary statistics for our dataset. These statistics show that students in different academic tracks and from different demographic backgrounds are similarly distributed across both years and tiers. Note that some students who were eligible for but did not receive a Tier 1 admission placement subsequently participated in the Tier 2 process. Students submitted their complete ROLs for all tiers at the same time, which was before the matching process was carried out, and no change was allowed once this process began.

Results

In this section, we first report our results regarding student strategies and then discuss our results regarding matching outcomes for the Tier 1 admissions process, using the Tier 2 process as our control. In the Supporting Information (SI Appendix henceforth), we use Tier 3 as the control condition as a robustness check.

Table 2 reports the summary statistics for the main outcome variables. At the individual strategy level, we investigate both the change in the number of colleges students rank (length of ROL), as well as the change in the prestige status of their top choice colleges. At the outcome level, we examine the

Fig. 1. Average length of rank-ordered lists (ROLs) across year and tier.



Notes: This figure reports the effect of changing the matching mechanism (Tier 1, red solid line with circles) on the average length of ROLs compared to the baseline with no mechanism change (Tier 2, green dashed line with triangles); error bars indicate the 95% confidence interval of the mean.

proportion of students admitted to their top choices, as well as the stability of the matching outcomes, using several measures to ensure robustness.

Student Preference Ranking Strategies. In the college admissions process, students within a given tier are asked to rank order anywhere from one to five colleges. Figure 1 reports the average length of the ROLs in 2008 and 2009 by tier. The red solid (green dashed) line refers to the length of Tier 1 (2) students' ROLs. From Figure 1, we see an increase in the ROL length for Tier 1 students from 2008 to 2009 by approximately one more college, whereas the average ROL length for Tier 2 students remains the same across two years.

Table 3 presents the results from nine OLS specifications. On the left panel, the dependent variable is the length of students' submitted ROLs. The independent variables (omitted) include Y2009 (Y2008), Tier1 (Tier2), Y2009 × Tier1, STEM (Humanities), Rural (Urban), Female (Male), and Percentile Ranking. To determine a student's percentile ranking, we calculate rankings in each of the eight markets based on student test scores on the respective National College Entrance Exams, as matching is carried out separately by year (2008/2009), tier (1/2), and academic track (humanities/STEM). To correct for different market sizes, we then normalize student rankings to their percentile rankings in their respective markets (the top ranked student in each market is 1 (100%), and the

Table 2. Summary statistics for outcome variables

	Tier 1		Tier 2		Tier 3	
	2008	2009	2008	2009	2008	2009
Length of rank-ordered list (1-5)	3.676	4.473	4.166	4.239	3.245	3.392
Top choice college prestige index (0-1)	0.519	0.458	0.403	0.368	0.458	0.406
First choice accommodation rate	0.740	0.501	0.673	0.672	0.638	0.536
Stability based on cutoff score	0.118	0.070	0.273	0.292	0.237	0.358
Stability based on college prestige	0.109	0.086	0.134	0.117	0.170	0.189
Stability based on score distance	0.494	0.374	0.521	0.508	0.328	0.374

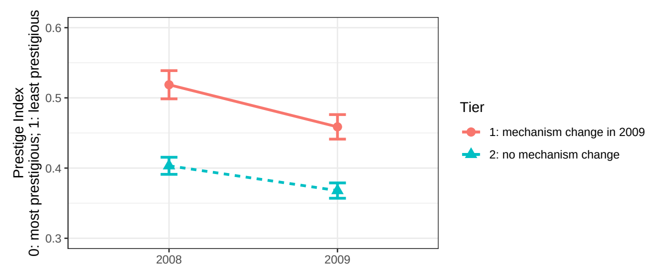
Notes: (i) The prestige index (0: most prestigious; 1: least prestigious) is calculated by ranking colleges based on the average scores of admitted students in year 2006 and 2007 from the best to the worst, within each STEM/humanities track, tier, and year bracket (a total of eight); the rankings are then normalized to 0 to 1 by dividing the rankings by the total number of colleges within each bracket; (ii) the first choice accommodation rate measures the percentage of students who are admitted by their first choice colleges within each tier; (iii) a matching is stable when there does not exist any student-college pair where both prefer each other to their current matches; (v) the measurement for stability based on cutoff score is described in the **Matching Outcomes** section in the main text, whereas the measurement for stability based on college prestige or score distance is relegated to SI; for each of the three measures, the larger the number the more unstable the matching outcome is.

bottom ranked 0). This measure of student rankings as our independent variable is used in all subsequent regressions.

The results in the left panel of Table 3 show a positive significant coefficient for our main treatment effect, Y2009 × Tier1, indicating that the change from IA to PA in the Tier 1 admissions process in 2009 increases the average ROL length by 0.724 ($p < 0.01$). That is, a Tier 1 student lists approximately one more college in 2009. Since the Tier 2 ROL length remains stable across the two years, the change in the length of the Tier 1 ROL is likely due to the change in the matching mechanism. Continuing with Table 3, we see that the coefficient for the Tier 1 dummy is negative and significant, indicating that the average length of the Tier 1 ROL is shorter than the corresponding Tier 2 ROL in 2008. This finding may reflect the importance of a student's first choice under IA, whereas lower-ranked colleges, such as a student's fourth or fifth choice, are not that useful under IA. Under PA, however, students have an incentive to include a less-prestigious college as their fifth choice as insurance. Finally, the results in column (3) in Table 3 show that higher ranked students as well as those in the STEM fields tend to submit shorter lists while women tend to submit longer lists.

Next, we investigate whether students list more prestigious colleges as their top choices under PA, as predicted by theory. We use two measures to compute our *prestige* index. First, we compute a local prestige index, using province-specific calculations. We rank colleges from most (1) to least (n) prestigious, as measured by the average scores of students within a tier or track market. We calculate these rankings separately for 2006 and 2007 and then average the two to obtain a final prestige score for each college. These ranks are then normalized to range from 0 (most prestigious) to 1 (least prestigious) within each of the eight markets. Since not all colleges that admitted students in 2008 and 2009 did so in 2006 and 2007, observations with these colleges as top choices (2.2%) are dropped from our analysis. Compared to alternative measures, our local prestige index utilizes the same data and statistics published and distributed to students and their parents by the Sichuan Educational Examination Authority in the *Gaokao Guide* (Sichuan: UESTC Press, 2009). While our prestige index is highly correlated with the published national rankings of colleges,[¶] using the average score of admitted students provides a more accurate aggregation of students' revealed preferences for colleges compared to national rankings, as students without complete preferences over colleges often use cutoff and average test scores to assess a given college's prestige.[‖]. Second, we use the national ranking of colleges as an alternative prestige measure, which has the advantage of being stable across years, even though it may not necessarily reflect the local preferences of students in Sichuan.

Fig. 2. Average local prestige index of first-choice college by year and tier.



Notes: This figure reports the effect of changing the matching mechanism (Tier 1, red solid line with circles) on the local prestige of students' first choice-colleges compared to the baseline with no mechanism change (Tier 2, green dashed line with triangles); the prestige index (0: most prestigious; 1: least prestigious) is calculated by ranking colleges based on the average scores of admitted students in year 2006 and 2007 from the best to the worst, within each STEM/humanities track, tier, and year bracket (a total of 8); the rankings are then normalized to 0 to 1 by dividing the rankings by the total number of colleges within each bracket; error bars indicate the 95% confidence interval of the mean.error bars indicate the 95% confidence interval of the mean.

Figure 2 presents the average local prestige index for students' first-choice colleges by year and tier, with 0 (1) indicating the most (least) prestigious college. From Figure 2, we see that, on average, students choose more prestigious colleges in both tiers in 2009, compared to the 2008 choices, with a more pronounced increase for Tier 1 students.

We next examine the effect of the change from IA to PA on the prestige level of students' first choices. In this analysis, the dependent variable is the local prestige level (from 0 to 1) of the student-reported top choice colleges (columns 4-6),

[¶] The correlation coefficient between the national ranking and our prestige index is 0.68 ($p < 0.001$, $n = 476$) for the STEM track, and 0.67 ($p < 0.001$, $n = 379$) for the humanities track. We use 2009 Chinese College Rankings published by the Chinese Alumni Network as to obtain our national rankings. This data is chosen because it is the most complete published rankings encompassing more than 500 colleges each year.

[‖] This is also the reason why top universities in China announce when they have a cutoff score higher than those of their rivals. See, e.g., https://cn.nytimes.com/education/20150701/c01sino-rivalry/en-us/. "If these students are taken by the competitors, then you'll be forced to lower your own cutoff score; once your cutoff score is lower than those of your competitors, you lose half of the battle of recruitment."

Table 3. Effects of matching mechanisms on the length of rank-ordered lists and the prestige of reported top choices (OLS)

Dependent variable:	Length of ROL			Local prestige index of top choices			National ranking of top choice colleges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Y2009	0.073*	0.073*	0.076	-0.035***	-0.037***	-0.038***	0.051***	0.052***	0.052***
	(0.041)	(0.041)	(0.049)	(0.011)	(0.012)	(0.012)	(0.003)	(0.003)	(0.004)
Tier1	-0.489**	-0.490**	-0.452**	0.115***	0.112***	0.109***	-0.191***	-0.200***	-0.198***
	(0.201)	(0.211)	(0.210)	(0.014)	(0.005)	(0.006)	(0.005)	(0.004)	(0.005)
Y2009 $\times$ Tier1	0.724***	0.724***	0.708***	-0.025**	-0.021*	-0.019	-0.044***	-0.046***	-0.047***
	(0.146)	(0.152)	(0.151)	(0.012)	(0.013)	(0.013)	(0.010)	(0.011)	(0.012)
Percentile Ranking		-0.784***	-0.776***		-0.722***	-0.721***		-0.294***	-0.293***
		(0.170)	(0.168)		(0.026)	(0.026)		(0.015)	(0.015)
STEM			-0.306***			0.024***			-0.022
			(0.033)			(0.005)			(0.011)
Rural			-0.053			0.043***			0.009
			(0.033)			(0.011)			(0.006)
Female			0.133***			-0.019***			0.009
			(0.024)			(0.007)			(0.009)
Constant	4.166***	4.557***	4.767***	0.403***	0.767***	0.722***	0.433***	0.587***	0.592***
	(0.038)	(0.083)	(0.097)	(0.029)	(0.020)	(0.029)	(0.013)	(0.011)	(0.009)
Observations	7,876	7,876	7,876	7,706	7,706	7,706	6,757	6,757	6,757
R-squared	0.021	0.053	0.070	0.021	0.449	0.455	0.117	0.220	0.223
Y2009 + Y2009 $\times$ Tier1	0.797***	0.797***	0.784***	-0.060***	-0.058***	-0.057***	0.006	0.006	0.005
	(0.181)	(0.188)	(0.194)	(0.009)	(0.006)	(0.006)	(0.009)	(0.008)	(0.008)
Tier1 + Y2009 $\times$ Tier1	0.235***	0.234***	0.257***	0.091***	0.091***	0.090***	-0.236***	-0.246***	-0.244***
	(0.086)	(0.087)	(0.089)	(0.017)	(0.013)	(0.012)	(0.011)	(0.011)	(0.011)

Notes: (i) Standard errors clustered at the high school level are in parentheses; *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. (ii) In the first model (Columns (1), (4) and (7)), the dependent variables are regressed on the year and tier dummies and their interactions using OLS. (iii) The second model (Columns (2), (5) and (8)) adds students' percentile rankings (0: lowest; 100%: highest) as control variable. (iv) The third model (Columns (3), (6), and (9)) further adds students' track and demographic information as additional control variables. (v) The (local) prestige index (0: most prestigious; 1: least prestigious) is calculated by ranking colleges based on the average scores of admitted students in year 2006 and 2007 from the best to the worst, within each STEM/humanities track, tier, and year bracket (a total of eight); the rankings are then normalized to 0 to 1 by dividing the rankings by the total number of colleges within each bracket. (vi) The national ranking (0: highest ranked; 1: lowest ranked) is calculated by putting colleges into bins based on their national rankings in year 2008 and 2009 (top 2 colleges, Peking and Tsinghua, are in bin 1, top 3 – 10 in bin 2, and every ten colleges in each subsequent bins) to account for correlated but heterogeneous preferences; then the bin numbers are normalized to [0,1] within each tier by dividing the numbers with total number of bins in that tier.

or the national ranking of top choice colleges (columns 7-9). The independent variables (omitted) again include Y2009 (Y2008), Tier1 (Tier2), Y2009 × Tier1, STEM (Humanities), Rural (Urban), Female (Male), and Percentile Ranking. The results on the right panel of Table 3 show a negative and significant coefficient for our main treatment effect, Y2009 × Tier1, indicating that students list more prestigious colleges as their Tier 1 first choices in 2009, with a magnitude of 2.5% ($p < 0.05$) using the local prestige index, and 4.4% ($p < 0.01$) using the national ranking of colleges. This result is consistent with the theoretical prediction that students are more likely to pick prestigious colleges as their top choice under PA since they are also able to include a safer choice in their ROL (Hypothesis 2). Using the local prestige index, we find that the effect becomes insignificant when control variables are added (column (6)). We further find that students' Tier 2 first choices in 2009 are ranked 3.5% higher than their corresponding rankings in 2008 ($p < 0.001$). It is not clear what drives this effect. Finally, the results in column (6) in Table 3 show that higher ranked students as well as women list more prestigious colleges as their Tier 1 first choices in 2009 (-0.72 and -0.019, respectively, $p < 0.001$). When we use Tier 3 as the control, the effect is also insignificant (see Table S6 in SI), indicating that the evidence is mixed. By contrast, using the national ranking as a measure of prestige, the treatment effect is robust to our model specifications.

We now summarize our treatment effect of the type of mechanism on student preference ranking strategies:

Result 1. Changing the Tier 1 admissions mechanism from IA to PA leads to an increase in the length of a student's rank-ordered list by approximately one more college, as well as a 4.4% increase in the national ranking of students' top-choice colleges.

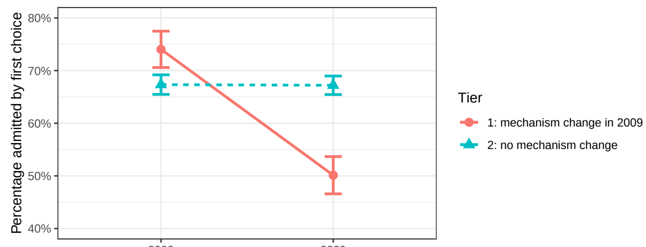
These empirical results are consistent with the theoretical prediction that students view the parallel mechanism (PA) as providing insurance or a fallback if they do not receive their ideal top choice, compared to the immediate acceptance mechanism (IA). Indeed, students appear to capitalize on the intuition that PA allows them to list more colleges and more prestigious colleges in the first tier without jeopardizing their admission chances to lower-ranked colleges within that tier.

Matching Outcomes. Next we investigate the effects of the type of mechanism and its subsequent behavioral changes on matching outcomes. First, we examine the effects of the type of mechanism change on the likelihood that a student is admitted by her reported first choice college. Recall that Hypothesis 3 predicts that IA will assign a larger number of students to their reported first choices than PA, as students have an incentive to aim higher under PA.

Figure 3 depicts the first choice accommodation rate by year and tier. As predicted by theory, we indeed see a drastic drop in the proportion of students admitted to their reported top-choice colleges in Tier 1 in 2009 (red solid line), in contrast to no change for Tier 2 admission rates (green dashed line). We next formally investigate this phenomenon through a regression analysis.

The left panel of Table 4 reports the results of our regression analysis using three probit specifications: the effects of the mechanism change on the likelihood of first choice accommodation (1), with student percentile ranking (2), and with demographic controls (3). The independent variables (omitted) again include Y2009 (Y2008), Tier1 (Tier2), Y2009 × Tier1, STEM (Humanities), Rural (Urban), Female (Male), and Percentile Ranking.

Fig. 3. First choice accommodation rate by year and tier.



Notes: This figure reports the effect of changing the matching mechanism (Tier 1, red solid line with circles) on the first choice accommodation rate compared to the baseline with no mechanism change (Tier 2, green dashed line with triangles); first choice accommodation rate measures the percentage of students who are admitted by their first choice colleges within each tier; error bars indicate the 95% confidence interval of the mean.

Consistent with our theoretical prediction (Hypothesis 3), we find that the coefficient for our main treatment effect, Y2009 × Tier1, is negative and significant, indicating that students are 24 percentage points ($p < 0.01$) less likely to be admitted by their reported top choices in the Tier 1 admissions process in 2009, whereas the likelihood of being admitted by first-choice colleges in Tier 2 in 2009 does not change compared to the previous year (-0.001, $p > 0.10$). Additionally, looking at the covariates, we find that students from rural areas are 3.9 percentage points more likely to be admitted into their reported first choices under PA. Finally, we see that students with a one-percentile increase in their entrance exam scores increase their likelihood of being admitted by their reported first choice by 0.558% under PA ($p < 0.001$). Since PA incentivizes students to aim high, we also find a decrease in the acceptance rate of top choice colleges after the change to PA.

In addition to examining the effect of the mechanism change on first-choice accommodation, we are interested in the performance of each mechanism in terms of matching stability. Recall that Hypothesis 4 predicts that PA will be more stable than IA. To measure stability, we first need to know students' preferences over colleges. In our study, we examine students' revealed preferences as indicated in their ROLs. This approach allows us to forego the assumption that students have identical preferences. With this measure, we assume that students preserve their preference order in their ROL, that is, they always list their more preferred colleges above their less preferred ones within the same choice band under PA, which is implied by Remark 3 in CK (14).

To identify unstable matchings, we consider an outcome to be unstable in two possible situations. First, an outcome is considered unstable if a student in Tier 1 has a listed college above her admitted college (within the same tier) whose cutoff score is lower than her test score, indicating justified envy. Second, an outcome is considered unstable if a student ends up in a Tier 2 college or lower even though her test score is high enough to obtain admission into one of her listed Tier 1 colleges. For Tier 2 observations, the first condition is the same, whereas the second condition changes to receiving admission

Table 4. Effects of matching mechanisms on first choice accommodation and stability (Probit)

Dependent variable:	Admitted to first choice			Unstable Matching		
	(1)	(2)	(3)	(4)	(5)	(6)
Y2009	-0.001	0.001	-0.001	0.019	0.017	0.018
	(0.023)	(0.021)	(0.022)	(0.024)	(0.023)	(0.023)
Tier1	0.067	0.077*	0.073*	-0.154***	-0.157***	-0.156***
	(0.052)	(0.039)	(0.039)	(0.022)	(0.020)	(0.020)
Y2009 × Tier1	-0.238***	-0.248***	-0.246***	-0.067***	-0.063***	-0.064***
	(0.050)	(0.047)	(0.048)	(0.028)	(0.025)	(0.026)
Percentile Ranking		0.558***	0.558***		-0.326***	-0.327***
		(0.017)	(0.016)		(0.025)	(0.023)
STEM			0.038***			-0.016**
			(0.014)			(0.008)
Rural			0.039***			-0.023**
			(0.012)			(0.009)
Female			0.005			-0.013
			(0.009)			(0.008)
Observations	6566	6566	6566	6300	6300	6300
Y2009 + Y2009 × Tier1	-0.239***	-0.247***	-0.247***	-0.048***	-0.046***	-0.046***
	(0.042)	(0.041)	(0.040)	(0.014)	(0.010)	(0.010)
Tier1 + Y2009 × Tier1	-0.171***	-0.171***	-0.173***	-0.221***	-0.219***	-0.220***
	(0.026)	(0.027)	(0.026)	(0.017)	(0.020)	(0.020)

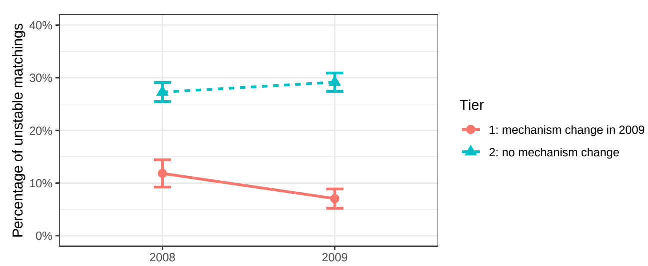
Notes: (i) Standard errors clustered at the high school level are in parentheses; *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Marginal effects are reported, calculated at the mean level of the covariates. (ii) In the first model (Columns (1) and (4)), dependent variables (for Columns 1-3, whether a student is admitted by the first-choice college and Columns 4-6, whether a matching is unstable; 0 = False, 1 = True) are regressed on the year and tier dummies and their interactions using a Probit model. (iii) The second model (Columns (2) and (5)) adds students' percentile rankings (0: lowest; 100%: highest) as a control variable. (iv) The third model (Columns (3) and (6)) further adds students' track and demographic information as additional control variables. (v) A matching outcome is considered unstable if a student in Tier 1(2) has a listed college above her admitted college (within the same tier) whose cutoff score is lower than her test score, or if she ends up in a Tier 2(3) college or lower even though her test score is high enough to obtain admission into one of her listed Tier 1(2) colleges.

to a college below Tier 2. While this approach ensures that all identified unstable matchings are truly unstable, it captures only a subset of all possible violations. For example, if the incentives of the IA mechanism lead a student to drop a highly desirable college from his list, violations of stability involving moving that student to an unlisted college are not detected. To address this issue, we use two alternative stability measures in SI Appendix. The first one uses college prestige as an approximation of students' preferences over colleges, which gives us an (almost) complete student preference profiles over colleges. The second one uses a "wasted" score, or a consequence of an unstable matching, as an indirect measure. We discuss the pros and cons of each measure in SI Appendix.

Using our main stability measure, we report the proportion of unstable matchings by year and tier in Figure 4. From 2008 to 2009, we see that the proportion of unstable matchings decreases for Tier 1 students (solid red line), whereas that for Tier 2 students remains almost constant (green dashed line). To examine this effect further, we next conduct a regression analysis on the same outcome variable.

The right panel of Table 4 (columns 4-6) reports the results of our regression analysis of the effects of the mechanism change on matching stability. The dependent variable here is whether the student's match is unstable. The independent variables (omitted) again include Y2009 (Y2008), Tier1 (Tier2), Y2009 × Tier1, STEM (Humanities), Rural (Urban), Female (Male), and Percentile Ranking. From the results in the table, we see that the coefficient for our main treatment effect, Y2009 × Tier1, is negative and significant, indicating that the move to PA decreases the number of unstable outcomes by 6.7 percentage points ($p < 0.01$).

Fig. 4. Proportion of unstable matching by year and tier.



Notes: This figure reports the effect of changing the matching mechanism (Tier 1, red solid line with circles) on matching stability compared to the baseline with no mechanism change (Tier 2, green dashed line with triangles); a matching outcome is considered unstable if a student in Tier 1(2) has a listed college above her admitted college (within the same tier) whose cutoff score is lower than her test score, or if she ends up in a Tier 2(3) college or lower even though her test score is high enough to obtain admission into one of her listed Tier 1(2) colleges; error bars indicate the 95% confidence interval of the mean.

We summarize our matching outcome analysis findings below:

Result 2. Changing the Tier 1 admissions mechanism from the IA to PA leads to a 24 percentage point decrease in the admissions students receive from their reported top-choice colleges, and a 6 percentage point decrease in the likelihood of unstable matchings.

Our observed first choice accommodation result is consistent with theoretical predictions (Hypothesis 3): students are indeed more focused on getting into their reported first choices under IA. The stability result is also consistent with theoretical predictions (Hypothesis 4) in that PA results in fewer unstable outcomes. This latter result is robust to different measures of stability, including a cardinal measure of the distance between a student's exam score and the cutoff score (see SI for details).

To provide greater confidence in our findings, we conduct a robustness test excluding the bottom 20% of Tier 1 students and the top 20% of Tier 2 students from our analysis. We do so to address the potential concern that the switch to PA in the Tier 1 process may impact the composition of students who participate in the Tier 2 process, as different mechanisms may leave different students unadmitted after the Tier 1 process concludes. Recall that, of the students with the highest scores, the computer algorithm considers 120% of the Tier 1 quotas for Tier 1 admissions, with an end goal of admitting the number of students equal to the Tier 1 quotas. This leaves 20% of the students rejected from the Tier 1 process. These students then enter the Tier 2 admissions process, and so on. This is important as our difference-in-differences estimates rely on the fact that the mechanism for Tier 2 does not change between 2008 and 2009. Excluding these students from our analyses yields similar results as those from our main analyses. Finally, we re-run our analyses using Tier 3 students as the control condition and find similar results except in the case of the local prestige index. SI Appendix summarizes the results from these robustness checks.

Conclusion

The assignment of students to colleges is one of the most important education policy issues throughout the world, with significant social welfare and economic development implications attached to the process. In China alone, ten million high school students participate in the college admission process each year. Since 2001, the process of allocating available slots to students has changed from the immediate acceptance mechanism to various versions of the parallel mechanism. While the parallel mechanism has been shown to have numerous benefits on a theoretical level (14), its benefits have been examined empirically mostly in a laboratory setting (51). By contrast, our study examines the effect of the parallel mechanism on student strategies and matching outcomes in a natural experiment using a unique data set with individual-level ranking strategies before and after the adoption of the new mechanism.

Specifically, we analyze a natural experiment using difference-in-differences estimators. Although some theoretical properties of matching mechanisms cannot be directly tested empirically due to the lack of students' true preferences, we can draw some analogies between the lab and the field using revealed preferences as seen in students' rank-ordered lists of their preferred colleges. We find that when the mechanism changes from IA to PA, students list better colleges as their first choices. We also find that students list more colleges in their rank-ordered lists under PA. These behavioral responses lead to more stable matching outcomes.

As college admissions reforms continue in China and other parts of the world, theoretical, experimental, and empirical analyses of ongoing reforms not only deepen our understanding of the science of market design, but also offer insights into how education and labor market policies should consider the adoption of better mechanisms in their implementation.

ACKNOWLEDGMENTS. We thank Yinghua He, Fuhito Kojima, Nancy Kotzian, Dorothea Kübler, Steve Leider, Sherry Xin Li, Tracy Xiao Liu, Alvin Roth, Utku Unver, Xiaohan Zhong, seminar participants at Stanford University, Tsinghua University, the University of Michigan, WZB and the Tsinghua BEAT Conference, and three anonymous referees for helpful discussions and comments. Chen and Kesten gratefully acknowledge financial support from the National Science Foundation through grants no. SES-0720943 and SES-0962492. Jiang gratefully acknowledges financial support from the National Natural Science Foundation of China through grant no. 71703038.

1. PR Milgrom, *Putting Auction Theory to Work.* (Cambridge University Press, New York), (2004).
2. P Milgrom, Auction market design: Recent innovations. *Annu. Rev. Econ.* **11**, 383–405 (2019).
3. JO Ledyard, K Szakaly-Moore, Designing organizations for trading pollution rights. *J. Econ. Behav. & Organ.* **25**, 167 – 196 (1994).
4. AE Roth, The evolution of labor market for medical interns and residents: A case study in game theory. *J. Polit. Econ.* **92**, 991–1016 (1984).
5. A Roth, New physicians:a natural experiment in market organization. *Science* **250**, 1524–1528 (1990).
6. JH Kagel, AE Roth, The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *Q. J. Econ.* **115**, 201–235 (2000).
7. M Niederle, AE Roth, Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match. *J. Polit. Econ.* **111**, 1342–1352 (2003).
8. Abdulkadiroğlu, T Sönmez, School choice: A mechanism design approach. *Am. Econ. Rev.* **93**, 729–747 (2003).
9. A Abdulkadiroğlu, PA Pathak, AE Roth, The New York City high school match. *Am. Econ. Rev.* **95**, 364–367 (2005).
10. A Abdulkadiroğlu, PA Pathak, AE Roth, T Sönmez, The Boston public school match. *Am. Econ. Rev.* **95**, 368–371 (2005).
11. AE Roth, T Sönmez, MU Ünver, Kidney Exchange. *The Q. J. Econ.* **119**, 457–488 (2004).
12. AE Roth, Experiments in market design in *The Handbook of Experimental Economics*, eds. JH Kagel, AE Roth. (Princeton University Press, Princeton, New Jersey) Vol. 2, pp. 290–346 (2015).
13. T Larroucau, I Rios, Do "short-list" students report truthfully? strategic behavior in the chilean college admissions problem. (2019).
14. Y Chen, O Kesten, Chinese college admissions and school choice reforms: A theoretical analysis. *J. Polit. Econ.* **125**, 99–139 (2017).
15. S Braun, N Dwenger, D Kübler, Telling the truth may not pay off: an empirical study of centralized university admissions in germany. *The BE J. Econ. Analysis & Policy* **10** (2010).
16. A Westkamp, An analysis of the german university admissions system. *Econ. Theory* **53**, 561–589 (2013).
17. S Braun, N Dwenger, D Kübler, A Westkamp, Implementing quotas in university admissions: An experimental analysis. *Games Econ. Behav.* **85**, 232 – 251 (2014).
18. J Grenet, Y He, D Kübler, Decentralizing centralized matching markets: Implications from early offers in university admissions, WZB Discussion Paper SP II 2019-208 (2019).
19. ML Balinski, T Sönmez, A tale of two mechanisms: Student placement. *J. Econ. Theory* **84**, 73–94 (1999).
20. AE Roth, A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the united kingdom. *Am. Econ. Rev.* **81**, 414–440 (1991).
21. HI Ergin, T Sönmez, Games of school choice under the boston mechanism. *J. Public Econ.* **90**, 215 – 237 (2006).
22. Y Chen, T Sönmez, School choice: An experimental study. *J. Econ. Theory* **127**, 202–231 (2006).
23. PA Pathak, T Sönmez, Leveling the playing field: sincere and sophisticated players in the boston mechanism. *Am. Econ. Rev.* **98**, 1636–1652 (2008).
24. H Nie, A game theoretic analysis of college admissions. *China Econ. Q.* **6**, 899–916 (2007).
25. D Gale, LS Shapley, College admissions and the stability of marriage. *Am. Math. Mon.* **69**, 9–15 (1962).
26. PA Pathak, T Sönmez, School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *The Am. Econ. Rev.* **103**, 80–106 (2013).
27. O Kesten, On two competing mechanisms for priority-based allocation problems. *J. Econ. Theory* **127**, 155 – 171 (2006).
28. B Wu, X Zhong, Matching inequality and strategic behavior under the boston mechanism: Evidence from china's college admissions. *Games Econ. Behav.* **123**, 1 – 21 (2020).
29. W Ha, L Kang, Y Song, College matching mechanisms and matching stability: Evidence from a natural experiment in china. *J. Econ. Behav. & Organ.* **175**, 206 – 226 (2020).

30. Y Song, K Tomoeda, X Xia, Sophistication and cautiousness in college applications. Manuscript (2019).
31. B Wu, X Zhong, Matching mechanisms and matching quality: Evidence from a top university in china. *Games Econ. Behav.* **84**, 196 – 215 (2014).
32. L Wei, A design for college and graduate school admission. *China Econ. Q.* **9**, 349–362 (2009).
33. J Pais, Á Pintér, School choice and information: An experimental study on matching mechanisms. *Games Econ. Behav.* **64**, 303 – 328 (2008).
34. CR Featherstone, M Niederle, Boston versus deferred acceptance in an interim setting: An experimental investigation. *Games Econ. Behav.* **100**, 353 – 375 (2016).
35. C Calsamiglia, G Haeringer, F Klijn, Constrained school choice: An experimental study. *Am. Econ. Rev.* **100**, 1860–74 (2010).
36. R Hakimov, D Kübler, Experiments on matching markets: A survey. WZB Discussion Paper (2019).
37. S Mongell, AE Roth, Sorority rush as a two-sided matching mechanism. *Am. Econ. Rev.* **81**, 441–464 (1991).
38. JR Carvalho, T Magnac, College choice and entry exams. Working paper (2010).
39. Y He, Gaming the Boston school choice mechanism in Beijing. Rice University Working Paper (2014).
40. C Calsamiglia, C Fu, M Güell, Structural estimation of a model of school choices: The boston mechanism vs. its alternatives. *J. Polit. Econ.* (2019).
41. N Agarwal, P Somaini, Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica* **86**, 391–444 (2018).
42. M De Haan, PA Gautier, H Oosterbeek, B van der Klaauw, The performance of school assignment mechanisms in practice. CEPR Discussion Paper No. DP10656. (2015).
43. A Abdulkadiroğlu, N Agarwal, PA Pathak, The welfare effects of coordinated assignment: Evidence from the New York City high school match. *Am. Econ. Rev.* **107**, 3635–89 (2017).
44. G Fack, J Grenet, Y He, Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *Am. Econ. Rev.* **109**, 1486–1529 (2019).
45. A Kapor, CA Neilson, SD Zimmerman, Heterogeneous beliefs and school choice mechanisms, (National Bureau of Economic Research), Technical report (2018).
46. A Rees-Jones, S Skowronek, An experimental investigation of preference misrepresentation in the residency match. *Proc. Natl. Acad. Sci.* **115**, 11471–11476 (2018).
47. X Zhong, N Cheng, Y He, Where have all the flowers gone? a game-theoretic analysis of the national college entrance exam preference submission mechanisms. *China Econ. Q.* **3**, 763–778 (2004).
48. JW Lien, J Zheng, X Zhong, Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory. *Exp. Econ.* **19**, 116–150 (2016).
49. M Jiang, When do stable matching mechanisms fail? the role of standardized tests in college admissions, (Working paper), Technical report (2016).
50. S Pan, The instability of matching with overconfident agents. *Games Econ. Behav.* **113**, 396 – 415 (2019).
51. Y Chen, O Kesten, Chinese college admissions and school choice reforms: An experimental study. *Games Econ. Behav.* **115**, 83 – 100 (2019).
52. Q Wu, AE Roth, The lattice of envy-free matchings. *Games Econ. Behav.* **109**, 201 – 211 (2018).
53. Y Kamada, F Kojima, Fair matching under constraints: Theory and applications. Manuscript (2019).